

- Nghiên cứu các tài liệu trong và ngoài nước về lĩnh vực dự báo và dự báo ngắn hạn; - Hội thảo với các chuyên gia am hiểu về lĩnh vực này; - Khảo sát thực tiễn để tính toán, phân tích, đánh giá, tổng kết.

Kết quả nghiên cứu

Đề tài đã đưa ra được các chuyên đề nghiên cứu và một báo cáo tổng hợp, một báo cáo tóm tắt cho đề tài. Báo cáo tổng hợp ngoài phần mở đầu còn gồm: Phần một “Tổng quan về dự báo Thống kê”; Phần hai “Các chỉ tiêu thống kê xã hội và lựa chọn các chỉ tiêu thống kê xã hội cho dự báo”; Phần ba “Các phương pháp dự báo các chỉ tiêu thống kê xã hội”; Phần bốn “Một số phương pháp dự báo cho một số chỉ tiêu thống kê xã hội”; Phần năm “Kết quả dự báo một số chỉ tiêu thống kê xã hội”; và cuối cùng là phần “Kết luận và khuyến nghị”.

PHẦN I

TỔNG QUAN VỀ DỰ BÁO THỐNG KÊ VÀ CÁC PHƯƠNG PHÁP DỰ BÁO CÁC CHỈ TIÊU THỐNG KÊ XÃ HỘI

I. Tổng quan về dự báo thống kê

Mục tiêu của phần này là nhằm tổng quan nội dung của dự báo các chỉ tiêu thống kê xã hội. Tuy nhiên, để dễ theo dõi các kết quả của đề tài, một số vấn đề chung về dự báo được đề cập. Cụ thể là: Khái niệm, định nghĩa dự báo; Khả năng của dự báo và mức độ tin cậy đối với các kết quả dự báo; Các cách dự báo; Đo lường mức độ thành công và thất bại của các dự báo; Khó khăn chủ yếu gặp phải trong quá trình dự báo.

1.1. Khái niệm, định nghĩa dự báo

Theo từ điển bách khoa toàn thư, Dự báo là quá trình thu thập, xử lý số liệu trong quá khứ và hiện tại để xác định xu hướng vận động của các hiện tượng trong tương lai nhờ vào một số mô hình toán học và thống kê. Trong lý thuyết cũng như thực tế của công tác dự báo, dự báo thường được hiểu theo nghĩa là dự đoán các giá trị của một biến nào đó ngoài mẫu số liệu sẵn có, tức là *dự đoán cho các giá trị của biến đó ở những thời gian hoặc không gian khác*. Trong thực tế dự báo là một chủ đề rất rộng.

1.2. Khả năng của dự báo

Do dự báo chỉ đơn giản là một nhận định về tương lai nên người ta có thể dự báo về bất kỳ điều gì, từ chỉ số giá tiêu dùng của tháng tới cho đến tình hình thời tiết ngày mai, từ mực nước biển tính bình quân dâng cao thêm bao nhiêu vào cuối thiên niên kỷ thứ ba, cho tới dân số của trái đất vào ngày hôm đó, và giá trị của chỉ số VN index tại thời điểm đầu năm 2012. Người ta cũng không khẳng định các kết quả dự báo nhất thiết phải là hữu ích. Như vậy việc dự báo các chỉ tiêu thống kê - xã hội cũng không nằm ngoài quan niệm nêu trên.

1.3. Mức độ tin cậy của các dự báo

Mức độ tin cậy của dự báo phụ thuộc vào mức độ tốt xấu của cơ sở xây dựng dự báo. Các dự đoán đơn giản không mang lại độ tin cậy cao, các dự báo dựa trên các cách tiếp cận đã được kiểm định đúng đắn có thể mang lại kết quả tốt hơn. Tuy nhiên, ngay cả khi nó được dựa trên các cách tiếp cận này, nó dường như vẫn chưa thể đảm bảo cho một kết quả sát với thực tế. Tương lai quá bất định chính là khó khăn gặp phải trong quá trình dự báo.

1.4. Các cách dự báo

Có nhiều cách để đưa ra dự báo. Một số cách dự đoán có thể kể ra là: dựa trên các mô hình chính thống tiến hành phân tích thống kê để từ đó tìm ra quy luật của hiện tượng để dự báo; Tiến hành các phân tích thống kê không dựa trên các mô hình tham số; Dựa trên các tính toán không chính thức; Dựa vào phương pháp nội suy giản đơn, tung đồng xu, đoán, và “linh cảm”;... Rất khó để đánh giá tần suất sử dụng của mỗi phương pháp trong thực tế, nhưng hầu hết các phương pháp này đều được thực hiện trong cuộc sống hàng ngày của con người. Tuy nhiên, ở đây sẽ chỉ tập trung vào các cách tiếp cận thống kê chính thống.

Trong thực tiễn công tác dự báo, các nhà khoa học thường sử dụng các phương pháp dự báo như đoán, đánh giá của chuyên gia, ngoại suy, các chỉ số dự báo sớm, các cuộc điều tra, các mô hình chuỗi thời gian và các hệ phương trình kinh tế lượng.

Đ đoán chủ yếu dựa vào may mắn. Đây là phương pháp đòi hỏi ít giả thiết nhất so với các phương pháp khác, nhưng lại không phải là một phương pháp có mức độ tin cậy cao. Thường khi đoán “đúng”, người ta mới công bố kết quả, còn đoán sai lại không được nhắc tới. Đoán thường có mức độ bất định cao nên

không thể đánh giá mức độ sai số. Nếu nhiều người cùng đoán, một số sẽ đúng một cách ngẫu nhiên, nhưng nó khó có thể được coi là một cách tiếp cận khoa học.

Đánh giá chuyên gia là dựa trên ý kiến của các chuyên gia để tiếp cận dự báo. Tuy nhiên, nó ít có giá trị nếu được sử dụng một cách đơn độc. Ngay cả khi những thành công mang tính hệ thống của phương pháp chuyên gia cũng không ai có thể dự đoán liệu chuyên gia có thành công trong lần dự đoán tiếp theo hay không.

Ngoại suy sẽ rất tốt nếu như xu thế vẫn tiếp diễn, nhưng điều này cũng không chắc chắn, bởi thường thấy là xu thế của hiện tượng ở các thời điểm khác nhau thường khác nhau. Ngoài ra, các dự báo cho kết quả sát thực nhất khi chúng dự báo sự thay đổi trong xu thế và điều này thì phương pháp ngoại suy thường không làm được.

Dự báo dựa trên các chỉ số dự báo sớm đòi hỏi một mối quan hệ ổn định giữa các biến dự báo sớm và các biến sẽ diễn ra sau đó. Khi có thông tin về yếu tố đi trước, chẳng hạn có thông tin về đơn đặt hàng khi tiến hành sản xuất, thì khi đấy các chỉ số dự báo sớm sẽ hữu ích. Tuy nhiên, nếu không có các thông tin này, việc dự báo dựa vào các chỉ số dự báo sớm sẽ cho thông tin sai lầm.

Các mô hình chuỗi thời gian mô tả đặc điểm của dữ liệu trong quá khứ là phương pháp dự báo khá phổ biến, đây là phương pháp được sử dụng bên cạnh phương pháp hệ phương trình kinh tế lượng (cụ thể là ở dạng đa biến).

Hệ phương trình kinh tế lượng là công cụ chính để dự báo kinh tế, xã hội. Những phương trình này nhằm tìm cách mô hình hóa hành vi của các nhóm tác nhân kinh tế, xã hội có thể quan sát được (người tiêu dùng, nhà sản xuất, người lao động, nhà đầu tư,...) với giả định về mức độ hợp lý trong hành động của các tác nhân này ở một mức độ vừa phải dựa theo kinh nghiệm lịch sử. Các mô hình kinh tế lượng và chuỗi thời gian là những phương pháp dự báo chính trong kinh tế, xã hội nhưng “các đánh giá”, “các chỉ số dự báo sớm”, và thậm chí cả những tiên đoán cũng có thể thay đổi các kết quả dự báo.

1.5. Đo lường mức độ thành công và thất bại của dự báo

Một dự báo có thể được đánh giá là thành công nếu nó gần sát với kết cục thực tế. Nhưng đánh giá này còn tùy thuộc vào việc đo lường thế nào là gần. Đối với người không làm chuyên môn, một dự báo chi tiết mà không chính xác có thể được xem là thất bại. Mặt khác, một dự báo rất chính xác nhưng không cụ thể cũng vậy. Các chuyên gia tán thành quan điểm – “tiêu

chuẩn vàng” cho một dự báo thành công là kết quả dự báo phải chính xác, chi tiết và kịp thời.

Trong đánh giá kết quả dự báo, còn dùng hai khái niệm khác là “Dự báo không chệch” và dự báo có “phương sai nhỏ”. Khái niệm dự báo “không chệch” đồng nghĩa với kết quả dự báo có giá trị trung tâm giống như kết cục được dùng trong cách phân tích kỹ thuật để đo lường độ chính xác, còn khái niệm phương sai nhỏ, chỉ một phạm vi tương đối hẹp xung quanh kết cục đo lường mức độ chi tiết.

Ngoài vấn đề về độ chệch và phương sai, các dự báo điểm thường được đánh giá trên nhiều tiêu chí khác như sử dụng thông tin hiệu quả. Các dự báo cũng thường bao gồm cả khoảng dự báo, và đôi khi là mật độ hoàn chỉnh của kết quả dự báo. Trong trường hợp như vậy, dự báo được coi là được xác định một cách đầy đủ. Do trong thực tế hiếm khi xảy ra trường hợp chỉ cần một phương pháp dự báo duy nhất có thể dự báo cho mọi hiện tượng được quan tâm nên các phương pháp dự báo khác nhau thường được đưa ra để so sánh với nhau sau khi đã có kết quả thực tế.

1.6. Phương pháp phân tích đặc điểm của các phương pháp dự báo

Đặc điểm của các phương pháp dự báo có thể được tìm hiểu thông qua thực nghiệm và các ví dụ mô phỏng. Người ta thường sử dụng giải tích toán và các phương pháp số học hoặc sử dụng máy tính để làm việc này. Người ta cũng có thể thử nghiệm với các ý tưởng được bàn luận ở trên để đo lường mức độ thành công hoặc thất bại của các dự báo bằng cách kiểm định các kết hợp của các mô hình dự báo về tính chất bao hàm. Các phương pháp dự báo có thể được so sánh bằng phương pháp Monte Carlo (hay mô phỏng ngẫu nhiên), trong đó người ta sẽ tạo ra dữ liệu ảo và theo đó các mô hình được so sánh với nhau trong các phép thử lặp lại, để từ đó tính ra được những phương pháp này hoạt động ra sao trong một môi trường có kiểm soát về lựa chọn.

So sánh các phương pháp dự báo bằng thực nghiệm thường dựa vào quan sát các kết quả của các phương pháp trong nhiều chuỗi thời gian. Bởi vì quá trình tạo dữ liệu không nằm trong sự kiểm soát của người nghiên cứu nên người ta biết về nó chưa thật sâu sắc. Kết quả của các phương pháp dự báo dựa vào bất kỳ một chuỗi thông tin nào đều có thể phụ thuộc vào đặc điểm riêng của chúng, và do vậy sẽ hạn chế khả năng ứng dụng mang tính khái quát.

1.7. Các quy luật đặc biệt của chuỗi số liệu theo thời gian có ảnh hưởng nhiều nhất đến kết quả dự báo

Nhiều chuỗi thời gian có những đặc điểm rất đặc thù như tính mùa vụ, dao động theo chu kỳ, tăng trưởng xu thế, tính phụ thuộc kế tiếp nhau, thay đổi mức độ dao động,... Việc không xem xét những đặc điểm này (ví dụ tính mùa) có thể dẫn đến kết quả dự báo không sát thực (ví dụ điểm ngoặt hay một xu thế) nếu như những đặc điểm này có liên hệ với nhau. Vì vậy, khi tiến hành dự báo có thể cần phải đặc biệt chú ý tới chúng. Với lý do đó, các mô hình được xây dựng luôn cố gắng nắm bắt những đặc điểm này, và sử dụng các cách tiếp cận khác nhau để xử lý vấn đề.

1.8. Các khó khăn chủ yếu gặp phải trong quá trình dự báo

Một trong những khó khăn chính khi dự báo hiện tượng liên tục phát triển theo thời gian là nó thỉnh thoảng gặp phải những trục trặc, và đôi khi có thể là các cú sốc lớn không thể dự đoán trước được. Do vậy, những mối quan hệ ổn định trước đây giữa các biến sử dụng cho dự báo có thể bị thay đổi, và nếu dùng chúng để dự báo thì có thể dẫn tới những sai lầm dự báo lớn và kéo dài.

1.9. Các phương án khắc phục khó khăn

Các khoảng dự báo cho biết mức độ bất định của dự báo, nhưng nó chỉ có thể phản ánh sự bất định đã biết - rút ra từ ước lượng mô hình, giả định các cú sốc tương lai giống với các cú sốc trong quá khứ - trong khi đó các thành phần xác định có thể thay đổi ngoài dự kiến. Như đã đề cập, do không biết những yếu tố không biết nên rất khó có thể tính đến yếu tố bất định chưa biết này.

Một ví dụ đơn giản có thể làm sáng tỏ vấn đề này. Giả sử một biến cần dự báo trong thực tế được tạo ra bằng phương trình sau:

$$Y_t = \delta + \Phi Y_{t-1} + a_t \quad (1)$$

Trong đó $\{a_t\}$ là một chuỗi các biến ngẫu nhiên phân phối chuẩn và độc lập với trung bình bằng 0 và phương sai σ_a^2 , được viết là $\{a_t\} \sim IN[0, \sigma_a^2]$. Tại thời kỳ T, người ta biết giá trị biến này sẽ nhận trong thời kỳ tiếp theo (T+1) sẽ là $y_{T+1} = \delta + \Phi y_T$ cộng với giá trị thực tế của cú sốc ngẫu nhiên. Ở đây, chúng ta biết phân phối của cú sốc đó, do vậy chúng ta biết y_{T+1} sẽ phân phối chuẩn xung quanh giá trị bình quân là $\delta + \Phi y_T$ với phương sai là σ_a^2 . Do vậy, chúng ta có

thể xây dựng khoảng tin cậy cho giá trị dự báo trung tâm ($y_{T+1/T} = \delta + \Phi y_T$). Dưới dạng công thức xác suất, kết cục sẽ nằm trong khoảng tin cậy sau:

$$P_r \left(Y_{T+1} \in \left\{ y_{T+1/T} \pm z_{\frac{\alpha}{2}} \sigma_a \right\} \right) = 1 - \alpha \quad ,$$

Trong đó $z_{\frac{\alpha}{2}}$ là giá trị để cho xác suất mà biến phân phối chuẩn nhỏ hơn nó bằng $\frac{\alpha}{2}$. Khoảng này cho biết nếu có K trường hợp $\{Y_{T+1}\}$ xảy ra, khi đó sẽ có $(1-\alpha) \times K$ trường hợp trong số đó, giá trị thực tế xảy ra sẽ nằm trong khoảng $\pm Z_{\frac{\alpha}{2}} \sigma_a$ quanh giá trị dự báo. Do mức độ bất định trong dự báo điểm được biết chính xác, vậy vấn đề nằm ở đâu?

Thứ nhất, trong thực tế, không bao giờ biết giá trị của các tham số trong mô hình, $\{\delta, \phi, \sigma_a^2\}$, do vậy chúng sẽ được thay thế bằng cách hệ số ước lượng. Tuy nhiên, tính bất định tăng thêm ở bước này là một phần trong tính bất định đã được biết: do biết rằng sẽ xuất hiện bất định này nên có thể tính chúng vào trong dự báo. Thứ hai, có thể không biết chắc chắn về dạng mô hình vì vậy có thể có những phản ứng trễ xa hơn so với những gì được đưa vào trong (1), hoặc có thể cần sử dụng dạng log thay vì giá trị tuyệt đối,... Vấn đề bất định có thể nảy sinh từ việc chưa biết dạng chỉ định mô hình khó khắc phục hơn. Vấn đề thứ ba là trường hợp tương lai không còn giống với quá khứ, chẳng hạn như $\{a_t\} \sim IN[\mu, \tilde{\sigma}_a^2]$ với $t > T$, trong đó hoặc $\mu \neq 0$ và/hoặc $\tilde{\sigma}_a^2 \neq \sigma_a^2$. Lưu ý rằng $\mu \neq 0$ tương đương với việc thay đổi δ thành $\delta + \mu$. Tất nhiên, không có lý do xác đáng nào để giả định rằng dạng phân phối của cú sốc này sẽ không thay đổi. Thay đổi giá trị bình quân sẽ ảnh hưởng tới tính chính xác của dự báo, và thay đổi phương sai của nhiễu có nghĩa là sẽ hoặc ước lượng quá cao hoặc ước lượng quá thấp mức độ bất định xung quanh dự báo điểm. Thay đổi phân phối sai số sẽ gây ra tính không chính xác trong dự báo mật độ.

Để làm sáng tỏ những vấn đề này, giả sử rằng trong (1), hệ số chặn δ thay đổi thành $\delta + \mu$ tại $T+1$, và để đơn giản hóa, $\phi = 0$, như vậy quá trình này sẽ thành:

$$Y_t = \delta + \mu 1(t > T) + a_t \quad (2)$$

Trong đó $1(t > T)$ nhận giá trị bằng 1 khi $t > T$ và không trong trường hợp còn lại. Thời kỳ dự báo $T+2$ tại thời điểm $T+1$ mà không nhận ra giá trị bình quân đã thay đổi, tính trung bình chúng ta sẽ mắc sai lầm bằng μ trong khi đó

nếu chúng ta sử dụng “chỉ số dự báo bước ngẫu nhiên”, tức là $\hat{y}_{T+2/T+1} = y_{T+1}$ thì sai số bình quân sẽ bằng không! Ngược lại, nếu chúng ta đánh giá lại dự báo vào T+2 đúng bằng sai lầm mà chúng ta mắc phải khi dự báo thời kỳ T+1 thì sai số bình quân sẽ lại bằng không. Nhưng cái giá phải trả là tăng được mức độ chính xác sẽ làm giảm mức độ chi tiết. Đơn giản như sau: khoảng tin cậy sẽ bao quát kết cục thực sự diễn ra trong $(1-\alpha)\times 100\%$ số lần sẽ xấp xỉ $\left\{ \hat{y}_{T+2/T+1} \pm z_{\frac{\alpha}{2}} \sigma_{\alpha} \right\}$, nó lớn gấp hai lần so với dự báo trước 1 thời kỳ khi sử dụng

mô hình hiệu chỉnh và không có sự thay đổi của giá trị bình quân.

II. Các phương pháp dự báo các chỉ tiêu thống kê xã hội

2.1. Các loại dự báo

Trên các góc độ khác nhau, phân dự báo ra thành các loại khác nhau. Dựa vào thời kỳ nghiên cứu và tầm dự đoán, chia dự báo ra thành hai loại. Loại thứ nhất là dự báo ngắn hạn và loại thứ hai là dự báo dài hạn.

Dự báo ngắn hạn thường được coi là dự báo cho các chuỗi số liệu có thời kỳ nghiên cứu dưới một năm (tuần, tháng, quý,...) hoặc có tầm dự báo từ 1 đến 5 thời kỳ tiếp theo. Tuy nhiên, các quy định này cũng chỉ là tương đối, vì vậy không câu nệ khi có người gọi khác đi so với cách xác định loại dự báo trên.

Trên một góc độ khác, có thể chia các phương pháp dự báo thành 2 nhóm chính: (i) Nhóm thứ nhất, dự báo trên cơ sở giả định về sự phụ thuộc của biến dự báo vào số liệu quá khứ với các phương pháp thường sử dụng như ngoại suy dựa trên phát hiện xu hướng theo thời gian, AR, MA, ARMA, ARIMA, ARDL, v.v...; (ii) Nhóm thứ hai, dự báo dựa trên cơ sở sử dụng các mô hình mô tả các quan hệ cơ cấu.

Nhóm phương pháp thứ nhất đơn giản hơn về kỹ thuật tính toán cũng như yêu cầu về số liệu song không cho phép dự báo được tác động của các thay đổi chính sách tạo nên sự thay đổi về các quan hệ cơ cấu.

2.2. Các phương pháp dự báo

Nhiều chỉ tiêu thống kê vừa có đặc trưng của các chỉ tiêu thống kê kinh tế và vừa có các đặc trưng của các chỉ tiêu thống kê xã hội. Vì vậy, nhiều phương pháp dự báo không những được áp dụng trong lĩnh vực thống kê kinh tế mà còn được áp dụng cả cho lĩnh vực thống kê xã hội và ngược lại. Dưới đây là các phương pháp dự báo cơ bản được áp dụng cho dự báo các chỉ tiêu thống

kê nói chung. Trong phần tổng quan ở trên đã đi qua các phương pháp dự báo chính sau đây: 1. Đoán, “quy tắc ngón tay cái” hoặc “các mô hình không chính thức”; 2. Đánh giá của chuyên gia; 3. Ngoại suy; 4. Các chỉ số dự báo sớm; 5. Các cuộc điều tra; 6. Các mô hình chuỗi thời gian và; 7. Các hệ kinh tế lượng. Trong mục này sẽ tập trung vào trình bày về các phương pháp dự báo và một số nội dung cơ bản của dự báo theo phương pháp kinh tế lượng.

2.2.1. Dự báo điểm và dự báo khoảng

Giả sử cần dự báo véc tơ các biến y (có thể gồm một hệ thống các chỉ tiêu thống kê xã hội nào đó), bài toán dự báo điểm hình là bài toán dự đoán các giá trị đối với y tại thời gian tương lai $T+h$, khi biết trước T quan sát y_1, y_2, \dots, y_T , và có thể biết trước các quan sát của một vài biến khác. Thời gian T thường được lấy là hiện tại và khoảng thời gian dương h được gọi là tầm dự báo. Dự báo điểm được ký hiệu là:

$$y_{T+h} \quad (3)$$

Ký hiệu trên biểu thị giá trị dự đoán của y tại thời điểm $T+h$. Trong chừng mực mà các giá trị đúng của các biến tại thời gian này, y_{T+h} , được xác định theo một phân phối xác suất thì dự báo điểm (3) nói chung được lấy là giá trị kỳ vọng của phân phối như đã ước lượng tại thời gian T từ các số liệu y_1, y_2, \dots, y_T . Giá trị kỳ vọng này có thể đặt trong ngoặc nếu ta dự báo bằng khoảng dự báo với mức độ (khoảng) tin cậy nào đó - thí dụ, khoảng tin cậy 90%:

$$P(y_{T+h} \in \bar{y}_{T+h} \pm 0,9) \quad (4)$$

được định nghĩa bằng:

$$P(\bar{y}_{T+h} - 0,9 \leq y_{T+h} \leq \bar{y}_{T+h} + 0,9) \quad (5)$$

Một loại dự báo khác dựa trên phân phối xác suất của y_{T+h} là dự báo xác suất, nó nêu ra xác suất của một phát biểu nào đó về giá trị tương lai của biến.

$$P(y_{T+h} \geq y_0) = \alpha \quad (6)$$

Ở đây y_0 được chỉ định và α là xác suất xảy ra của sự kiện. Một thí dụ là dự báo thời tiết dựa vào xác suất mưa. Loại dự báo này không được sử dụng trong kinh tế cũng như xã hội học rộng rãi như các dự báo điểm và dự báo khoảng, nhưng chúng có thể được sử dụng rộng rãi hơn trong tương lai khi phương pháp luận để xây dựng và đánh giá chúng được phát triển.

Có một vài cách tiếp cận khác nhau đối với dự báo. Mục này giới thiệu những cách tiếp cận khác với cách tiếp cận kinh tế lượng, và mục tiếp theo trình bày cách tiếp cận kinh tế lượng. Cách tiếp cận lâu đời nhất của dự báo là cách tiếp cận theo ý kiến chuyên gia, trên cơ sở những đánh giá của các chuyên gia am hiểu về hiện tượng được nghiên cứu. Một ví dụ điển hình của cách tiếp cận này là dựa vào kết quả của các cuộc điều tra về dự tính, ví dụ như những cuộc điều tra về dự tính đầu tư tư bản, dự tính về tiêu dùng, trong đó người ta yêu cầu chính những người ra quyết định dự báo những hành động tương lai của họ. Nhìn chung, những nhân tố liên quan đến dự báo như ngân sách và lượng hàng hoá bán được, và những điều kiện tín dụng, thường không được xét đến một cách cụ thể trong cách tiếp cận dự báo này mà thường chỉ được cân nhắc và đánh giá chủ quan bởi chuyên gia.

Một biến thể hiện đại của phương pháp tiếp cận dự báo chuyên gia là phương pháp Delphi. Ở phương pháp này, người ta tập hợp đánh giá của một nhóm chuyên gia để đưa ra kết quả dự báo. Mỗi chuyên gia được hỏi ý kiến và rồi những dự báo của họ được trình bày dưới dạng thống kê tóm tắt cho mọi người. Việc trình bày những ý kiến trả lời này thường được tiến hành giấu tên không có sự tiếp xúc mặt đối mặt (chẳng hạn thông qua câu hỏi bằng thư) để tránh những vấn đề tương tác trong nhóm nhỏ có thể tạo nên những sai lệch nhất định trong kết quả. Sau đó người ta yêu cầu các chuyên gia duyệt xét lại những dự báo của họ trên cơ sở tóm lược của tất cả các dự báo và có thể những thông tin bổ sung thêm. Quá trình này được lặp lại cho đến khi nhóm chuyên gia đạt được sự thống nhất ý kiến.

Một cách tiếp cận hình thức hơn là dự báo dựa trên tính y: dựa trên giả thiết hệ thống có một động lượng nào đó, với tương lai là bản sao của quá khứ. Loại đơn giản nhất là dự báo theo nguyên trạng: dự báo giá trị hiện tại của biến số sẽ tiếp tục được chuyển qua tương lai. Giả sử cần dự báo một biến đơn theo cách tiếp cận dự báo theo nguyên trạng và dự báo này là dự báo điểm. Ta sẽ có:

$$y_{T+1} = y_T \quad (7)$$

Cách tiếp cận dự báo này còn được gọi là dự báo “ngây thơ I”. Một thí dụ về loại dự báo này là ta dự báo rằng thời tiết ngày mai trùng với thời tiết hôm nay. Dự báo này có xu hướng đúng với một tỷ lệ thời gian rất cao.

Một loại đơn giản khác của dự báo dựa trên tính y là dự đoán có sự thay đổi như nhau từ thời kỳ này sang thời kỳ tiếp theo, tức là:

$$\hat{y}_{T+1} - y_T = y_T - y_{T-1}, \text{ hay } \hat{y}_{T+1} = 2y_T - y_{T-1} \quad (8)$$

Cách tiếp cận dự báo này còn được gọi là dự báo “ngây thơ II”. Một dạng dự báo dựa trên tính ì khác là dự đoán có sự thay đổi tỷ lệ như nhau, cụ thể là:

$$(\hat{y}_{T+1} - y_T) / y_T = (y_T - y_{T-1}) / y_{T-1}, \text{ hay } \hat{y}_{T+1} = y_T + \frac{y_T}{y_{T-1}}(y_T - y_{T-1}) \quad (9)$$

Một dạng tổng quát của dự báo dựa trên tính ì là mô hình tự hồi quy, với phương trình dự báo sau:

$$\hat{y}_{T+1} = \sum_{j=0}^{\infty} a_j y_{T-j} \quad (10)$$

Ở đây, giá trị dự báo thu được từ một tổ hợp tuyến tính có trọng số của tất cả các giá trị quá khứ của biến. Các hệ số a_j có thể được chỉ định tiên nghiệm như trong (7) và (8) hoặc chúng có thể được ước lượng bằng thống kê.

Một cách tiếp cận dự báo khác là ngoại suy xu thế dựa trên cơ sở những hàm đơn giản của thời gian, thí dụ hàm xu thế tuyến tính:

$$y_t = a + bt \quad (11)$$

Dựa vào hàm này, giá trị dự báo tại T+1 là:

$$\hat{y}_{T+1} = a + b(T+1) \quad (12)$$

Ở đây a và b hoặc là được mặc nhiên công nhận hoặc là được ước lượng bằng thống kê. Thực tế, mô hình này là một trường hợp đặc biệt của dự báo dựa trên tính ì (8) trên cơ sở lượng thay đổi tuyệt đối không đổi từ thời kỳ này sang thời kỳ khác, ở đây:

$$\hat{y}_{T+1} - y_T = y_T - y_{T-1} = b \quad (13)$$

Tương tự, xu thế hàm mũ

$$y_t = Ae^{\alpha t}, \quad (14)$$

Cho kết quả dự báo:

$$\hat{y}_{T+1} = Ae^{\alpha(T+1)}, \quad (15)$$

là một trường hợp đặc biệt của (9), trong đó dự đoán trên cơ sở thay đổi tương đối không đổi, bởi:

$$\frac{(\hat{y}_{T+1} - y_T)}{y_T} = \frac{(y_T - y_{T-1})}{y_{T-1}} = e^{\alpha} - 1 \quad (16)$$

Lấy logarit của (14) có:

$$\ln y_t = \ln A + \alpha t \quad (17)$$

do đó dự báo tại thời gian $T+h$ có thể được viết

$$\ln \hat{y}_{T+h} = \ln A + \alpha(T+h), \quad (18)$$

Ở đây α là tốc độ thay đổi của y . Đây là mô hình được dùng để dự báo tăng trưởng. Thí dụ, khi sử dụng để dự báo ngược theo thời gian, mô hình dự đoán rằng $y = 1$ khi

$$\ln \hat{y}_{T+h} = \ln A + \alpha(T+h) = 0, \quad (19)$$

nghĩa là khi tầm h bằng h^* , ở đây

$$T + h^* = -\ln A/\alpha \quad (20)$$

thu được từ ước lượng A và α .

Còn một cách tiếp cận khác đối với dự báo là phương pháp các chỉ báo sớm. Để sử dụng cách tiếp cận này, một dự báo đối với y dựa trên cơ sở một biến x có liên quan, gọi là chỉ báo sớm, trong đó y tại thời gian t phụ thuộc vào x tại thời gian trước đó $T-\theta$. Như vậy,

$$y_t = f(x_{t-\theta}) \quad (21)$$

Như vậy giá trị dự đoán của y tại thời gian $T+h$ là

$$\hat{y}_{T+h} = f(x_{T+h-\theta}) \quad (22)$$

Thực tế, phương pháp này thường không được áp dụng để dự báo mức của biến mà là cho dự báo tốc độ thay đổi của nó. Mô hình:

$$\Delta y_{T+h} = g(\Delta x_{T+h-\theta}) \quad (23)$$

được sử dụng để dự đoán các điểm ngoặt, thí dụ như dự báo sự suy sụp dựa vào một biến báo hiệu sự suy sụp sau đó của một biến khác. Các chỉ báo dẫn đường được chọn trên cơ sở ghi chép trong dự đoán các điểm ngoặt quá khứ (có vai trò dẫn đường). Thí dụ, cung tiền tệ, trong đó một thay đổi trong tốc độ tăng trưởng cung tiền tệ nói chung dẫn đến một thay đổi trong tốc độ tăng trưởng của thu nhập quốc gia trong 9 đến 12 tháng.

2.3. Cách tiếp cận kinh tế lượng đối với dự báo

2.3.1. Các dự báo ngắn hạn

Cách tiếp cận kinh tế lượng bằng mô hình tuyến tính để dự báo dựa vào các phương trình dạng rút gọn sau:

$$\vec{y}_t = \vec{y}_{t-1} \Pi_1 + \vec{z}_t \Pi_2 + \vec{u}_t \quad (24)$$

Ở đây, \vec{y}_t là véc tơ của g biến nội sinh cần dự báo; \vec{z}_t là véc tơ của k biến ngoại sinh; \vec{y}_{t-1} , các biến nội sinh trễ, và \vec{z}_t đều là các biến xác định trước; và \vec{u}_t là véc tơ g số hạng nhiễu ngẫu nhiên. Các ma trận hệ số được ước lượng là Π_1 và Π_2 , các ma trận số với g2 và kg phần tử. Ma trận hệ số dạng rút gọn ước lượng này có thể thu được trực tiếp từ một ước lượng của dạng rút gọn hoặc gián tiếp từ dạng cấu trúc ước lượng được. Khoảng thời gian gắn với trễ trong (24) phụ thuộc vào hệ thống được xét, phụ thuộc vào việc khoảng dự báo “mở rộng ra” nhanh thế nào.

Một cách lý tưởng nêu mô hình (24) được dựa trên lý thuyết nào đó và được biểu hiện bằng các phương trình dạng cấu trúc. Đôi khi, để chọn mô hình, đặc biệt là để chọn các biến ngoại sinh người ta sử dụng các số liệu (thí dụ, sử dụng số liệu và phân tích tương quan giữa các biến để chọn biến ngoại sinh). Tuy nhiên, cách tiếp cận này cũng có thể gặp phải trường hợp giữa các biến có tương quan giả tạo (thí dụ, tương quan của tỷ suất sinh với số lượng cò, hoặc tương quan giữa các biến có cùng bản chất chẳng hạn như biến đầu tư với biến tiết kiệm có cùng một bản chất). Cách tiếp cận như thế có thể cho các ước lượng ăn khớp tốt, bao gồm các giá trị R^2 cao, nhưng không cho kết quả dự báo tốt. Khi đã chọn được mô hình kinh tế lượng, được thể hiện bằng các phương trình dạng rút gọn, phương trình dự báo ngắn hạn cho kỳ tiếp theo sẽ là:

$$\hat{y}_{T+1} = y_T \hat{\Pi}_1 + \hat{z}_{T+1} \hat{\Pi}_2 + \hat{u}_{T+1}. \quad (25)$$

Trong công thức dự báo này, giá trị của các biến nội sinh ở kỳ sau bao gồm hai thành phần: thành phần hệ thống và thành phần đánh giá.

Thành phần hệ thống thứ nhất là $\vec{y}_t \hat{\Pi}_1$. Nó chỉ sự phụ thuộc vào giá trị hiện tại của biến nội sinh, được lấy trọng số là các hệ số ước lượng được trong $\hat{\Pi}_1$. Số hạng này bao hàm sự phụ thuộc hệ thống của mỗi biến nội sinh vào các giá trị trước của chúng do hiện tượng tự tương quan hoặc hiện tượng ảnh trễ gây ra. Trong nhiều trường hợp, chỉ trên cơ sở các giá trị trễ đã có thể thu được kết quả dự báo rất tốt, như trong trường hợp dự báo dựa trên tính ý. Thành phần hệ thống thứ hai trong (25) là $\vec{z}_{T+1} \hat{\Pi}_2$. Nó là kết quả dự đoán dựa trên giá trị tương lai của các biến ngoại sinh \vec{z}_{T+1} và các hệ số ước lượng được trong

$\hat{\Pi}_2$. Thành phần này phản ánh sự phụ thuộc của các biến nội sinh vào các biến ngoại sinh của mô hình. Vì \vec{z}_T là các biến ngoại sinh, nên bản thân chúng được xác định trên cơ sở các nhân tố khác không được đưa ra trong mô hình kinh tế lượng. Vì vậy, hợp lý hơn cả là các biến này phải được dự báo trên cơ sở các nhân tố khác với các nhân tố trong chính mô hình này, tức là chính \vec{z}_{T+1} lại là kết quả dự báo từ một mô hình kinh tế lượng khác. Thí dụ, một số công ty lớn dự báo lượng hàng bán được, việc làm, v.v... của họ trên cơ sở một mô hình kinh tế lượng đặc biệt cho công ty hoặc ngành của họ. Một mô hình như thế thường coi các biến kinh tế vĩ mô chủ yếu như thu nhập cá nhân hoặc chi tiêu đầu tư là biến ngoại sinh. Do đó, để dự đoán lượng hàng sẽ bán được cần tiến hành dự báo các biến kinh tế vĩ mô chủ yếu trên và như vậy các biến ngoại sinh này trở thành biến nội sinh của các mô hình kinh tế lượng vĩ mô khác. Nhiều khi, bản thân một số biến ngoại sinh nào đó của mô hình kinh tế lượng vĩ mô được dự đoán trên cơ sở những cuộc điều tra về dự tính đối với hành vi tiêu dùng các quyết định chi tiêu tư bản. Ngoài các output của mô hình kinh tế lượng khác và các điều tra về dự tính, các giá trị tương lai của các biến ngoại sinh \vec{z}_{T+1} đôi khi thu được trên cơ sở ý kiến chuyên gia hoặc ngoại suy các xu thế quá khứ của các biến này. Thành phần thứ ba trong (25) là thành phần đánh giá \vec{u}_{T+1} , được gọi là nhân tố bổ sung có thể diễn giải như là ước lượng của các giá trị tương lai của nhiễu (thành phần ngẫu nhiên), hay một cách khác đi, như là những điều chỉnh của các hệ số chặn trong mỗi phương trình dạng rút gọn. Các nhân tố bổ sung này làm dự báo kinh tế lượng hoàn chỉnh hơn : thành phần thứ nhất trong (25) thu tóm ảnh hưởng của các biến nội sinh quá khứ; thành phần thứ hai thu tóm ảnh hưởng của tất cả các biến ngoại sinh được đưa vào mô hình; thành phần thứ ba, các nhân tố bổ sung, thu tóm ảnh hưởng của tất cả các nhân tố khác, kể cả các biến bị bỏ sót trong mô hình. Các nhân tố bổ sung dựa trên cơ sở những đánh giá về các nhân tố không được chứa tường minh trong mô hình. Thí dụ, trong mô hình kinh tế lượng vĩ mô có thể không tính đến các hoạt động đình công, nhưng nếu đình công có thể nảy sinh trong kỳ dự báo thì những dự báo về sản xuất nên được điều chỉnh xuống một cách thích hợp. Nhiều nhân tố khác có thể không được đưa vào trong mô hình vì chúng hiếm xảy ra hoặc vì khó thu được các số liệu, nhưng điều đó không có nghĩa là buộc phải bỏ qua chúng trong quá trình dự báo. Thực vậy, thật là thiếu sót nếu không xem xét thích đáng các nhân tố này chỉ vì chúng không có mặt trong mô hình. Theo nghĩa đó, dự báo bằng một mô hình kinh tế lượng không

đơn giản là một bài tập máy móc mà là sự pha trộn những xem xét khách quan và chủ quan.

Dự báo kinh tế lượng trong (25) được gọi là dự báo ex-ante (có trước-ngoại suy) bởi vì nó đúng là một dự báo trước khi sự kiện xảy ra. Trái lại, dự báo ex-post (có sau-nội suy), thực hiện sau sự kiện, thay các giá trị dự đoán của các biến ngoại sinh \vec{z}_{T+1} bằng các giá trị thực \vec{z}_{T+1} của nó và thay các nhân tố bổ sung u_{T+1} bằng giá trị kỳ vọng zero của các số hạng nhiễu ngẫu nhiên. Như vậy dự báo ex-post là: $(z_{T+1} - \widehat{z}_{T+1})$

$$\hat{y}_{T+1} = y_T \hat{\Pi}_1 + \hat{z}_{T+1} \hat{\Pi}_2 \quad (26)$$

Quan hệ giữa dự báo ex-ante y_{T+1} và dự báo ex-post y_{T+1} là:

$$\hat{y}_{T+1} = \hat{y}_{T+1} + (z_{T+1} - \hat{z}_{T+1}) \hat{\Pi}_2 - \hat{u}_{T+1} \quad (27)$$

nên \hat{y}_{T+1} có thể thu được từ \hat{y}_{T+1} sau khi quan sát z_{T+1} bằng cách điều chỉnh cả đối với sai số trong dự đoán các biến ngoại sinh $(z_{T+1} - \hat{z}_{T+1})$ lẫn các nhân tố bổ sung \hat{u}_{T+1} .

Trở lại dự báo tất định ex-ante (25), các cách tiếp cận khác nhau đối với dự báo giới thiệu trong mục trước tất cả có thể diễn giải theo thuật ngữ của dự báo kinh tế lượng. Dự báo theo ý kiến chuyên gia là trường hợp đặc biệt trong đó không có phần hệ thống, do đó dự báo có thể được biểu thị như dự báo thuần túy đánh giá:

$$\hat{y}_{T+1} = \hat{u}_{T+1} \quad (28)$$

Ở đây $\hat{\Pi}_1 = 0$ và $\hat{\Pi}_2 = 0$. trong dự báo dựa trên tính ý thành phần hệ thống thứ nhất $y_T \hat{\Pi}_1$ được nhấn mạnh, vì nó phản ánh sự phụ thuộc vào các giá trị quá khứ của cùng các biến đó. Thí dụ, dự báo theo nguyên trạng:

$$\hat{y}_{T+1} = y_t \quad (29)$$

trùng ứng với trường hợp $\hat{\Pi}_1 = I$ và $\hat{\Pi}_2 = 0$, $\hat{u}_{T+1} = 0$. Các dự báo dựa trên tính ý khác, như mô hình tự hồi quy có thể được biểu diễn sử dụng (25) đơn giản bằng cách thêm các số hạng đối với y_{T+2} , y_{T+3} , v.v... Ngoại suy xu thế có thể được giải thích theo các thuật ngữ của (25) hoặc bởi các biến nội sinh trễ hoặc bằng việc đưa thời gian vào như một biến ngoại sinh. Các chỉ báo dẫn đường có thể được biểu diễn theo cách tiếp cận kinh tế lượng sử dụng các biến

ngoại sinh trễ. Các biến dự tính có thể được đưa trực tiếp vào mô hình, và việc đưa chúng vào thường dẫn tới những dự báo được cải thiện hơn.

Thực tế, các dự báo kinh tế lượng như trong (25) có một số ưu việt so với các cách tiếp cận khác. Thứ nhất, tiếp cận kinh tế lượng cho ta một cấu trúc hữu ích để xem xét các nhân tố khác nhau có thể, như các giá trị quá khứ của các biến dự đoán, các giá trị của các biến liên quan và các nhân tố khác. Thứ hai, nó đủ rộng để cho phép xử lý nhiều xem xét khác nhau, kể cả việc tổng hợp các nhân tố hệ thống và đánh giá khác nhau. Thứ ba, nó dẫn tới dự báo các biến có liên quan tương hợp với nhau, vì chúng tất cả phải thoả mãn các đòi hỏi của mô hình, đặc biệt sự nhận diện nó. Thứ tư, nó dẫn tới các dự báo có điều kiện trên các giá trị dự đoán của các biến ngoại sinh tương lai \hat{z}_{T+1} , các nhân tố bổ sung \hat{u}_{T+1} , các ma trận hệ số $\hat{\Pi}_1$ và $\hat{\Pi}_2$ và các giá trị hiện tại của các biến nội sinh y_T . Do đó có thể phân tích tầm quan trọng tương đối của mỗi yếu tố này của dự báo và kiểm định độ nhạy của dự báo trước những thay đổi trong mỗi yếu tố, đặc biệt khi có thêm số liệu mới. Thứ năm, có thể tái tạo một dự báo liên quan – dự báo ex-post (26). Thứ sáu, và có lẽ quan trọng nhất, nó có lịch sử tốt về độ chính xác so với các cách tiếp cận khác mà mỗi trong chúng có thể diễn giải như sự nhấn mạnh một khía cạnh của dự báo kinh tế lượng nhưng không kể đến các khía cạnh khác. Thực tế, những dự báo tốt nhất nói chung kết hợp một mô hình kinh tế lượng, kể cả các chỉ báo dẫn đường, các dữ liệu dự tính và phân tích chuỗi thời gian với các nhân tố đánh giá, biểu thị bởi các nhân tố bổ sung.

2.3.2. Các dự báo dài hạn

Các dự báo dài hạn là các dự báo mà đối với chúng tầm dự báo, h , lớn hơn một mức định trước h_0 nào đó. Như vậy dự báo điểm dài hạn là:

$$\hat{y}_{T+h}, \quad h > h_0, \quad (30)$$

bao gồm các giá trị dự đoán của tất cả các biến nội sinh h thời kỳ về sau. Mức h_0 phụ thuộc chủ yếu vào bản chất của các biến dự báo, đặc biệt là tốc độ khoảng dự báo “mở rộng ra” theo thời gian. Nó cũng phụ thuộc vào mục đích của dự báo. Ví dụ, trong dự báo tổng thu nhập quốc nội, với các mục đích nhất định, các dự báo dài hạn có thể có bậc tám năm, trong khi với các mục đích khác chúng có thể có bậc tám quý. Tuy nhiên, thường thì trong khung cảnh này các dự báo dài hạn đề cập đến tầm lớn hơn 5 năm.

Các dự báo dài hạn có thể thu được bằng cách phát triển liên tiếp các dự báo ngắn hạn – nghĩa là bằng việc lặp lại các dự báo thu được từ dạng rút gọn,

như trong (25). Kỹ thuật này áp dụng được đối với mô hình kinh tế lượng tuyến tính hoặc phi tuyến. Sử dụng cách tiếp cận này, các dự báo dài hạn có thể được viết như sau :

$$\hat{y}_{T+1} = y_T \hat{\Pi}_1^h + \sum_{j=0}^{h-1} \hat{z}_{T+h-j} \hat{\Pi}_2 \hat{\Pi}_1^j + \sum_{j=0}^{h-1} \hat{u}_{T+h-j} \hat{\Pi}_1^j \quad (31)$$

Trong công thức này các dự báo dài hạn có điều kiện là biết rõ các giá trị hiện tại của các biến nội sinh y_T ; biết rõ các giá trị dự báo của các biến ngoại sinh đến tận thời điểm $T+h$: $\hat{z}_{T+1}, \hat{z}_{T+2}, \dots, \hat{z}_{T+h}$; biết rõ các nhân tố bổ sung đối với tất cả các thời kỳ dự báo : $\hat{u}_{T+1}, \hat{u}_{T+2}, \dots, \hat{u}_{T+h}$; và biết rõ các ma trận hệ số ước lượng $\hat{\Pi}_1, \hat{\Pi}_2$. Như vậy, có thể phân tích ảnh hưởng của những thay đổi trong các dự báo dài hạn. Dự báo dài hạn ex-post tương ứng là:

$$\hat{\hat{y}}_{T+1} = y_T \hat{\Pi}_1^h + \sum_{j=0}^{h-1} \hat{\hat{z}}_{T+h-j} \hat{\Pi}_2 \hat{\Pi}_1^j \quad (32)$$

và các dự báo ngẫu nhiên có thể được xây dựng như trong trường hợp dự báo ngắn hạn.

2.4. Độ chính xác của dự báo

Khi đọc kết quả của một dự báo nào đó, ngắn hoặc dài hạn, người ta quan tâm chủ yếu tới mức độ chính xác cũng như ảnh hưởng của nó đối với quyết định chính sách sẽ đưa ra để hành động của mình. Mỗi quan tâm thứ hai thuộc về chủ đề đánh giá chính sách. Ở đây không bàn luận tới. Mục này bàn về độ chính xác của dự báo.

Trong một nghiên cứu về độ chính xác của các dự báo kinh tế lượng, cần phải xem xét đến các loại sai số khác nhau. Thứ nhất, sai số do lựa chọn mô hình không chính xác gây ra. Mô hình, vốn là một sự đơn giản hoá thực tế và vì vậy thường bỏ sót những ảnh hưởng nhất định và đơn giản hoá các ảnh hưởng khác. Thứ hai, sai số do số liệu được sử dụng trong việc ước lượng mô hình gây ra. Thứ ba, sai số khâu ước lượng các tham số của mô hình gây ra. Thứ tư, sai số trong dự báo các biến ngoại sinh và trong các nhân tố bổ sung gây ra. Cuối cùng, sai số do sự thiếu chính xác của các số liệu “thực tế” mà dự báo được so sánh với chúng. Tất nhiên, một số sai số có thể bù trừ nhau, dẫn tới sự chính xác giả tạo trong dự báo. Có một vài thước đo về độ chính xác của dự báo. Để đơn giản, ở đây ta chỉ xét các dự báo ngắn hạn tất định ex-ante, nhưng có thể sử dụng cách tiếp cận này để nghiên cứu độ chính xác của các dự báo dài hạn.

Sai số tuyệt đối \hat{e}_{T+1} của dự báo ngắn hạn trong (25) có được nhờ kết hợp phương trình này với phương trình (24):

$$\hat{e}_{T+1} = y_{T+1} - \hat{y}_{T+1} = (y_T \Pi_1 + z_{T+1} \Pi_2 + u_{T+1}) - (y_T \hat{\Pi}_1 + z_{T+1} \hat{\Pi}_2 + \hat{u}_{T+1}) \quad (33)$$

Trong đó \hat{e}_{T+1} có thể được viết dưới dạng:

$$\hat{e}_{T+1} = y_{T+1}(\Pi_1 - \hat{\Pi}_1) + z_{T+1}(\Pi_2 - \hat{\Pi}_2) + (z_{T+1} - \hat{z}_{T+1})\hat{\Pi}_2 + (u_{T+1} - \hat{u}_{T+1}) \quad (34)$$

Phương trình được viết lại theo cách này nhằm giúp nghiên cứu về sai số rõ ràng hơn, bởi vì mỗi một số hạng trong công thức trên của sai số có một ý nghĩa nhất định. Số hạng thứ nhất biểu thị sai số do ước lượng không đúng ma trận hệ số Π_1 , các sai số này được gán quyền số là y_T . Số hạng thứ hai biểu thị các sai số do ước lượng ma trận hệ số Π_2 gây ra, các sai số này được gán quyền số là các giá trị đúng của các biến ngoại sinh ở tương lai z_{T+1} . Số hạng thứ ba biểu thị sai số trong dự đoán các biến ngoại sinh này ở tương lai, được gán quyền số là ma trận hệ số ước lượng $\hat{\Pi}_2$. Số hạng thứ tư biểu thị sai số do các số hạng nhiễu ngẫu nhiên gây ra, ở đây \hat{u}_{T+1} là các nhân tố bổ sung. Cả bốn số hạng, nhưng đặc biệt số hạng cuối cùng, có thể phần nào chịu ảnh hưởng của sự thay đổi trong cơ cấu kinh tế qua thời gian, và đây có thể là nguyên nhân chính của sai số dự báo. Tất nhiên, một số trong các sai số này có thể và nói chung bù trừ nhau. Sai số \hat{e}_{T+1} là một biến ngẫu nhiên, vì $\hat{\Pi}_1$, $\hat{\Pi}_2$ và u_{T+1} tất cả là các biến ngẫu nhiên. Vào lúc dự đoán, y_T được giả định là một lượng đã biết (và vì vậy không phải là ngẫu nhiên). Lấy kỳ vọng, nếu các ma trận hệ số $\hat{\Pi}_1$, $\hat{\Pi}_2$ là các ước lượng không chệch thì khi giả định z_{T+1} và \hat{u}_{T+1} là tất định sẽ có:

$$E(\hat{e}_{T+1}) = (z_{T+1} - \hat{z}_{T+1})\Pi_2 - \hat{u}_{T+1}, \quad (35)$$

bởi vì u được giả định là có trung bình bằng 0.

Như vậy, sai số kỳ vọng bao gồm sai số trong dự đoán các biến ngoại sinh được gán quyền số là ma trận hệ số đúng Π_2 , trừ đi các nhân tố bổ sung. Dự báo \hat{y}_{T+1} là một dự báo không chệch nếu sai số tuyệt đối định nghĩa trong (33) có kỳ vọng bằng 0:

$$E(\hat{e}_{T+1}) = 0. \quad (36)$$

Trong trường hợp này các giá trị trung bình của các giá trị thực và dự báo trùng nhau:

$$E(y_{T+1}) = E(\hat{y}_{T+1}) \quad (37)$$

Từ (5.3), dự báo là không chệch nếu

$$\hat{u}_{T+1} = (z_{T+1} - \hat{z}_{T+1})\Pi_2, \quad (38)$$

Điều này suy ra rằng trừ trường hợp các biến ngoại sinh là dự báo không có sai số (hoặc $\Pi_2 = 0$), việc không đưa nhân tố bổ sung vào, hoặc tương đương là đặt chúng bằng 0, sẽ dẫn đến các dự báo chệch. Để dễ thấy, trình bày bằng hình học độ chính xác của dự báo trong trường hợp dự báo một biến đơn (lấy hình 2 ở Báo cáo tổng hợp). Thay đổi (tính theo) phần trăm thực tế, được chỉ ra trên trục hoành, là

$$A_T = \frac{y_{T+1} - y_T}{y_T} \times 100 \quad (39)$$

và thay đổi phần trăm dự báo, được chỉ ra trên trục tung là:

$$F_T = \frac{\hat{y}_{T+1} - y_T}{y_T} \times 100 \quad (40)$$

Trong hình này, đường 45 độ là đường các dự báo hoàn hảo, mà đối với nó thay đổi phần trăm thực tế và dự báo bằng nhau. Góc phần tư thứ nhất chứa các điểm mà đối với nó dự báo là tăng và đối với nó thực tế xảy ra tăng, góc phần tư thứ ba chứa các điểm mà đối với nó dự báo là giảm và đối với nó thực tế xảy ra giảm. Các góc phần tư thứ hai và thứ tư chứa các sai số điểm ngoặt (nghĩa là, dự báo không đúng chiều hướng thay đổi): trong góc phần tư thứ hai dự báo là tăng nhưng biến thực tế giảm về giá trị và trong góc phần tư thứ tư dự báo là giảm nhưng biến thực tế tăng về giá trị.

Một dự báo đối với các biến nào đó là tăng 3% có thể gắn với một gia tăng thực tế, ví dụ trung bình là 5%. Một cách điển hình, theo thuật ngữ của các định nghĩa AT và FT ở trên,

$$|AT| > |FT| \quad (41)$$

Trong trường hợp y_T dương, ở đây y_{T+1} và \hat{y}_{T+1} cả hai đều lớn hơn y_T , (41) suy ra rằng y_{T+1} lớn hơn \hat{y}_{T+1} – nghĩa là, các mức thực tế lớn hơn các mức dự báo. Khi đó từ (34) suy ra rằng

$$y_T(\Pi_1 - \hat{\Pi}_1) + z_{T+1}(\Pi_2 - \hat{\Pi}_2) + (z_{T+1} - \hat{z}_{T+1})\Pi_2 + (u_{T+1} - \hat{u}_{T+1}) > 0. \quad (42)$$

Lấy kỳ vọng, bất đẳng thức này suy ra là

$$\hat{u}_{T+1} < (z_{T+1} - \hat{z}_{T+1})\Pi_2, \quad (43)$$

trái với (38), ở đó đẳng thức xảy ra với các dự báo không chệch.

Trong trường hợp đơn giản nhất với một biến ngoại sinh đơn, trong đó $\Pi_2 > 0$, bất đẳng thức (43) dựa trên cả ước lượng thấp giá trị tương lai của biến ngoại sinh ($\hat{z}_{T+1} < z_{T+1}$) lẫn ước lượng thấp nhân tố bổ sung (nghĩa là bỏ qua nó).

Một giải thích có thể của sự ước lượng thấp một cách hệ thống của thay đổi là: tất cả các phương pháp dự báo dựa trên cơ sở một mô hình cụ thể hoặc không cụ thể nào đó, và tất cả các mô hình, do sự đơn giản hoá thực tế của nó, đều bao hàm một giả định rằng có các biến thích ứng nào đó hoặc không đổi hoặc được tính đến trong số hạng nhiễu ngẫu nhiên. Khi các biến này thực tế có sự thay đổi và chúng ảnh hưởng lên các biến dự báo thì kết quả của sự đơn giản hoá này là sự ước lượng thấp có hệ thống được nhận thấy. Kết quả là một sai lệch (sự chệch) bảo thủ trong dự báo, che lấp cả các dự báo tăng lẫn giảm dưới biên độ thực tế của nó.

Một thước đo đại số của độ chính xác tổng thể của một số dự báo là hệ số không bằng nhau. Nếu FT_i và AT_i tương ứng là những thay đổi phần trăm dự báo và thực tế, đối với thời kỳ (hoặc biến) i , lấy giá trị từ 1 đến m , hệ số không bằng nhau đối với tập hợp dự báo này là:

$$U_T = \frac{\sqrt{(1/m) \sum_{i=1}^m (A_{Ti} - F_{Ti})^2}}{\sqrt{(1/m) \sum_{i=1}^m A_{Ti}^2}} \quad (44)$$

Ở đây, tử số là căn của sai số bình phương trung bình dự báo, trong khi mẫu số là căn của sai số bình phương trung bình khi giả định thay đổi dự báo là 0. Trường hợp dự báo hoàn hảo là trường hợp trong đó $U = 0$. Nếu $FT_i = 0$, do đó thay đổi phần trăm dự báo là 0, nghĩa là dự báo theo nguyên trạng đối với tất cả các biến được nghiên cứu, thì $U = 1$. Do đó trường hợp $U = 1$ tương đương với một dự báo theo nguyên trạng. Tất nhiên, U có thể lớn hơn 1, trong trường hợp đó các dự báo là kém hơn, theo nghĩa tổng thể, so với các dự báo theo nguyên trạng. Tổng quát hơn, nếu

$$FT_i = k AT_i, \quad (45)$$

ở đây k là một hằng số lớn hơn 0 (nghĩa là tránh được sai số điểm ngoặt) nhưng nhỏ hơn 1 (nghĩa là có sự ước lượng thấp một cách hệ thống của sự thay đổi) thì

$$U_T = 1-k \quad (46)$$

Thí dụ, nếu thay đổi phần trăm dự báo tất cả bằng 75% của thay đổi phần trăm thực thì $UT = 0,25$.

Hệ số không bằng nhau có thể được phân rã thành một số số hạng, phản ánh những nguyên nhân khác nhau của sự không chính xác. Để thực hiện sự phân rã như vậy, thuận tiện là xét bình phương của (44). Bình phương của tử số có thể viết là:

$$\frac{1}{m} \sum_{i=1}^m (A_{Ti} - \bar{A})^2 + (\bar{A} - \bar{F})^2 + (\bar{F} - F)^2 \quad (47)$$

ở đây \bar{A} và \bar{F} tương ứng là những thay đổi phần trăm thực tế và dự báo trung bình. Lấy bình phương các số hạng trong ngoặc dẫn đến biểu thức:

$$U_{Ti}^2 = \frac{1}{(1/m) \sum A_{Ti}^2} \left[(\bar{A} - \bar{F})^2 - (S_A - S_F)^2 + 2(1-r)S_A S_F \right] \quad (48)$$

Ở đây S_A và S_F tương ứng là các độ lệch tiêu chuẩn của A và F, và r là hệ số tương quan giữa A và F. Trong biểu thức này bình phương của hệ số không bằng nhau được phân rã thành ba số hạng. Số hạng thứ nhất là bình phương của sai khác giữa các trung bình, thước đo sự chệch – nghĩa là, khuynh hướng trung tâm không bằng nhau của các thay đổi phần trăm thực tế và dự báo. Số hạng thứ hai là bình phương của sai khác giữa các độ lệch tiêu chuẩn, một thước đo của mức độ thay đổi không bằng nhau. Số hạng thứ ba gắn với hệ số tương quan và là một thước đo của tính hiệp biến không hoàn hảo. Số hạng cuối cùng này biểu thị một sai số ngẫu nhiên không hệ thống không thể tránh. Tuy nhiên, hai số hạng đầu biểu thị sai số hệ thống một cách lý tưởng là phải tránh.

2.5. Kinh nghiệm dự báo bằng các mô hình kinh tế lượng vĩ mô: những bài học tổng quát

Kinh nghiệm mới gần đây hơn với dự báo sử dụng mô hình kinh tế lượng đã chứng tỏ tầm quan trọng và giá trị của các nhân tố bổ sung. Các nhân tố bổ sung này, phản ánh đánh giá của chuyên gia về các nhân tố không được đưa vào mô hình, nói chung cải thiện đáng kể khả năng dự báo của mô hình. Các dự báo với những điều chỉnh chủ quan như vậy nói chung chính xác hơn so với các dự báo thu được từ việc áp dụng thuần túy máy móc mô hình kinh tế lượng, kể cả các dự báo ex-post. Việc kết hợp một mô hình kinh tế lượng với ý kiến chuyên gia theo cách này tận dụng những điểm tốt nhất của mỗi phương pháp.

Kinh nghiệm với các mô hình kinh tế lượng vĩ mô cũng chỉ ra rằng các mô hình như vậy nói chung có khả năng sản sinh ra những dự báo chính xác đối với các biến nhất định, nhưng cũng sản sinh ra các dự báo kém chính xác hơn nhiều đối với các biến khác. Nói chung, như ta có thể mong đợi, các biến chuyển động chậm và trơn được dự đoán chính xác hơn các biến thể hiện có phương sai cao và dao động lớn từ thời kỳ này qua thời kỳ khác.

PHẦN II

CÁC CHỈ TIÊU THỐNG KÊ XÃ HỘI VÀ LỰA CHỌN CÁC CHỈ TIÊU THỐNG KÊ XÃ HỘI CHO DỰ BÁO

Một điều cho đến nay không ai có thể phủ nhận được là xã hội có ổn định người dân mới được sống hạnh phúc, xã hội có ổn định mới phát triển được kinh tế từ đó nâng cao đời sống vật chất của người dân. Với lý do đó, các nhà quản lý và điều hành đất nước luôn muốn có thông tin phản ánh các vấn đề về xã hội. Mặt khác, càng muốn biết xu thế phát triển trong tương lai của các hiện tượng xã hội- kinh tế như thế nào. Như vậy, các chỉ tiêu thống kê xã hội-kinh tế giữ vai trò quan trọng trong việc điều hành đất nước. Việc biết được xu thế phát triển trong tương lai của chúng trở nên quan trọng. Tuy nhiên, không thể một lúc tiến hành dự báo được tất cả các chỉ tiêu, vả lại không phải tất cả các chỉ tiêu đều có đủ thông tin để tiến hành dự báo, vì vậy cần phải có sự chọn lựa và ưu tiên. Trong phần này chúng tôi sẽ tập trung vào trình bày hệ thống các chỉ tiêu thống kê xã hội-kinh tế và quá trình lựa chọn chúng để tiến hành dự báo.

I. Tổng quan về các chỉ tiêu thống kê xã hội ở Việt Nam

Theo thống kê của Liên Hợp quốc, các lĩnh vực xã hội bao gồm các nhóm chỉ tiêu như: Dân số, phân bố dân số theo vùng địa lý; việc định cư và điều kiện nhà ở, gia đình, hộ gia đình; tình trạng hôn nhân và tình hình sinh đẻ; sức khỏe và điều kiện y tế; tình trạng học vấn và điều kiện giáo dục; tình hình tham gia hoạt động kinh tế; các nhóm kinh tế và động lực xã hội; thu nhập, tiêu dùng và tài sản của dân cư; tình hình an ninh, an sinh xã hội; tham gia hoạt động văn hóa truyền thông; sử dụng thời gian nhàn rỗi.

Ở Việt Nam hệ thống các chỉ tiêu thống kê xã hội về cơ bản đã theo chỉ dẫn của thống kê LHQ. Tuy nhiên, còn thiếu các thống kê sau đây: Di cư quốc tế; Tiêu thụ năng lượng trong hộ gia đình; Tiêu thụ rượu và thuốc lá; Đền bù việc làm; Các điều kiện lao động và đào tạo; Các nhóm kinh tế-xã hội và

những biến động theo thời gian; Động lực nội thế hệ và động lực giữa các thế hệ; Phạm vi bảo hộ đối với sự mất thu nhập; Sử dụng và quy mô của bảo hộ đối với sự mất thu nhập; Sử dụng thời gian nhàn rỗi; Thời gian nhàn rỗi và các hoạt động văn hoá, các phương tiện và chi phí; Tần suất và mức độ trầm trọng của các vụ phạm tội và tình hình nạn nhân; Các đặc điểm của phạm nhân và xử lý phạm nhân; Các cơ quan và cán bộ tư pháp hình sự; Các lĩnh vực khác (Môi trường tự nhiên; Các mối quan hệ xã hội).

II. Thực trạng số liệu các chỉ tiêu thống kê xã hội ở Việt Nam, đề xuất danh sách các chỉ tiêu thống kê xã hội chủ yếu để tiến hành dự báo ngắn hạn

1. Nhóm chỉ tiêu dân số

Các chỉ tiêu trong nhóm dân số cơ bản đã đáp ứng nhu cầu của người dùng tin chủ yếu, bao gồm các cơ quan chính phủ và các viện nghiên cứu.

Các chỉ tiêu trong nhóm dân số đã: Sử dụng các khái niệm/ định nghĩa, phương pháp tính toán rõ ràng, theo chuẩn quốc tế; Phân tổ theo các phân tổ chủ yếu, gồm khu vực thành thị, nông thôn, vùng, tỉnh/ thành phố, giới tính, tuổi/ nhóm tuổi, và dân tộc; Công bố kết quả của các cuộc tổng điều tra dân số thông qua các ấn phẩm, đĩa CD-ROM, các cơ sở dữ liệu vĩ mô và vi mô, và được đưa lên trang WEB của Tổng cục Thống kê. Số liệu dân số hàng năm được công bố trong các niên giám Thống kê.

Nguồn thu thập số liệu chính về số liệu dân số là các cuộc Tổng điều tra dân số và nhà ở thực hiện 10 năm một lần. Số liệu dân số hàng năm được tính toán từ số liệu tổng điều tra dân số và số liệu điều tra chọn mẫu biến động dân số và KHHGD hàng năm.

Các chỉ tiêu trong nhóm dân số còn có các nguồn số liệu khác. Cụ thể như: Nguồn từ ngành dân số và kế hoạch hoá gia đình; Nguồn từ ngành công an; Nguồn từ ngành tư pháp.

Hai ngành Dân số - KHHGD và Tư Pháp còn có hệ thống đăng ký và thu thập số liệu về mức độ chết. Tuy nhiên, ngoài những hạn chế như chỉ tiêu tỷ suất sinh thô, cả 2 hệ thống này nói chung đều không đăng ký và thu thập được các trường hợp sinh ra chỉ được ít phút, ít ngày đã chết (chết trước khi khai sinh).

Ngành Công an cũng có hệ thống thu thập số liệu về di chuyển của dân số từ “Sổ đăng ký hộ tịch, hộ khẩu” do công an xã/phường quản lý. Tuy nhiên

số liệu về chuyển đi và chuyển đến của dân số tổng hợp được từ hệ thống sổ sách này chỉ bao gồm số người chuyển đi đã được ngành Công an cắt hộ khẩu thường trú và số người chuyển đến đã được ngành công an nhập hộ khẩu. (Số đăng ký hộ tịch, hộ khẩu cũng ghi chép các đối tượng KT3, KT4 nhưng không theo dõi được những biến động của họ theo thời gian). Chính vì vậy, số liệu di chuyển của dân số do ngành Công an tổng hợp cũng thường bị thiếu rất nhiều.

Các chỉ tiêu dân số tính toán từ các cuộc điều tra dân số của TCTK còn có một số vấn đề tồn tại sau đây mà chúng cần được khắc phục để hoàn thiện:

Do số liệu điều tra chọn mẫu biến động dân số và KHHGD hàng năm chỉ đại diện được cho cấp tỉnh trở lên nên số liệu dân số cũng chỉ tính được cho cấp tỉnh trở lên mà chưa có số liệu cho cấp huyện.

Số liệu dân số theo độ tuổi công bố hàng năm và 5 năm chỉ được tính toán và phân tổ theo nhóm 5 độ tuổi nên chưa đáp ứng yêu cầu sử dụng theo từng độ tuổi hoặc các nhóm tuổi khác theo yêu cầu của Bộ Giáo dục và Đào tạo, Bộ Y tế.

Số liệu số hộ chủ yếu thu thập và công bố qua các cuộc Tổng điều tra dân số thực hiện 10 năm 1 lần. Hàng năm, Tổng cục Thống kê cũng công bố số liệu về số hộ dựa vào kết quả suy rộng mẫu của các cuộc điều tra chọn mẫu dân số và KHHGD hàng năm, nhưng số liệu hộ của TCTK và các tỉnh có sự khác biệt. Số liệu chủ hộ chưa công bố theo giới tính nên chưa đáp ứng nhu cầu nghiên cứu những hộ có chủ hộ là nữ.

Số năm đi học trung bình của dân số năm 1989 chỉ tính chung cho toàn bộ dân số, không phân tổ theo giới cũng như chưa phân tổ đến các khu vực thành thị, nông thôn.

Tỷ lệ biết chữ của dân số cũng cần được mở rộng độ tuổi đến 10 tuổi trở lên vì một số ngành cần chỉ tiêu này. Trong bộ chỉ số MDGs cần phân tổ từ 15-24 khu vực thành thị/nông thôn.

2

Nguồn số liệu gồm các cuộc Tổng điều tra dân số và nhà ở thực hiện 10 năm một lần. Số liệu hàng năm được tính toán từ điều tra lao động việc làm, và chế độ báo cáo định kỳ theo 2 kênh: báo cáo của các Cục Thống kê về lao động do địa phương quản lý và các Bộ ngành về lao động do Trung ương quản lý. Tuy nhiên, một vấn đề vướng mắc trong nhóm chỉ tiêu lao động việc làm là cách tiếp cận. Trong các cuộc Tổng điều tra dân số và nhà ở, tình trạng hoạt

động kinh tế trong đó có thất nghiệp và đang làm việc được tiếp cận theo khái niệm hoạt động thường xuyên, tức là hoạt động trong 12 tháng qua. Cách tiếp cận này có ưu điểm là khắc phục được hạn chế về tính thời vụ của lao động, đặc biệt là lao động trong ngành nông nghiệp và tiểu thủ công nghiệp, nhưng lại bị hạn chế là mức độ chính xác không cao do phải hồi tưởng cho một thời gian dài (12 tháng qua). Trong khi đó các cuộc điều tra lao động việc làm theo các khuyến nghị của Tổ chức Lao động Quốc tế (ILO) lại tiếp cận hoạt động kinh tế theo khái niệm hoạt động hiện thời, tức là hoạt động trong 7 ngày qua, nhưng ở Việt Nam chỉ điều tra tại một thời điểm trong năm nên tính đại diện không cao do không loại được ảnh hưởng của thời vụ đến việc làm. Vì vậy điều tra Lao động - việc làm cần thực hiện thường xuyên hơn, trước hết là 6 tháng một lần, dần dần thực hiện mỗi quý một lần và tiến tới thực hiện mỗi tháng một lần. Điều này sẽ hạn chế được tính thời vụ của các hoạt động kinh tế, nhất là chỉ tiêu thất nghiệp và lao động làm việc trong các ngành nông nghiệp và tiểu thủ công nghiệp.

3

Các chỉ tiêu trong nhóm thông tin cơ bản đã đáp ứng nhu cầu của người dùng tin chủ yếu, bao gồm các cơ quan chính phủ, các viện nghiên cứu và các đối tượng thu thập nhưng sẽ thu thập được dễ dàng.

4. Nhóm chỉ tiêu về giáo dục và đào tạo

Các chỉ tiêu trong nhóm giáo dục và đào tạo cơ bản đã đáp ứng nhu cầu của người dùng tin chủ yếu, bao gồm các cơ quan chính phủ, các viện nghiên cứu và các đối tượng khác.

Các chỉ tiêu trong nhóm giáo dục và đào tạo đã: Sử dụng các khái niệm/định nghĩa, phương pháp tính toán rõ ràng, theo chuẩn quốc tế; Phân tổ theo các phân tổ chủ yếu, gồm loại hình, loại trường, giới tính, dân tộc, trình độ chuyên môn, tỉnh/TP, tuyển mới, lưu ban, bỏ học, cấp quản lý, loại cơ sở, hình thức đào tạo, ngành đào tạo cấp 2, trong nước, ngoài nước, trình độ chuyên môn; Công bố số liệu thông qua Niên giám Thống kê.

Nguồn số liệu chính về thống kê giáo dục và đào tạo là hệ thống báo cáo hành chính của Bộ Giáo dục và Đào tạo và Tổng cục dạy nghề thuộc Bộ Lao động Thương binh và Xã hội. Ngoài ra TCTK duy trì song song kênh báo cáo

từ các Cục Thống kê nhưng vẫn thu thập số liệu thứ cấp từ các Sở Giáo dục và Đào tạo.

Một số tồn tại: Đối với các chỉ tiêu cần số liệu dân số chia theo tuổi thì số liệu dân số của TCTK dự báo cho 5 nhóm tuổi, không phù hợp với độ tuổi học sinh. Dự báo theo nhiều phương án nên lựa chọn phương án nào để tính là một khó khăn. Ngoài ra số liệu dự báo dân số chưa chính xác, không phản ánh được biến động dân số do di cư, nên có một vài tỉnh số liệu học sinh đi học lớn hơn dân số cùng độ tuổi; Một số khái niệm chưa thật chuẩn để áp dụng, như: trẻ tàn tật để tính toán tình trạng nhập học, nước sạch để tính toán các chỉ tiêu về trường học. Một số chỉ tiêu việc phân tổ theo giới tính chưa thực sự chính xác do hệ thống báo cáo chưa đầy đủ, ví dụ: tỷ lệ học sinh lưu ban, bỏ học, tỷ lệ học sinh phân tổ theo nhóm tuổi và giới; Số liệu chủ yếu từ báo cáo hành chính nên khó tránh khỏi “bệnh thành tích”.

Một số số liệu của kênh TCTK khác số liệu của kênh giáo dục do thời điểm thu thập số liệu khác nhau, kênh của TCTK thường thu thập số liệu từ các Sở Giáo dục và Đào tạo sớm trong khi số liệu chính thức của kênh Giáo dục thường có muộn hơn.

Hội LHPNVN đề nghị thêm chỉ tiêu: Số lượng và tỷ lệ người (theo giới tính) được tư vấn nghề, tư vấn việc làm; Số lượng và tỷ lệ người (theo giới tính) đi xuất khẩu lao động.

Chỉ tiêu: Số trường, lớp, phòng học mầm non đề nghị thêm phân tổ chất lượng phòng học.

Chỉ tiêu: Số trường, lớp, phòng học phổ thông đề nghị thêm phân tổ loại phòng học kiên cố, bán kiên cố, nhà tạm, phòng học 3 ca, phòng học xây dựng mới và mới được cải tạo, số trường tiểu học học 2 ca 1 ngày.

Chỉ tiêu: Số học sinh phổ thông đề nghị thêm số học sinh tiểu học học 2 ca/ ngày.

Chỉ tiêu: Tỷ lệ học sinh phổ thông tốt nghiệp: Đề nghị sửa chỉ tiêu này thành: học sinh dự thi, học sinh tốt nghiệp và báo cáo 2 lần trong năm (báo cáo nhanh: 17 ngày sau khi thi tốt nghiệp và báo cáo chính thức).

Chỉ tiêu: Tỷ lệ học sinh chuyển cấp và tỷ lệ học sinh hoàn thành cấp học

Chỉ tiêu này hiện vẫn được Bộ Giáo dục và Đào tạo tổng hợp trong các báo cáo. Tuy nhiên, số liệu phân tổ về giới tính chưa được chính xác. Do việc

Bộ Giáo dục và Đào tạo đã bỏ thi tốt nghiệp tiểu học và THCS nên rất nhiều sở Giáo dục không báo cáo số liệu của hai cấp học này.

5. Nhóm chỉ tiêu y tế và chăm sóc sức khỏe

Các chỉ tiêu trong nhóm y tế và chăm sóc sức khỏe cơ bản đã đáp ứng nhu cầu của người dùng tin chủ yếu, bao gồm các cơ quan chính phủ và các viện nghiên cứu và các đối tượng khác.

Các chỉ tiêu trong nhóm y tế và chăm sóc sức khỏe đã:

Sử dụng các khái niệm/ định nghĩa, phương pháp tính toán rõ ràng, theo chuẩn quốc tế; Phân tổ theo các phân tổ chủ yếu, gồm giới tính, độ tuổi, dân tộc, thành thị, nông thôn, tỉnh/TP; Công bố số liệu thông qua Niên giám Thống kê Y tế.

Nguồn số liệu chính về thống kê y tế và chăm sóc sức khỏe là hệ thống báo cáo hành chính của Bộ Y tế.

Một số tồn tại: Hiện còn một số phân tổ chưa được Bộ Y tế thống nhất như cơ sở y tế, giường bệnh phân theo loại hình.

Trong thực tế do 17/18 chỉ tiêu trong nhóm Y tế và chăm sóc sức khỏe là báo cáo hành chính hàng năm, do đó việc phân tổ thành thị, nông thôn giới tính, độ tuổi, dân tộc... hiện gặp nhiều khó khăn. Chỉ tiêu Nhân lực y tế báo cáo hành chính không phân tổ theo loại hình, loại cơ sở y tế.

Ngành Y tế chưa thu thập số liệu đối với y tế tư nhân và y tế có vốn đầu tư nước ngoài.

Chỉ tiêu: Số bác sĩ tính trên 10.000 chưa thống nhất được cách tính do có hai quan điểm. Quan điểm thứ nhất cho rằng chỉ cần tính số bác sĩ đang làm trong các cơ sở y tế trực tiếp liên quan đến khám, chữa bệnh; trong khi quan điểm thứ hai cho rằng cần tính thêm cả số bác sĩ làm công tác quản lý, văn phòng không trực tiếp tham gia khám chữa bệnh, ví dụ khối văn phòng ở Bộ Y tế, Sở Y tế, Phòng y tế huyện và khối sản xuất kinh doanh khác.

6

hao và du lịch cơ bản đã đáp ứng nhu cầu của người dùng tin chủ yếu, bao gồm các cơ quan chính phủ, các viện nghiên cứu và các đối tượng khác.

niệm/ định nghĩa, phương pháp tính toán rõ ràng, theo chuẩn quốc tế;

Phân tổ theo các phân tổ chủ yếu, gồm khu vực thành thị, nông thôn, tỉnh/thành phố.

7. Nhóm chỉ tiêu về mức sống dân cư

Các chỉ tiêu trong nhóm mức sống dân cư cơ bản đã đáp ứng nhu cầu của người dùng tin chủ yếu, bao gồm các cơ quan chính phủ, các viện nghiên cứu và các đối tượng khác.

Các chỉ tiêu trong nhóm mức sống dân cư đã sử dụng các khái niệm/định nghĩa, phương pháp tính toán rõ ràng, theo chuẩn quốc tế; Phân tổ theo các phân tổ chủ yếu, gồm khu vực thành thị, nông thôn, vùng, tỉnh/thành phố, giới tính, nhóm tuổi, và dân tộc.

- Công bố kết quả của các cuộc khảo sát mức sống thông qua các ấn phẩm, đĩa CD-ROM, các cơ sở dữ liệu vi mô, và được đưa lên trang WEB của Tổng cục Thống kê, Niên giám Thống kê.

Nguồn thu thập số liệu chính về số liệu mức sống dân cư là các cuộc khảo sát mức sống thực hiện hai năm một lần trong những năm gần đây, từ 2002.

Tuy nhiên có một số vấn đề cần giải quyết để hoàn thiện:

- Đối với hai chỉ tiêu Thu nhập bình quân nhân khẩu một tháng và chi tiêu bình quân nhân khẩu một tháng cần bổ sung thêm các phân tổ: nhóm dân tộc chủ hộ, trình độ học vấn chủ hộ.

- Chỉ tiêu Chênh lệch thu nhập nhóm 5 so với nhóm 1 trong “Hệ thống chỉ tiêu thống kê quốc gia” có đưa phân tổ nhóm thu nhập là không có ý nghĩa nên đề nghị bỏ phân tổ này.

- Đối với chỉ tiêu Tỷ lệ nghèo cần bổ sung một số phân tổ: giới tính chủ hộ, nhóm dân tộc chủ hộ, trình độ học vấn chủ hộ. Ngoài ra cần thống nhất cách hiểu trong cách tính chỉ tiêu này khi có lạm phát, tức là cần phân biệt rõ điều chỉnh chuẩn nghèo khi mô hình mức sống thay đổi và cập nhật chuẩn nghèo khi có lạm phát.

- Chỉ tiêu Tỷ lệ hộ, nhân khẩu thiếu đói còn gặp khó khăn và chưa thống nhất trong xác định các hộ thiếu đói. Ngoài ra để tạo điều kiện thuận lợi trong việc thu thập thông tin để tính chỉ tiêu này thì cần có sự thống nhất và phối hợp chặt chẽ giữa Tổng cục Thống kê và Bộ Lao động, Thương binh và Xã hội trong xác định hộ thiếu đói và trách nhiệm thu thập thông tin.

- Đối với chỉ tiêu: Chỉ số khoảng cách nghèo cần được công bố trong những năm có Khảo sát mức sống hộ gia đình.

8

nhu cầu của người dùng tin.

. Nguyên nhân là do các số liệu về tội phạm đã được ngành Tư pháp thu thập, nhưng chưa được công bố. Chỉ tiêu bạo hành gia đình chưa có chế độ báo cáo. Chỉ tiêu “Số người được trợ giúp pháp lý” sẽ khó thu thập đầy đủ.

III. Lựa chọn các chỉ tiêu thống kê xã hội cho dự báo

3.1. Nguyên tắc lựa chọn các chỉ tiêu thống kê xã hội cho dự báo

- Là các chỉ tiêu quan trọng phản ánh đời sống chính trị xã hội và kinh tế của đất nước;
- Là các chỉ tiêu có thể thu thập thông tin một cách có hệ thống với độ tin cậy cao.

Cơ sở để chọn các chỉ tiêu cho dự báo trong thực tế phải đảm bảo các tiêu chí, cụ thể là: các chỉ tiêu do Đại hội Đảng toàn quốc nêu ra; các chỉ tiêu do Quốc hội thông qua; các chỉ tiêu thiên niên kỷ của Việt Nam; các chỉ tiêu Liên Hợp quốc, hay cơ quan thống kê khu vực quan; chỉ tiêu thuộc Hệ thống chỉ tiêu thống kê quốc gia đã được Thủ tướng Chính phủ phê duyệt.

3.2. Đề xuất các chỉ tiêu thống kê xã hội chủ yếu để tiến hành dự báo ngắn hạn

cấu hộ; Tỷ số giới tính của dân số; Tỷ số giới tính của trẻ em mới sinh; Tỷ suất sinh; Tỷ suất chết; Tỷ suất chết của trẻ em dưới 1 tuổi; Tỷ suất chết của trẻ em dưới 5 tuổi; Tuổi thọ trung bình tính từ lúc sinh.

Các chỉ tiêu này là các chỉ tiêu cơ bản phản ánh quy mô, tốc độ và chất lượng về mặt thể chất của dân số, nên chúng cần được dự báo để có chính sách về dân số hợp lý.

tế.

Các chỉ tiêu này phản ánh quy mô và tốc độ tăng lao động của xã hội. Mặt khác chúng cũng phản ánh tình trạng việc làm và sự phân biệt đối xử giữa nam và nữ vì vậy cần được dự báo sớm để các nhà quản lý có những điều chỉnh kịp thời.

Tỷ lệ hộ có TV; Tỷ lệ hộ có máy tính.

Các chỉ tiêu này phản ánh điều kiện thông tin của người dân, một điều kiện quan trọng giúp cho người dân nắm bắt tình hình kinh tế xã hội,...

; Chi tiêu cho giáo dục bình quân đầu người của hộ gia đình.

Các chỉ tiêu này là các chỉ tiêu cơ bản của lĩnh vực giáo dục, vì vậy chúng cần được dự báo sớm để các nhà quản lý nắm bắt trước mà có những chính sách hiệu chỉnh theo mong muốn.

Số cơ sở y tế; S

HIV, số bệnh nhân AIDS, s

-

; Chi tiêu cho y tế bình quân đầu người của hộ gia đình.

Các chỉ tiêu này phản ánh điều kiện y tế cũng như sức khỏe của người dân, các yếu tố ảnh hưởng trực tiếp tới chất lượng dân số vì vậy cũng cần có được các thông tin dự báo.

; Chi tiêu cho văn hóa, thể thao bình quân đầu người của hộ gia đình.

Các chỉ tiêu này phản ánh đời sống văn hoá cơ bản của người dân, một chỉ tiêu được nhà nước quan tâm.

; Hệ số bất bình đẳng trong phân phối thu nhập (hệ số Gini); Tiêu dùng năng lượng bình quân đầu người; Mức và tỷ lệ chi tiêu của chính phủ và cá nhân cho các dịch vụ y tế và chăm sóc sức khỏe ban đầu.

Các chỉ tiêu này phản ánh điều kiện đời sống vật chất của người dân. Một lĩnh vực tất cả các quốc gia đều quan tâm do vậy cũng cần có thông tin sớm cho công tác quản lý điều hành.

Các chỉ tiêu này phản ánh điều kiện an toàn xã hội. Một điều kiện quan trọng của đời sống cộng đồng, vì vậy cũng cần có thông tin dự báo để kịp thời điều chỉnh các hành vi xã hội.

PHẦN III
MỘT SỐ PHƯƠNG PHÁP DỰ BÁO VÀ KẾT QUẢ DỰ BÁO CHO
MỘT SỐ CHỈ TIÊU THỐNG KÊ XÃ HỘI

I. Dự báo cho các chỉ tiêu thống kê dân số

Số lượng dân và cơ cấu theo giới tính và độ tuổi của dân số giữ vai trò quan trọng trong đời sống kinh tế cũng như xã hội của một quốc gia nói chung và của một vùng nói riêng. Lý do rất dễ thấy là dân số và cơ cấu dân số tác động lên các mặt của đời sống xã hội rất mạnh. Cơ cấu dân số thiên sang trẻ, đời sống xã hội rất dễ bị mất cân đối về y tế, giáo dục, việc làm và đặc biệt là định hướng các hành vi xã hội. Cơ cấu dân số thiên sang già, đời sống xã hội rất dễ bị mất cân đối trong khâu nuôi dưỡng người già, chăm sóc bệnh tật, ...

1.1. Dự báo dân số theo dãy số thời gian

a. Mô hình theo dãy số thời gian

Mô hình mô tả sự phát triển của dân số theo thời gian theo phát hiện của Malthus có dạng:

(1) $P_t = P_0 e^{rt}$, trong đó P_t là dân số thời kỳ báo cáo, P_0 là dân số thời kỳ gốc, r là tốc độ tăng dân số, t là thời gian.

Trong thực tế có người sử dụng công thức sau đây để mô tả sự phát triển của dân số theo thời gian:

(2) $P_t = P_0 (1+r)^t$, trong các ký hiệu như ở công thức (1).

Tuy nhiên, về bản chất công thức (2) và công thức (1) đều giống nhau. Một mô hình mô tả quá trình rời rạc (mô hình (2)), còn một mô hình mô tả quá trình liên tục (mô hình (1)). Về mặt toán học, khi t tiến đến vô cùng mô hình (2) sẽ trở về mô hình (1). Với lý do đó và trong thực tế sự phát triển của dân số là một quá trình liên tục, nên mô hình (1) được sử dụng để dự báo dân số theo dãy số thời gian.

Một mô hình khác hay được nước Mỹ sử dụng để dự báo dân số là hàm số Logistic:

(3) $P_t = \frac{K}{1+e^{-rt}}$, trong K biểu thị mức tối hạn của dân số, còn các ký hiệu khác như ở công thức (1).

Phương pháp dự báo dân số theo thời gian có ưu điểm là đơn giản, dễ thực hiện, song có một số nhược điểm sau:

+ Chỉ sử dụng kết quả của các cuộc tổng điều tra mới tính được tốc độ tăng của dân số, vì vậy độ chính xác của phương pháp dự đoán này phụ thuộc rất nhiều vào độ chính xác của thông tin thu được từ các cuộc tổng điều tra.

+ Thông thường việc áp dụng phương pháp này dựa chủ yếu vào kết quả của hai cuộc tổng điều tra dân số, vì vậy ngoài việc phụ thuộc vào độ chính xác của hai cuộc điều tra nó còn phụ thuộc vào khoảng cách của hai cuộc tổng điều tra dân số. Thế nhưng, do tổng điều tra tốn kém nên thường 10 năm mới tổ chức một lần. Khoảng cách giữa hai cuộc tổng điều tra dài như vậy chắc chắn sẽ làm cho tốc độ tăng dân số ước lượng được từ mô hình sẽ kém sát thực, vì vậy kết quả dự báo cũng kém sát thực.

+ Dự báo dân số theo mô hình này không cho được cơ cấu dân số theo giới tính và nhóm tuổi. Thế mà, nhu cầu dân số theo giới tính và nhóm tuổi lại rất cao trong công tác lập kế hoạch phát triển kinh tế xã hội.

b. Mô hình cân bằng dân số

Một mô hình đơn giản khác được dùng để dự đoán dân số là mô hình cân bằng dân số. Mô hình này có dạng:

(4) $P_t = P_0 + B_{0,t} - D_{0,t} + IM_{0,t} - OM_{0,t}$, trong đó P_t là dân số thời kỳ báo cáo, P_0 là dân số thời kỳ gốc, $B_{0,t}$ là số sinh trong kỳ, $D_{0,t}$ là số chết trong kỳ, $IM_{0,t}$ dân nhập cư trong kỳ, $OM_{0,t}$ là dân xuất cư trong kỳ.

1.2. Phương pháp dự báo thành phần

Một phương pháp khác hay được sử dụng trong dự báo dân số là phương pháp thành phần. Về mặt cơ bản, phương pháp thành phần áp dụng mô hình cân bằng dân số. Tuy nhiên, mô hình này được biến đổi về dạng:

(5) $P_t = (P_0 - D_{0,t}) + B_{0,t} \pm NM_{0,t}$, trong đó các ký hiệu giống như ở mô hình (4), riêng $NM_{0,t}$ là mức di cư thuần túy.

Ở mô hình (5), dân số (P_t) được biểu diễn bằng ba thành phần chính: thành phần thứ nhất ($P_0 - D_{0,t}$) biểu thị số lượng dân số ở thời kỳ gốc còn lại sau khi đã trừ đi số người chết trong kỳ dự báo. Thành phần thứ hai ($B_{0,t}$) biểu thị số người mới được sinh ra trong thời kỳ dự báo. Thành phần thứ ba ($NM_{0,t}$) biểu thị di cư thuần túy trong thời kỳ dự báo.

Ở phương pháp thành phần, khi tiến hành dự báo thành phần thứ nhất sử dụng cơ cấu dân số theo giới tính và độ tuổi ở thời kỳ gốc, sau đó nhân cơ cấu

này với hệ số sống tương ứng của thời kỳ này để có dân số còn tồn tại ở thời kỳ dự báo. Do giữa nam và nữ có hệ số sống khác nhau và ngay trong một giới thì ở độ tuổi khác nhau cũng có hệ số sống khác nhau nên khi dự báo (chuyển tuổi cho dân số gốc) thành phần thứ nhất người ta sử dụng hai bộ hệ số sống khác nhau cho hai giới.

Để dự báo thành phần thứ hai ($B_{0,t}$), đầu tiên dự báo tỷ lệ sinh đặc trưng theo nhóm tuổi của phụ nữ trong thời kỳ dự báo sau đó sử dụng chúng để tính tổng số trẻ em được sinh ra trong thời kỳ dự báo.

Để dự báo cho thành phần thứ ba cần thu thập thông tin (qua cơ quan chức năng) về số người di cư quốc tế (xuất cư và nhập cư) của dân số. Tuy nhiên, do hầu hết các nước đều có chính sách cấm nhập cư nên lượng người xuất và nhập cư quốc tế ít vì vậy khi tiến hành dự báo dân số thường người ta giả thiết thành phần này không xuất hiện ($NM_{0,t}=0$).

1.3. Quy trình dự báo theo phương pháp thành phần

Các bước khi tiến hành dự báo dân số theo phương pháp thành phần như sau:

+ Bước 1: Để tiến hành dự báo dân số của một nước theo phương pháp thành phần đòi hỏi phải có các loại số liệu cơ bản là cơ cấu dân số theo giới tính và nhóm 5 độ tuổi ở thời điểm xuất phát, bảng sống của nam và nữ, tỷ lệ sinh đặc trưng theo 7 nhóm tuổi sinh đẻ của phụ nữ và nếu có tình trạng di cư quốc tế mạnh thì cần có cả tỷ lệ di cư thuần túy theo giới tính và nhóm tuổi. Tuy nhiên, do nhiều nước đóng cửa với di cư quốc tế nên thành phần này thường coi như không xuất hiện. Như vậy bước đầu tiên khi tiến hành dự báo theo phương pháp thành phần là thu thập thông tin về cơ cấu dân số theo giới tính và nhóm tuổi ở thời điểm khởi đầu của dự báo, thu thập thông tin về tỷ lệ sinh đặc trưng theo nhóm tuổi, thu thập và tính toán thông tin về hệ sống của dân số.

Bước 2: Sau khi đã thu thập thông tin cần tiến hành đánh giá chất lượng số liệu và hiệu chỉnh chúng nếu thấy cần thiết. Ở bước này có một số kỹ thuật về nhân khẩu học được áp dụng. Ví dụ như để đánh giá hiện tượng báo tuổi sai sử dụng chỉ số Mayer (cho độ tuổi) hoặc chỉ số UN Joint Score (cho nhóm 5 độ tuổi). Để đánh giá mức độ thiếu hụt của dân số có thể sử dụng tỷ lệ giới tính,...

Bước 3: Nghiên cứu số liệu gốc xây dựng các phương án dự báo.

Bước 4: Thiết lập bảng cơ sở dữ liệu gốc (cơ cấu dân số theo giới tính và nhóm tuổi, mức sinh và mô hình sinh, mức chết và mô hình chết,...).

Bước 5: Thực hiện việc dự báo số người ở thời điểm gốc hiện còn sống sau 5 năm, 10 năm, 15 năm,... (Thực hiện phép chuyển tuổi).

Để tiến hành ước lượng này cần phải xác định mức chết của dân số trong thời kỳ dự báo thông qua chỉ tiêu tuổi thọ bình quân lúc sinh hoặc tỷ lệ chết của trẻ sơ sinh (dưới 1 tuổi). Sau đó sử dụng bảng sống mẫu của Coale Demeny để tính hệ số sống sử dụng cho công việc chuyển tuổi của dân số gốc để dự báo số người ở thời điểm gốc hiện còn sống sau 5 năm, 10 năm, 15 năm,...

Bước 6: Dự báo số sinh ở các thời kỳ dự báo. Để dự báo được số trẻ được sinh ra trong thời kỳ dự báo, cần dự báo tỷ lệ sinh tổng cộng (TFR) ở thời kỳ này và dạng sinh đẻ của phụ nữ (được xác định thông qua tỷ lệ sinh đặc trưng theo nhóm tuổi của phụ nữ). Có được tỷ lệ sinh đặc trưng theo nhóm tuổi của phụ nữ và biết được số phụ nữ ở các nhóm tuổi trong độ tuổi sinh đẻ ta nhân chúng với nhau sẽ có được số trẻ được sinh ra trong thời kỳ dự báo.

II. Các phương pháp dự báo lực lượng lao động

Để dự báo lực lượng lao động, thường người ta sử dụng một trong hai phương pháp, đó là dự báo theo dãy số thời gian hoặc dự báo theo phương pháp nhân khẩu học (phương pháp dự báo thành phần).

Cũng giống như dự báo cho các chỉ tiêu thống kê bằng phương pháp dãy số thời gian khác, để dự báo về lực lượng lao động bằng phương pháp dãy số thời gian, người ta phải thu thập thông tin về số người thuộc lực lượng này qua các thời kỳ để lập thành một dãy số thời gian. Trên cơ sở dãy số liệu này tìm ra một phương trình toán học (mô hình) thích hợp mô tả quy luật phát triển của dãy số và sử dụng nó để dự đoán. Việc ước lượng các tham số của phương trình thường dựa vào phương pháp bình phương nhỏ nhất.

Dự báo lực lượng lao động theo dãy số thời gian có ưu điểm là: Thứ nhất, nó đơn giản, dễ làm; Thứ hai, nhanh chóng cho ra kết quả dự báo cần thiết. Tuy nhiên, nó có các nhược điểm sau đây: Thứ nhất, khó thu thập thông tin cho một dãy số liệu dài. Mặt khác, các thông tin có được thường mức độ sát thực thấp và số liệu giữa các thời kỳ thường có phạm vi không đồng nhất do vậy kết quả dự báo thu được cũng có mức độ sát thực thấp; Thứ hai, số liệu dự

báo thu được chỉ là một con số tổng hợp mà không tách ra được theo giới tính và nhóm tuổi nên tác dụng của dự báo cũng bị hạn chế theo.

Dự báo lực lượng lao động bằng phương pháp thành phần là phương pháp thường kết hợp giữa dự báo dân số với dự báo tỷ lệ lực lượng lao động nói chung và ở từng nhóm (độ) tuổi nói riêng. Ở phương pháp này cần thực hiện hai bước. Bước thứ nhất, sử dụng phương pháp thành phần để dự báo dân số theo giới tính và nhóm tuổi. Bước thứ hai, dự đoán tỷ lệ lực lượng dân số của các nhóm và sử dụng chúng cùng với dân số đã dự đoán được ở bước trên để dự đoán số lượng lực lượng lao động của các nhóm tuổi và của toàn bộ dân số.

Việc dự báo tỷ lệ lực lượng lao động nói chung và tỷ lệ lực lượng lao động đặc trưng theo nhóm tuổi nói riêng thường được thực hiện theo phương pháp dãy số thời gian. Dựa vào dãy số thời gian dự báo được tỷ lệ lực lượng lao động và sau đó kết hợp với dự báo phân bố của lực lượng lao động dự báo tỷ lệ lực lượng lao động đặc trưng theo từng nhóm tuổi.

Phương pháp dự báo lực lượng lao động theo phương pháp thành phần có các ưu điểm sau:

Thứ nhất, tận dụng được kết quả của dự báo dân số và các ưu điểm của dự báo dân số theo phương pháp thành phần (kết quả dự báo dân số theo phương pháp thành phần thường sát với thực tế hơn các phương pháp khác); Thứ hai, nó cho biết phân bố của lực lượng lao động theo giới tính và nhóm tuổi. Thông tin này rất bổ ích cho những người làm công tác hoặc định các chính sách về lao động, việc làm.

Với các ưu điểm trên, phương pháp thành phần thường hay được sử dụng trong dự báo lực lượng lao động. Bên cạnh các ưu điểm trên, phương pháp dự báo lực lượng theo phương pháp thành phần có các nhược điểm là nó tiến hành phức tạp, đòi hỏi phải thu thập nhiều loại thông tin hơn so với phương pháp dãy số thời gian.

III. Các phương pháp dự báo tỷ lệ thất nghiệp và số người thất nghiệp

Cũng giống như một số chỉ tiêu thống kê dân số khác, các chỉ tiêu tỷ lệ thất nghiệp và số lượng người thất nghiệp có thể dự báo bằng các phương pháp như: dãy số thời gian, hồi quy tương quan đa nhân tố và phương pháp thành phần. Nội dung của phương pháp dãy số thời gian và hồi quy đa nhân tố (biến) đã được trình bày ở phần phương pháp luận dự báo. Do không có nguồn số liệu

để áp dụng phương pháp hồi quy đa nhân tố, ở đây chúng tôi chỉ giới thiệu một số nét cơ bản cơ bản của việc áp dụng phương pháp hồi quy đa nhân tố áp dụng cho dự báo tỷ lệ người thất nghiệp và kinh nghiệm chọn biến của Mỹ trong dự báo chỉ tiêu này.

Trong thực tế dự báo, mô hình đa nhân tố dạng tuyến tính được áp dụng để dự báo chỉ tiêu tỷ lệ thất nghiệp và nó có thể ứng dụng theo không gian và cũng có thể được ứng dụng theo thời gian. Theo không gian có nghĩa là thông tin về các biến được thu thập từ các doanh nghiệp, hoặc các địa phương để từ đó ước lượng các thông số của mô hình và để tiến hành dự báo, trước tiên dự báo các biến của mô hình sau đó lắp vào mô hình để tính tỷ lệ thất nghiệp. Theo thời gian có nghĩa là số liệu của các biến theo thời gian được thu thập để ước lượng các tham số của mô hình, sau đó để tiến hành dự báo phải tiến hành một bước tiếp theo là dự báo từng yếu tố của mô hình để lắp vào phương trình dự báo.

Trong mô hình dự báo tỷ lệ thất nghiệp của Mỹ người ta đã chọn các yếu tố sau đây làm biến độc lập của mô hình:

+ Tỷ lệ % các chủ doanh nghiệp báo cáo là có kế hoạch tăng thêm lực lượng lao động;

+ Tỷ lệ % các chủ doanh nghiệp báo cáo là ít nhất một công việc sẽ được doanh nghiệp mở ra.

+ Tỷ lệ chủ doanh nghiệp kỳ vọng khối lượng hàng hoá bán ra của họ sẽ tăng lên.

Phân tích định tính thấy việc lựa chọn các biến đó để dự báo tỷ lệ thất nghiệp là hợp lý. Khi tỷ lệ các chủ doanh nghiệp có kế hoạch tăng thêm lực lượng lao động sẽ dẫn đến họ tuyển dụng lao động vào doanh nghiệp, nhờ vậy số người tham gia lực lượng lao động sẽ tăng lên và điều đó đồng nghĩa với việc tỷ lệ thất nghiệp giảm đi. Tương tự như vậy, khi tỷ lệ các chủ doanh nghiệp báo cáo sẽ mở ra ít nhất một công việc mới vậy cũng sẽ làm tăng khả năng tuyển dụng thêm lao động và nhờ vậy tỷ lệ thất nghiệp cũng giảm đi. Tỷ lệ doanh nghiệp mở ra ít nhất một công việc mới càng cao, tỷ lệ thất nghiệp càng thấp. Còn tỷ lệ các doanh nghiệp kỳ vọng khối lượng hàng hoá của họ bán ra tăng cao làm tỷ lệ lực lượng lao động tăng, vì một trong hai giải pháp để tăng hàng hoá bán ra là tuyển dụng thêm nhân công để sản xuất ra được nhiều hàng hoá, nhờ vậy tỷ lệ thất nghiệp giảm.

Tương tự như dự báo lực lượng lao động, dự báo số lượng người thất nghiệp cũng được tiến hành theo một trong hai phương pháp: phương pháp dãy số thời gian và phương pháp nhân khẩu học (phương pháp thành phần). Quy trình áp dụng phương pháp dãy số thời gian để dự báo số lượng người thất nghiệp là quy trình cũng giống như đã áp dụng cho dự báo các chỉ tiêu thống kê khác.

Ở phương pháp nhân khẩu học thường dựa vào kết quả dự báo dân số theo phương pháp thành phần sau đó tiến hành dự báo tỷ lệ thất nghiệp theo giới tính và nhóm tuổi, sử dụng hai kết quả này để tính ra số người thất nghiệp của thời kỳ dự báo. Với cách làm này có được số người có khả năng thất nghiệp chi tiết ở từng nhóm tuổi cho từng giới tính.

IV. Dự báo thu nhập hộ gia đình và chỉ số giá tiêu dùng (Tỷ lệ lạm phát)

4.1. Dự báo thu nhập hộ gia đình

Nâng cao mức sống của dân cư là một mục tiêu rất cơ bản của mọi quốc gia. Dân có giàu, nước mới mạnh, mặt khác, đời sống vật chất lên cao kéo theo đời sống tinh thần cũng đi lên. Ở nước ta, mức thu nhập của người dân còn thấp, vì vậy nâng cao mức thu nhập của các hộ gia đình trở thành một nhiệm vụ lớn của Đảng và Nhà nước. Với lý do này chỉ tiêu thống kê “mức thu nhập bình quân một hộ” trở thành một chỉ tiêu thống kê quan trọng cần được theo dõi thường xuyên. Vì vậy, cứ hai năm một lần chúng ta tổ chức một cuộc điều tra khảo sát mức sống hộ gia đình. Tuy nhiên, bên cạnh nhu cầu thông tin phản ánh tình hình thu nhập thực tế của các hộ gia đình, cần có các thông tin dự báo, bởi có như vậy mới biết khả năng sắp tới thu nhập của các hộ gia đình như thế nào để từ đó có những biện pháp thúc đẩy mức thu nhập của người dân lên cao.

Để dự báo thu nhập bình quân một đầu người người ta cũng thường sử dụng một trong hai phương pháp, đó là dự báo theo dãy số thời gian hoặc dự báo theo phương pháp tương quan hồi quy đa nhân tố. Quy trình thực hiện các phương pháp này giống như đã được trình bày ở các phần dự báo trên. Tuy nhiên, còn có thể dự báo chỉ tiêu này còn có thể dự báo bằng phương pháp kết hợp, đó là dự báo riêng cho chỉ tiêu thu nhập, đại diện là chỉ tiêu GDP, và dự báo riêng cho chỉ tiêu số lượng dân số sau đó sử dụng kết quả dự báo được để tính thu nhập bình quân đầu người. Phương pháp này có ưu điểm là cho phép tận dụng hai nguồn số liệu khác nhau để dự báo.

4.2. Dự báo chỉ số giá tiêu dùng

Lạm phát là một chỉ tiêu kinh tế- xã hội được nhiều quốc gia quan tâm và luôn muốn biết tương lai của chỉ tiêu này như thế nào. Thực tế cho thấy, nếu tỷ lệ lạm phát cao và kéo dài liên tục, đời sống của dân chúng sẽ bị ảnh hưởng nặng nề vì vậy dẫn đến các hệ quả xã hội nghiêm trọng. Với các lý do như vậy, nhiều quốc gia luôn tìm các giải pháp để không chế tỷ lệ lạm phát ở mức dưới hai con số và cố gắng duy trì ở mức khoảng từ 3-5% năm.

Cũng giống như ở các chỉ tiêu thống kê kinh tế-xã hội khác, người ta thường dùng các phương pháp dự đoán theo dãy số thời gian, dự đoán theo mô hình hồi quy đa nhân tố, dự đoán theo phương pháp tự hồi quy và phương pháp đánh giá của chuyên gia để dự đoán chỉ tiêu lạm phát.

Trên bình diện macro, người ta còn sử dụng mô hình sau đây để dự đoán chỉ số lạm phát:

$$p_t = \alpha_1 p_t^e + \beta_1 (y_t / \hat{y}_t) + \gamma w_t + \sigma_e p w_t$$

$$w_t = \alpha_2 p_t^e + \beta_2 (N_{t-1} / L_{t-1}) \text{ đường cong Philip}$$

$$p_t^e = \theta p_{t-1} + (1 - \theta) p^*$$

Trong đó: p là chỉ số lạm phát; p^e là giá trị p kỳ vọng; p^* là chỉ số lạm phát mục tiêu; y là tổng sản phẩm được sản xuất ra, \hat{y} là tổng sản phẩm tiềm năng; N là số lượng lao động, L lượng cung lao động; w là lạm phát lương.

Để dự đoán chỉ số lạm phát theo mô hình này, đòi hỏi phải có thông tin của các biến được thể hiện trong mô hình trong ít nhất là 5 năm. Trong điều kiện thông tin ở nước ta thiếu, mô hình này hiện nay chưa thể sử dụng được.

V. Kết quả dự báo một số chỉ tiêu thống kê xã hội

5.1. Kết quả dự báo dân số năm 2014, 2019 (theo phương pháp thành phần)

5.1.1. Về quy mô và tốc độ tăng dân số

Bảng 1 trình bày kết quả dự báo quy mô dân số theo ba phương án sinh (đại diện là chỉ tiêu Tỷ lệ sinh tổng cộng TFR) khác nhau. Do phương án dự báo II (PAII) dễ xảy ra trong thực tế nên ở đây chủ yếu trình bày kết quả dự báo của phương án này. Theo dự báo, vào ngày 1/ 4/ 2014 tổng số dân của Việt Nam vào khoảng 90,7 triệu người. So với các phương án dự báo khác, kết quả

dự báo của phương án II chỉ chênh khoảng từ 200 – 300 nghìn người, một con số nhỏ so với tổng số dân số nói chung. Vào ngày 1/ 4/ 2019 dân số Việt Nam vào khoảng 94,7 triệu người. Như vậy sau 10 năm (từ 2009-2019) dân số Việt Nam chỉ tăng khoảng 8,2 triệu người. So với thời kỳ trước (1989-1999), số lượng tăng dân số giảm đáng kể (8,2 triệu người so với 10,1 triệu người).

Bảng 1: Dân số và tốc độ tăng dân số Việt Nam trong giai đoạn 2014-2019

	1/4/2009			1/4/2014			1/4/2019		
	Nam	Nữ	Tổng	Nam	Nữ	Tổng	Nam	Nữ	Tổng
PAI	43136	43307	86443	45557	43580	90991	47920	43828	95528
PAII	43136	43307	86443	45403	43580	90710	47475	43828	94679
PAIII	43136	43307	86443	45249	43580	90410	47029	43828	93811

Bảng 2 trình bày kết quả dự báo một số chỉ tiêu dân số khác. Kết quả của bảng này cho thấy trong thời kỳ 2009-2014 có khoảng 7,4 triệu trẻ em được sinh ra và 1,6 triệu người bị mất đi. Như vậy, mức tăng dân số tự nhiên là khoảng 5,8 triệu người. Vào thời kỳ 2014-2019 có khoảng 7,2 triệu trẻ được sinh ra, khoảng 1,6 triệu người bị mất đi và số tăng tự nhiên là 5,6 triệu người. Trong bảng còn trình bày chỉ tiêu số người năm sống bình quân cho từng thời kỳ dự báo. Chỉ tiêu này được tính bằng phương pháp nội suy hàm lũy thừa cơ số tự nhiên (cơ số e; hàm EXP). Thông qua chỉ tiêu này, tính tỷ lệ sinh, chết và tăng dân số tự nhiên hợp logic hơn so với các phương pháp tính thông thường khác và cũng sát với thực tế hơn. Theo cách tính này, tỷ lệ sinh của Việt Nam thời kỳ 2009-2014 là 16,7 phần nghìn, còn tỷ lệ sinh vào thời kỳ 2014-2019 là 15,5 phần nghìn. Tỷ lệ chết của hai thời kỳ này tương ứng là 3,7 và 3,5 phần nghìn. Tỷ lệ tăng dân số của hai thời kỳ này vẫn đạt mức trên 1%: thời kỳ 2009-2014 là 1,3%; thời kỳ 2014-2019 là 1,2%.

Bảng 2: Dự báo một số chỉ tiêu dân số cơ bản

	2009-2014	2014-2019
Sinh (nghìn người)	7382	7174
Chết (nghìn người)	1622	1623
Tăng tự nhiên (nghìn người)	5760	5552
Số người năm sống (nghìn người)	442752	463365
Tỷ lệ sinh (%)	16,7	15,5

Tỷ lệ chết (%)	3,7	3,5
Tỷ lệ tăng dân số bình quân năm (%)	13,0	12,0
Tỷ lệ tăng dân số 15-59 tuổi (%)	12,4	5,9

Dân số trong độ tuổi lao động ở hai thời kỳ 2009-2014 và 2014-2019 vẫn tăng. Ở giai đoạn 2009-2014 tăng nhanh hơn so với ở giai đoạn 2014-2019 (1,24% so với 0,59%). Tốc độ tăng của lực lượng lao động giảm mạnh vào thời kỳ 2014-2019 là do có sự tác động mạnh của hiện tượng giảm sinh của các thời kỳ trước.

5.1.2. Về cơ cấu dân số

Bảng 3 trình bày cơ cấu dân số theo ba nhóm tuổi chính: 0-14, 15-59, và 60+. Kết quả dự báo cho thấy về giá trị tuyệt đối số trẻ em ở độ tuổi từ 0-14 ở cả ba thời điểm đều giữ mức ổn định ở khoảng 21,3 và 21,4 triệu em. Tuy nhiên, số người trong độ tuổi lao động và người già (trên 60) lại tăng đáng kể. Sau 10 năm số người trong độ tuổi lao động tăng tới 5,5 triệu người, còn số người già tăng khoảng 2,7 triệu người. Kết quả là cơ cấu tương đối của dân số thay đổi và thiên về tăng người già: tỷ trọng người già đã tăng lên từ mức 8,9% vào 1/4/ 2009 lên mức 11,0% vào 1/4/ 2019. Kết quả dự đoán cũng cho thấy, trong cả hai giai đoạn 2009-2014 và 2014-2019 cơ cấu dân số Việt Nam vẫn duy trì ở trạng thái có cơ cấu “dân số vàng”.

Kết quả dự báo còn cho thấy tốc độ tăng bình quân một năm của lực lượng lao động trong thời kỳ 2014-2019 giảm mạnh so với tốc độ tăng bình quân của thời kỳ 2009-2014: từ 1,24%/ năm xuống còn 0,59% một năm.

5.2. Kết quả dự báo lao động

5.2.1. Dự báo lao động đang làm việc hàng năm sử dụng phương pháp dãy số thời gian

Dựa vào dãy số liệu thống kê về lực lượng lao động đang làm việc do TCTK thu thập cho giai đoạn 2000-2009 tiến hành dự báo cho chỉ tiêu này vào các năm 2010, 2011, 2012, 2013,....

Sử dụng chương trình SPSS đã xác định được hàm số bậc hai mô tả tốt nhất quy luật phát triển theo thời gian của chỉ tiêu lao động đang làm việc. Điều này thể hiện rất rõ thông qua hệ số tương quan (R) và hệ số mô tả (R^2) ước lượng được. Hệ số tương quan thu được: $R=1$; $R^2=1$. Điều này cho thấy biến thời gian và chỉ tiêu này có mối quan hệ tất định. Hơn thế nữa, sai số dự báo tương đối chỉ ở mức 0,02%. Như vậy, ta có thể kết luận là hàm số bậc hai

phù hợp với việc mô tả sự biến động theo thời gian của chỉ tiêu lực lượng lao động và yên tâm sử dụng nó để dự báo. Hàm mô tả thu được có dạng cụ thể như sau:

$y_t = 3601,3 + 1054,25t + 11,9087t^2$, với $R=1$; $R^2=1$; Sai số tuyệt đối bình quân bằng 7,4 nghìn người; sai số tương đối bằng 0,02%. Hệ số mô tả $R^2=1$ cho thấy đây là mô hình tất định.

Theo kết quả tính toán, năm 2010 lao động đang làm việc của Việt Nam là 49,0 triệu người. Lao động đang làm việc của Việt Nam vào các năm 2011, 2012, 2013,... dự báo tương ứng sẽ là 50,4 triệu, 51,7 triệu và 53,1 triệu người. Với mức độ tin cậy là 95%, lực lượng lao động của các năm 2009, 2010, 2011 và 2012 sẽ nằm trong các khoảng tương ứng sau (đơn vị là triệu người): (47,7;49,1), (49,0;50,4), (50,4;51,7), (51,7;53,1).

Bên cạnh dự báo dân số đang làm việc cho toàn quốc, chúng tôi tiến hành dự đoán lao động đang làm việc cho khu vực thành thị và nông thôn. Kết quả dự báo được trình bày ở các bảng trên cho thấy vào các năm 2010, 2011, 2012, 2013 dân số đang làm việc của khu vực thành thị tương ứng là 13,2 triệu, 13,8 triệu, 14,5 triệu và 15,2 triệu người, còn của khu vực nông thôn các con số tương ứng là 35,8 triệu, 36,5 triệu, 37,2 triệu, 37,9 triệu người.

5.2.2. Dự báo dân số hoạt động kinh tế theo giới tính và nhóm tuổi

Để dự báo chỉ tiêu dân số hoạt động kinh tế theo giới tính và nhóm tuổi sử dụng phương pháp thành phần. Theo phương pháp này cần phải dự đoán dân số ở các nhóm tuổi và tỷ lệ dân số hoạt động kinh tế ở từng nhóm tuổi. Như vậy, ở đây sử dụng kết quả dự báo dân số đã được tiến hành. Bước tiếp theo cần dự báo tỷ lệ hoạt động kinh tế cho từng nhóm tuổi. Dựa vào kết quả của ba cuộc tổng điều tra dân số (1989, 1999 và 2009), chúng tôi dự đoán tỷ lệ này cho các năm sau.

Số liệu lịch sử cho thấy, dạng của tỷ lệ dân số hoạt động kinh tế theo nhóm tuổi của ba cuộc tổng điều tra dân số gần giống nhau: thấp ở các nhóm tuổi đầu sau tăng lên ở các nhóm tuổi trung niên rồi giảm đi ở các nhóm tuổi già. Về mức độ thì vào năm 1999 là cao nhất, còn ở năm 1989 và 2009 thì tương đương. Điều này có thể do ảnh hưởng của khủng hoảng kinh tế vào các năm 1997-1998. Ở hai năm 1989 và 2009 mức độ hoạt động kinh tế ở các nhóm tuổi trẻ và già có khác nhau. Vào năm 2009 nhóm tuổi trẻ có tỷ lệ hoạt động kinh tế thấp hơn so với năm 1989. Điều này được giải thích là nhờ sự

phát triển mạnh về kinh tế các thế hệ trẻ ngày nay được cha mẹ cho đi học nhiều hơn. Phân tích trên cho phép dự đoán là vào các năm sau này mức độ và dạng tỷ lệ hoạt động kinh tế sẽ tương tự như ở năm 2009.

Với giả thiết mức độ và dạng phân bố của dân số hoạt động kinh tế và dân số đang làm việc của các năm 2014 và 2019 giống như năm 2009 và dựa vào kết quả dự báo dân số của hai năm này, tiến hành dự báo dân số hoạt động kinh tế và dân số đang làm việc của hai giới tính theo từng nhóm tuổi. Trên cơ sở tỷ lệ dân số đang làm việc của năm 2009, tính được số người đang làm việc vào hai năm 2014 và 2019 như sau:

Bảng 3: Dân số đang làm việc theo giới tính và nhóm tuổi năm 2014 và 2019

(Nghìn người)

	2014			2019		
	Nam	Nữ	Tổng	Nam	Nữ	Tổng
15-19	1719,6	1425,0	3144,6	1594,3	1316,3	2910,6
20-24	3562,5	3037,0	6599,5	2937,8	2464,9	5402,8
25-29	3864,4	3257,0	7121,4	4111,8	3427,3	7539,1
30-34	3796,0	3239,9	7035,9	3896,5	3280,0	7176,5
35-39	3286,6	2871,6	6158,2	3705,2	3162,5	6867,6
40-44	3056,6	2730,9	5787,5	3169,2	2770,4	5939,6
45-49	2694,1	2461,6	5155,8	2879,7	2551,3	5431,0
50-54	2223,0	2065,6	4288,6	2395,5	2122,0	4517,6
55-59	1557,7	1417,5	2975,2	1844,2	1569,5	3413,7
60+	1321,6	1280,3	2601,9	1578,1	1417,4	2995,5
Tổng	27082,2	23786,5	50868,6	28112,2	24081,7	52193,9

5.3. Dự báo tỷ lệ và số người thất nghiệp

5.3.1. Dự báo tỷ lệ thất nghiệp

Để dự báo tỷ lệ thất nghiệp cho các năm, chúng tôi dựa vào dãy số liệu thống kê về tỷ lệ thất nghiệp của lực lượng lao động do TCTK thu thập cho giai đoạn 1996-2009. Phương pháp dự báo ở đây là tìm dạng hàm số toán phù hợp mô tả sự biến động của tỷ lệ thất nghiệp theo thời gian, sau đó dùng nó để

ngoại suy cho tương lai. Hàm số được xem là mô tả phù hợp sự biến động theo thời gian của tỷ lệ thất nghiệp là hàm số mũ bậc hai. Dựa vào số liệu thực tế chúng tôi ước lượng được hàm này như sau:

$y_t = 6,01 + 0,186t - 0,024t^2$, với $R = 0,935$; $R^2 = 0,874$; sai số tuyệt đối bình quân bằng 0,253 đơn vị %; và sai số tương đối bằng 4,4%.

Kết quả tính toán cho thấy hệ số tương quan R giữa tỷ lệ thất nghiệp (y) và thời gian (t) rất cao. Mặt khác bình phương của hệ số tương quan đạt mức 87,4% (tính theo phần trăm). Điều này có nghĩa là biến thời gian t mô tả được tới 87,4% sự biến động của tỷ lệ thất nghiệp hàng năm.

Kết quả dự đoán cho thấy nếu tình hình kinh tế vẫn như những năm trước đây, tỷ lệ thất nghiệp của Việt Nam sẽ tiếp tục giảm. Tuy nhiên, xu thế này sẽ không còn giữ vững được nữa vì năm 2008 xảy ra khủng hoảng tài chính thế giới và như vậy giả thiết “tình hình kinh tế vẫn như những năm trước” bị vi phạm dẫn đến sẽ có sự gia tăng tỷ lệ thất nghiệp vào năm 2009 và ít nhất là vẫn tăng vào năm 2010.

5.3.2. Dự báo số người thất nghiệp

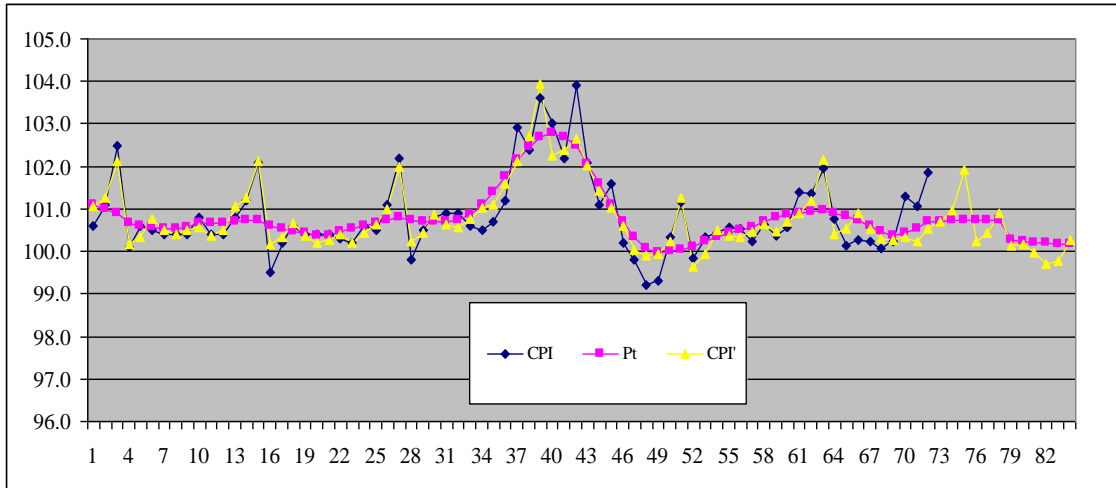
Ngoài việc sử dụng dãy số thời gian để dự báo tỷ lệ thất nghiệp hàng năm của dân số, đề tài còn sử dụng phương pháp thành phần để dự báo dân số thất nghiệp theo giới tính và nhóm tuổi. Ở phương pháp dự báo này giả thiết là tỷ lệ thất nghiệp theo giới tính và nhóm tuổi vào các năm 2014 và 2019 tương tự như vào năm 2009. Theo kết quả dự báo, năm 2014 và 2019 mỗi năm có khoảng 1,6 triệu người thất nghiệp, trong đó số người thất nghiệp ở độ tuổi từ 20-29 chiếm khoảng một phần ba tổng số người thất nghiệp.

5.4. Kết quả dự báo tỷ lệ lạm phát

Lạm phát là một chỉ tiêu thống kê vừa mang tính chất kinh tế vừa mang tính chất xã hội. Tỷ lệ lạm phát cao, kéo dài sẽ làm cho nền kinh tế suy sụp. Mặt khác nó cũng kéo theo hệ quả xã hội to lớn. Trong đề tài này tiến hành dự báo tỷ lệ lạm phát thông qua chỉ số giá tiêu dùng. Phương pháp dự báo được sử dụng là phương pháp gia quyền điều hòa, một phương pháp thay đổi uyển chuyển theo sự thay đổi của hiện tượng được nghiên cứu và đưa ra giả thuyết là tốc độ tăng ở thời kỳ hiện tại ảnh hưởng lớn nhất tới tốc độ tăng của hiện tượng trong tương lai.

Kết quả tính toán dự báo cho thấy tỷ lệ lạm phát thường cao hơn mức trung bình của các tháng rơi vào 3 tháng đó là tháng 1, tháng 2 và tháng 5. Kết quả tính toán trong bảng cũng cho thấy vào tháng 1 năm 2011 tỷ lệ lạm phát (so với tháng 12 năm 2010) là 0,96%, còn vào tháng 2 năm 2011 tỷ lệ lạm phát (so với tháng 1 năm 2011) là 1,93%.

Hình 1: Tỷ lệ lạm phát



5.5. Kết quả dự báo thu nhập bình quân một đầu người

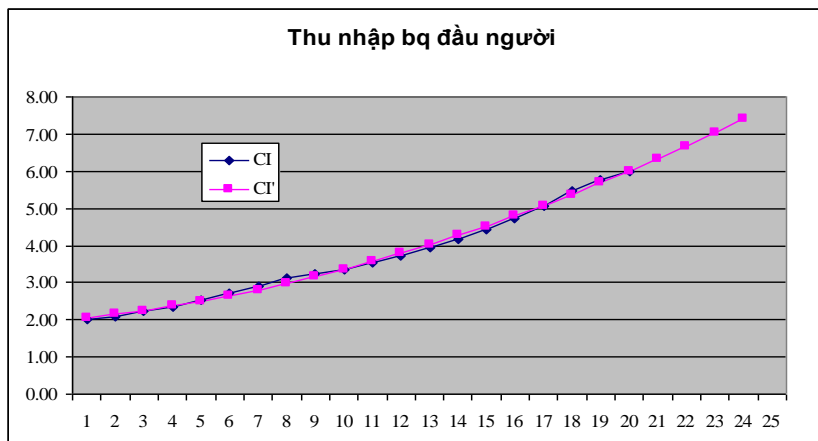
5.5.1. Dự báo thu nhập bình quân đầu người theo dãy số thời gian

Dựa vào dãy số liệu thống kê về thu nhập bình quân một người (CI) trong giai đoạn 1990-2009. Ta tiến hành dự báo bằng phương pháp dãy số thời gian cho các năm 2010, 2011 và 2012.

$$CI'_t = 1,9625 + 0,077t + 0,0062t^2$$

$R = 0,998$; $R^2 = 0,996$ Sai số tuyệt đối bình quân bằng 0,055 đơn vị; sai số tương đối bằng 1,67%.

Hình 2: Đồ thị thu nhập bình quân một đầu người



Sử dụng chương trình SPSS đã xác định được hàm số bậc hai mô tả tốt nhất quy luật phát triển theo thời gian của chỉ tiêu thu nhập bình quân một người. Điều này thể hiện rất rõ thông qua hệ số tương quan R và hệ số mô tả R^2 ước lượng được. Hệ số tương quan thu được bằng $R=0,998$. Kết quả này cho thấy giữa chỉ tiêu thu nhập bình quân một người và biến thời gian có mối quan hệ rất chặt. Hệ số mô tả thu được: $R^2 = 0,996$. Điều này cho thấy biến thời gian mô tả được tới 99,3% sự biến động của chỉ tiêu thu nhập bình quân một người. Hơn thế nữa, sai số dự báo tuyệt đối và tương đối đều ở mức độ thấp (các con số tương ứng là 0,073 đơn vị tính và 1,98%). Như vậy, ta có thể kết luận là hàm số bậc hai phù hợp với việc mô tả sự biến động theo thời gian của chỉ tiêu thu nhập bình quân một người và có thể yên tâm sử dụng nó để dự báo.

Theo kết quả tính toán, thu nhập bình quân một người của các năm 2010, 2011 và 2012 tương ứng sẽ là 6,33 triệu, 6,67 triệu và 7,03 triệu đồng. Với mức độ tin cậy là 95%, thu nhập bình quân một người của các năm này sẽ nằm trong các khoảng sau: năm 2010 - (6,12;6,53); năm 2011 - (6,45;6,90); và năm 2012- (6,79; 7,27).

5.5.2. Dự báo thu nhập bình quân một đầu người theo phương pháp kết hợp giữa dự báo thu nhập (GDP) và dự báo dân số (POP)

Khi tiến hành dự báo theo phương pháp này, tiến hành dự báo riêng rẽ cho chỉ tiêu GDP và tổng số dân (POP) sau đó kết hợp với nhau để dự báo cho chỉ tiêu thu nhập bình quân một người. Bảng dưới đây trình bày kết quả dự báo thu được.

Bảng 4: Dự báo thu nhập bình quân đầu người thông qua dự báo hai chỉ tiêu GDP và dân số

	GDP (nghìn tỷ đồng)	POP (000)	GDP'	POP'	Thu nhập bq 1 người
A	(1)	(2)	(3)	(4)	(5)
1990	131968	66017	133206	66046	2,02
1991	139634	67242	142805	67253	2,12
1992	151782	68450	153484	68450	2,24
1993	164043	69645	165242	69637	2,37
1994	178534	70825	178079	70814	2,51
1995	195567	71996	191994	71979	2,67
1996	213833	73157	206989	73135	2,83
1997	231264	74307	223063	74280	3,00

1998	244596	75456	240216	75415	3,19
1999	256272	76597	258447	76539	3,38
2000	273666	77635	277758	77653	3,58
2001	292535	78686	298148	78757	3,79
2002	313247	79727	319617	79850	4,00
2003	336242	80902	342165	80933	4,23
2004	362435	82032	365792	82005	4,46
2005	393031	83106	390497	83068	4,70
2006	425135	84156	416282	84119	4,95
2007	.	.	443146	85160	5,20
2008	.	.	471089	86191	5,47
2009	.	.	500111	87212	5,73

Nguồn số liệu: Tổng cục Thống kê, Niên giám thống kê các năm

Cột (1) và (2) của bảng 4 trình bày số liệu thực tế của hai chỉ tiêu GDP và dân số trung bình trong giai đoạn 1990-2006. Cột (3) và (4) trình bày giá trị của hai chỉ tiêu này được ước lượng dựa vào mô hình dự báo. Kết quả của các năm 2007, 2008 và 2009 là kết quả dự báo cho hai chỉ tiêu đó. Cột (5) là giá trị của chỉ tiêu thu nhập bình quân một đầu người được tính dựa vào giá trị được tính từ mô hình dự báo của hai chỉ tiêu GDP và dân số. Giá trị của chỉ tiêu này ở các năm 2007, 2008 và 2009 cũng là các giá trị dự báo.

So sánh giá trị dự báo của chỉ tiêu thu nhập bình quân đầu người bằng phương pháp trực tiếp và phương pháp kết hợp thấy không khác nhau là mấy.

Bảng 13: Kết quả dự báo chỉ tiêu thu nhập bình quân đầu người bằng hai phương pháp

	Phương pháp trực tiếp	Phương pháp kết hợp	Độ lệch giữa hai phương pháp
2007	5,21	5,20	0,010
2008	5,48	5,47	0,010
2009	5,76	5,73	0,030

5.5.3. Dự báo thu nhập bình quân đầu người theo phương pháp hồi quy đa biến

Dựa vào số liệu thu được từ cuộc điều tra khảo sát mức sống hộ gia đình 2004, chúng tôi xây dựng mô hình dự báo cho thu nhập bình quân thực tế của các hộ gia đình. Các bước được tiến hành như sau:

1) Phân tích định tính để lựa chọn các yếu tố có quan hệ với thu nhập của các hộ gia đình; 2) Tính ma trận tương quan giữa các yếu tố này (bao gồm cả chỉ tiêu thu nhập bình quân của hộ gia đình); 3) Lựa chọn các yếu tố để đưa vào mô hình; 4) Ước lượng các thông số của mô hình.

Qua phân tích định tính và dựa vào ma trận hệ số tương quan chúng tôi đã chọn các biến sau đây để xây dựng mô hình: Chi bình quân thực tế (CHIBQTTE), số năm học bình quân của những người từ 15 tuổi trở lên (NAMHBQ15), tỷ lệ trẻ em trong hộ (TLETREEM) và số giờ làm việc trong tuần (GIOBQTUA).

Qua khảo nghiệm thấy biến thu nhập bình quân của các hộ gia đình (THUBQTTE) cần được log hóa để quan hệ của nó có dạng tuyến tính, vì vậy mô hình có dạng:

$Ln(THUBQTTE) = a_0 + a_1CHIBQTTE + a_2NAMHOCBQ15 + a_3TLETREEM + a_4GIOBQTUA$
 Sử dụng bộ số liệu của cuộc điều tra mức sống hộ gia đình năm 2004, chúng tôi đã ước lượng được các hệ số của mô hình như sau:

Bảng 5: Hệ số của mô hình và mức ý nghĩa

	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	5.061	.016		322.917	.000
CHIBQTTE	.001	.000	.593	77.617	.000
NAMHBQ15	.038	.002	.187	23.267	.000
TLETREEM	-.288	.024	-.090	-12.052	.000
GIOBQTUA	.006	.000	.122	16.986	.000

a Dependent Variable: LNTHUBQT

Các thông số của mô hình được ước lượng ở cấp vi mô và hệ số mô tả của mô hình ở mức 75,6%. Ở cấp vi mô mà hệ số mô tả như vậy là cao. Mô hình có thể sử dụng để dự báo cho chỉ tiêu thu nhập bình quân của các hộ. Mô hình có dạng cụ thể là:

$Ln(THUBQTTE) = 5,061 + 0,001CHIBQTTE + 0,038NAMHOCBQ15 - 0,288TLETREEM + 0,006GIOBQTUA$
 Từ mô hình trên, nếu biết được mức chi tiêu bình quân của một hộ, số năm đi học bình quân của những người từ 15 tuổi trở lên, tỷ lệ trẻ em trong hộ và số giờ làm việc trong 1 tuần của chủ hộ ta sẽ dự báo được mức thu nhập của hộ này là bao nhiêu.

KẾT LUẬN VÀ KIẾN NGHỊ

I. Kết luận

1) Có thể sử dụng các mô hình kinh tế trắc lượng để dự báo cho một số chỉ tiêu thống kê kinh tế xã hội nói chung và các chỉ tiêu thống kê xã hội nói riêng;

2) Có các phương pháp đặc thù được sử dụng để dự báo các chỉ tiêu về dân số, lao động và thất nghiệp;

3) Phương pháp dự báo theo dãy số thời gian áp dụng có hiệu quả trong việc dự báo một số chỉ tiêu thống kê xã hội;

4) Còn nhiều chỉ tiêu thống kê xã hội không có được số liệu đầy đủ nên không thể tiến hành dự báo cho chúng được;

5) Còn nhiều chỉ tiêu thống kê xã hội không được thu thập và lưu trữ hệ thống nên cũng không thể tiến hành dự báo cho chúng;

6) Quá trình tiến hành dự báo cho thấy số liệu thống kê vẫn còn có hạn chế, vì vậy cần tăng cường hơn nữa các biện pháp để nâng cao chất lượng của chúng.

7) Quá trình tiến hành dự báo cũng cho thấy để chất lượng dự báo được nâng cao, trước khi tiến hành dự báo cần phải đánh giá chất lượng số liệu và tiến hành hiệu chỉnh nếu thấy cần thiết.

Có thể áp dụng phương pháp dự báo ngắn hạn cho các chỉ tiêu sau đây:

Số thứ tự	tiêu	Mức độ ưu tiên (1, 2)
01.		
1		1
2		1
3		1
4		1
5		1
6		1
7		1
8	Tỷ lệ lao động đang làm việc so với tổng dân số	1
9		1
10		1
11	Tỷ lệ thất nghiệp	1

Số thứ tự	tiêu	Mức độ ưu tiên (1, 2)
02.	Thông tin	
1	100 người dân	2
2		2
03.		
1		2
2		2
3		2
4		2
5		2
6		2
7	Số học sinh phổ thông bình quân 1 giáo viên	2
8	Số học sinh phổ thông bình quân 1 lớp học	2
9		2
10		2
11		2
12		2
13		2
14		2
15		2
16		2
17		2
18		2
19		2
20		2
04.		
1		1
2		1
3	quân 10.000 người dân	1
4		1
5		1
05.		2
1		2
06.		
1		1
2	Tỷ lệ hộ gia đình có một số đồ dùng lâu bền	2

II. Khuyến nghị

Với các kết luận trên, ban chủ nhiệm đề tài có các khuyến nghị sau:

1) Tiếp tục áp dụng các phương pháp đặc thù của lĩnh vực dân số để dự báo dân số và cơ cấu dân số cho tương lai. Tuy nhiên trước khi tiến hành dự báo cần sử dụng tỷ lệ giới tính theo độ tuổi để đánh giá số liệu đầu vào. Cần có những hiệu chỉnh số liệu hợp lý để kết quả dự báo sát với thực tế hơn;

2) Nên sử dụng phương pháp dự báo dân số thành phần để dự báo lực lượng lao động và số người thất nghiệp trong tương lai.

3) Đề nghị triển khai áp dụng các phương pháp dự báo ngắn hạn cho các chỉ tiêu đã trình bày ở bảng trên. Tuy nhiên, để dự báo được hết các chỉ tiêu trong bảng cần tiến hành bổ sung và hoàn thiện số liệu;

4) Để dự báo được triển khai có kết quả và cho nhiều chỉ tiêu cần củng cố mạng lưới thống kê ở các bộ, các ngành và tổ chức lưu trữ một cách có hệ thống, đồng bộ các số liệu thống kê.

III. Đề xuất lộ trình thực hiện

Để dự báo các chỉ tiêu xã hội có thể đưa vào ứng dụng thành công trong thực tế, đề tài đề xuất một lộ trình như sau:

a) Trong năm 2011, 2012 nên tiến hành dự báo các chỉ tiêu đề tài đã thực hiện dự báo và cho rằng có kết như các chỉ tiêu thuộc lĩnh vực dân số, lao động, chỉ tiêu về lạm phát (CPI),...

b) Cũng trong các năm 2011, 2012 triển khai hoàn thiện các cơ sở dữ liệu cho các chỉ tiêu có mức độ ưu tiên bằng 1 để năm 2013 tiến hành dự báo cho chỉ tiêu này nhằm mở rộng thêm phạm vi dự báo.

c) Bên cạnh việc triển khai dự báo, cần đẩy mạnh việc hỗ trợ cho các cơ quan chức năng hoàn thiện công tác thống kê nhằm có được các thông tin sát thực giúp cho công tác dự báo được mở rộng và có chất lượng.

TÀI LIỆU THAM KHẢO

1. Bates, J.M., C.W.J Granger (1969), "The combination of forecasts", *Operations Research Quarterly* 20: 451 -468.
2. Box, G.E.P and G.M. Jenkins (1976). *Time Series Analysis , Forecasting and Control* . San Francisco, Holden- Day , Inc
3. Brillinger, D.R.(1981), *Time Series: Data Analysis and Theory*, expanded edition (Holden - Day, San Francisco).
4. Clements M.P, and D.F. Hendry (1999), *Forecasting Non - stationary Economic Time Series* (MIT Press, Cambridge, MA).
5. Clemen, R.T (1989), "Combining forecasts: a review and annotated bibliography", *International Journal of Forecasting* 5: 559 - 583.
6. Diebol, F.X., and P. Pauly (1987), "Structural change and the combination of forecasts", *Journal of Forecasting* 6:21- 40.
7. Diebold, F.X., and J.A. Lopez (1996), "Forecast evaluation and combination", in: G.S.
8. Figlewski, S. (1983), "Optimal price forecasting using survey data", *Review of Economics and Statistics* 65: 813- 836.
9. Geweke, J. (1977), "The Dynamic Factor Analysis of Economic Time Seres", in: D.J. Aiger and A.S. Goldberger, eds., *Latent Variables in Socio-Economic Models*, (North-Holland, Amsterdam).
10. Granger, C.W.J and P. Newbold (1976). *Forecasting Economic Time Series*, Academic Press: New York
11. Green, W.H. (1993). *Econometric Analysis*, Macmillan Publishing Company.
12. Hendry, D.F., and M.P. Clements (2002), "Pooling of Forecasts", *Econometrics journal* 5: 1- 26.
13. Intriligator ,M.D., R.G. Bodkin ., Cheng Hsiao (1971). *Econometric Models, Techniques, and Applications*. Prentice-Hall International, Inc.
14. McNees,S.K.(1988). On the future of Macroeconomic Forecasting." *International Journal of Forecasting*, 4, 359-362.
15. Newbold, P., and D.I. Harvey (2002), "Forecast combination and encompassing", in: M.P. Clements and D.F. Hendry, eds., *C Companion to Econmic Forecasting* (Blackwll Press: Oxford) 268- 283.
16. Nguyễn Khắc Minh (2001). *Phân tích và dự báo tài chính*, NXB
17. Nguyễn Khắc Minh (2002). *Phương pháp kinh tế lượng trong phân tích và dự báo - Những điều cần chú ý*," *Tạp chí kinh tế & Phát triển*.58

18. Nguyễn Khắc Minh (2003). Các phương pháp phân tích và dự báo trong kinh tế, NXB KH&KT.
19. Stock, J.H., and M.W. Watson (2004a), "Combination forecasts of output growth in a seven - country data" , forthcoming, Journal of Forecasting.
20. Stock, J.H., and M.W. Watson (2004b), "An empirical comparison of methods for forecasting using many predictors", manuscript.
21. Timmermann, Alan (2006). " Forecast Combinations," in Handbook of Economic Forecasting, in Graham Elliott, Clive W.J. Granger and Alla Timmermann (eds) Elsevier.
22. Kết quả nghiên cứu đề tài dự báo ngắn hạn các chỉ tiêu thống kê kinh tế năm 2006-2007.