PHƯƠNG PHÁP NGHIÊN CỨU

1.1. Những vấn đề chung về dãy số thời gian

1.1.1. Khái niệm

Dãy số thời gian (DSTG) là dãy các trị số của chỉ tiêu thống kê được sắp xếp theo thứ tự thời gian. DSTG cho phép thống kê nghiên cứu xu hướng biến động của hiện tượng qua thời gian. Từ đó, tìm ra tính quy luật của sự phát triển đồng thời dự đoán được các mức độ của hiện tượng trong tương lai.

Một dãy số thời gian bao giờ cũng có hai bộ phận: thời gian và chỉ tiêu của hiện tượng nghiên cứu.

Thời gian có thể là ngày, tuần, tháng, quý, năm. Độ dài giữa hai thời gian liền nhau gọi là khoảng cách thời gian.

Chỉ tiêu của hiện tượng nghiên cứu bao gồm tên chỉ tiêu với đơn vị tính phù hợp và trị số của chỉ tiêu được sắp xếp theo thời gian (được gọi là các mức độ của dãy số thời gian), ký hiệu là yi (i = 1, 2,..., n).

1.1.2. Các thành phần của dãy số thời gian

Xu thế (T) phản ánh xu hướng biến động cơ bản của hiện tượng qua thời gian.

Biến động chu kỳ (C) phản ánh quy luật lặp lại của dãy số trong những khoảng thời gian nhất định, thường là vài năm.

Biến động thời vụ (S) là những biến động của hiện tượng nhất định có tính chất lặp đi lặp lại trong từng thời gian nhất định của năm.

Biến động ngẫu nhiên (I) là do các yếu tố ngẫu nhiên gây ra. Loại biến động này thường rất khó dự đoán do tính chất bất thường của nó.

1.1.3. Yêu cầu của dãy số thời gian

Để phân tích dãy số thời gian được chính xác thì yêu cầu cơ bản khi xây dựng dãy số thời gian là phải đảm bảo tính chất có thể so sánh được giữa các mức độ trong dãy số. Yêu cầu này được thể hiện trên ba điểm cụ thể là:

- Nội dung và phương pháp tính chỉ tiêu doanh thu qua thời gian phải được thống nhất.

- Phạm vi của doanh thu nghiên cứu qua thời gian phải được thống nhất.
- Các khoảng cách thời gian trong dãy số nên bằng nhau, nhất là đối với dãy số thời kỳ.

Trong thực tế, do nhiều nguyên nhân khác nhau, các yêu cầu trên có thể bị vi phạm. Do đó, trước khi tiến hành phân tích, cần có sự đánh giá và chỉnh lý dãy số cho phù hợp với các yêu cầu trên.

1.1.4. Các loại dãy số thời gian

Dãy số tuyệt đối: khi các mức độ của dãy số là tuyệt đối. Dãy số tuyệt đối bao gồm:

Dãy số tuyệt đối thời điểm: là các dãy số mà mức độ của nó là những số tuyệt đối thời điểm, phản ánh quy mô (khối lượng) tại những thời điểm nhất định. VD: giá trị tồn kho đầu mỗi quý.

Dãy số tuyệt đối thời kỳ: là dãy số mà các mức độ của nó là những số tuyệt đối thời kỳ, phản ánh quy mô (khối lượng) của hiện tượng được tích luỹ trong khoảng thời gian nhất định. VD: kết quả sản xuất kinh doanh trong từng năm.

Dãy số tương đối: khi các mức độ của dãy số là số tương đối. Ví dụ: tốc độ phát triển doanh thu của công ty.

Dãy số bình quân: khi các mức độ của dãy số là số bình quân. Ví dụ: tiền lương bình quân của lao động trong công ty .

1.1.5. Tác dụng của dãy số thời gian

Phân tích đặc điểm và tính quy luật, sự biến động của hiện tượng qua thời gian.

Dự đoán sự phát triển của hiện tượng trong tương lai.

1.2. Các chỉ tiêu phân tích đặc điểm biến động của hiện tượng qua thời gian

1.2.1. Mức độ bình quân theo theo thời gian

Mức độ bình quân theo thời gian là mức độ đại diện cho các mức độ tuyệt đối của một dãy số thời gian.

Đối với dãy số thời kỳ: mức độ bình quân theo thời gian được tính theo công thức.

$$\bar{y} = \frac{y_1 + y_2 + \cdots + y_{n-1} + y_n}{n}$$

Đối với dãy số thời điểm:

Dãy số thời điểm biến động đều và chỉ có 2 mức độ đầu kỳ và cuối kỳ:

$$\bar{y} = \frac{y_{đk} + y_{ck}}{2}$$

Dãy số thời điểm biến động không đều, có nhiều mức độ mà khoảng cách thời gian bằng nhau:

$$\bar{y} = \frac{\frac{y_1}{2} + y_2 + \cdots + y_{n-1} + \frac{y_n}{2}}{n}$$

Trong đó, yi (i = 1,2,…,n) là các mức độ của dãy số thời điểm có khoảng cách thời gian bằng nhau

Dãy số thời điểm có khoảng cách thời gian không bằng nhau:

$$\bar{y} = \frac{\sum y_i h_i}{\sum h_i}$$

Trong đó hi (i=1,2,...,n) là khoảng thời gian có mức độ yi (i=1,2,..,n)

1.2.2. Lượng tăng (giảm) tuyệt đối

Lượng tăng (giảm) tuyệt đối là chỉ tiêu phản ánh sự biến động về mức độ tuyệt đối của hiện tượng giữa hai thời gian.

Lượng tăng giảm tuyệt đối liên hoàn là chỉ tiêu phản ánh biến động về mức độ tuyệt đối của hiện tượng giữa 2 thời gian liền nhau:

$$\delta_i = y_i - y_{i-1} \quad (i = \overline{2,n})$$

Nếu δi > 0 phản ánh quy mô hiện tượng tăng, ngược lại nếu δi < 0 phản ánh quy mô hiện tượng giảm.

Lượng tăng giảm tuyệt đối định gốc là chỉ tiêu phản ánh sự biến động về mức độ tuyệt đối của hiện tượng trong những khoảng thời gian dài và thường lấy mức độ đầu tiên làm gốc cố định:

$$\Delta_i = yi - y_1 \qquad (i = \overline{2, n})$$

Mối liên hệ giữa lượng tăng (giảm) tuyệt đối liên hoàn và định gốc:

$$\delta_2 + \delta_3 + \dots + \delta_n = \Delta_n = y_n - y_1$$

Lượng tăng giảm tuyệt đối trung bình là chỉ tiêu bình quân của các lượng tăng(giảm) tuyệt đối liên hoàn của dãy số trong cả kỳ nghiên cứu:

$$\overline{\delta} = \frac{\delta_1 + \delta_2 + \dots + \delta_n}{n - 1} = \frac{\Delta_n}{n - 1} = \frac{y_n - y_1}{n - 1}$$

1.2.3. Tốc độ phát triển

Tốc độ phát triển là chỉ tiêu phản ánh xu hướng và tốc độ biến động của hiện tượng nghiên cứu qua thời gian, được tính bằng cách chia mức độ của hiện tượng ở kỳ nghiên cứu cho mức độ của hiện tượng ở kỳ gốc.

Tốc độ phát triển liên hoàn là chỉ tiêu phản ánh xu hướng và tốc độ biến động của hiện tượng giưa hai thời gian liền nhau.

$$t_i = \frac{y_i}{y_{i-1}} \qquad (i = \overline{2, n})$$

Trong đó ti là tốc độ phát triển liên hoàn thời gian I so với thời gian i-1 và có thể biểu hiện bằng lần hoặc %.

Tốc độ phát triển định gốc là chỉ tiêu phản ánh tốc độ và xu hướng biến động của hiện tượng ở những khoảng thời gian dài, được tính bằng cách so sánh mức dộ của hiện tượng ở kỳ nghiên cứu với mức độ ở kỳ được chọn làm gốc so sánh cố định (thường chọn kỳ đầu tiên).

$$Ti = \frac{y_i}{y_1} \qquad (i = \overline{2, n})$$

Trong đó Ti là tốc độ phát triển định gốc thời gian i so với thời gian đầu của dãy số và có thể biểu hiện bằng lần hoặc %.

Mối quan hệ giữa tốc độ phát triển liên hoàn và định gốc:

Tích các tốc độ phát triển liên hoàn bằng tốc độ phát triển định gốc tương ứng:

$$t_2 \text{ x } t_3 \text{x} \dots \text{x } t_n = T_n$$

(2) Thương của tốc độ phát triển định gốc ở thời gian i với tốc độ phát triển định gốc ở thời gian i – 1 bằng tốc độ phát triển liên hoàn giữa hai thời gian đó, tức là:

$$\frac{T_i}{T_{i-1}} = t_i \qquad \text{(với i=2, 3, …, n)}$$

Tốc độ phát triển trung bình là chỉ tiêu bình quân của các tốc độ phát triển liên hoàn trong cả kỳ nghiên cứu.

$$\bar{t} = \sqrt[n-1]{t_2 \cdot t_3 \dots t_n} = \sqrt[n-1]{T_n} = \sqrt[n-1]{\frac{y_n}{y_1}}$$

1.2.4. Tốc độ tăng (giảm)

Tốc độ tăng (giảm) là chỉ tiêu phản ánh nhịp độ tăng (giảm) tương đối giữa các mức độ của hiện tượng qua thời gian.

Tốc độc tăng (giảm) liên hoàn là chỉ tiêu phản ánh nhịp độ tăng (giảm) tương đối của hiện tượng giữa hai thời gian liền nhau.

$$a_i = \frac{\delta_i}{y_{i-1}} = \frac{y_j - y_{i-1}}{y_{i-1}} = t_{i-1} \qquad (i = \overline{2, n})$$

Tốc độ tăng (giảm) định gốc là chỉ tiêu phản ánh nhịp độ tăng (giảm) tương đối của hiện tượng giữa hai thời gian và thường lấy mức độ đầu làm gốc cố định.

$$A_i = \frac{\Delta_i}{y_1} = \frac{y_i - y_1}{y_1} = T - 1 \qquad (i = (\overline{2, n})$$

Tốc độ tăng (giảm) bình quân là chỉ tiêu phản ánh nhịp độ tăng (giảm) đại diện cho các tốc độ tăng (giảm) liên hoàn

$\bar{a} = \bar{t} - 1$ (nếu $\bar{t}$ biểu hiện bằng lần)

$\bar{a} = \bar{t} - 100$ (nếu $\bar{t}$ biểu hiện bằng %)

1.2.5. Giá trị tuyệt đối của 1% tăng (giảm) liên hoàn

Giá trị tuyệt đối của 1% tăng (giảm) liên hoàn là chỉ tiêu phản ánh cứ 1% tốc độ tăng (giảm) liên hoàn thì tương ứng hiện tượng nghiên cứu tăng thêm (hoặc giảm đi) một lượng là bao nhiêu.

$$g_i = \frac{\delta_i}{a_i(\%)} = \frac{\delta_i}{\frac{\delta_i}{y_{i-1}} \cdot 100} = \frac{y_{i-1}}{100}\ (với\ i = 2,3, \dots n)$$

Cần chú ý là chỉ tiêu này không tính đối với tốc độ tăng (giảm) định gốc vì nó luôn là một số không đổi và bằng $\frac{y_1}{100}$.

1.3. Các phương pháp biểu hiện xu hướng biến động cơ bản của dãy số thời gian

1.3.1. Phương pháp mở rộng khoảng cách thời gian

Nội dung: Mở rộng thêm khoảng cách thời gian bằng cách ghép một số thời gian liền nhau vào thành 1 khoảng thời gian dài hơn. Ví dụ: Mở rộng khoảng cách thời gian từ tháng thành quý, từ quý thành năm,…

Mục đích là để từ dãy số không có hoặc chưa thể hiện rõ tính quy luật thành dãy số xuất hiện tính quy luật (triệt tiêu ngẫu nhiên để biểu hiện xu hướng).

Vận dụng: Mở rộng khoảng cách thời gian được vận dụng với dãy số thời kỳ có khoảng cách thời gian tương đối ngắn, nhiều mức độ và chưa thấy rõ được xu hướng phát triển cơ bản của doanh thu. Thời gian dài – ngắn mang ý nghĩa tương đối, phụ thuộc vào đặc điểm của doanh thu và từng loại chỉ tiêu khác nhau

Hạn chế: Do ghép nhiều khoảng thời gian vào thành một nên số lượng các mức độ trong dãy số mất đi quá nhiều, đôi khi làm mất ảnh hưởng của các nhân tố cơ bản. Trường hợp sử dụng với những hiện tượng có tính chất thời vụ sẽ làm mất đi tính chất thời vụ của hiện tượng.

1.3.2. Phương pháp dãy số bình quân trượt

Khái niệm: Là phương pháp tính giá trị bình quân cho một nhóm các mức độ nhất định của dãy số bằng cách loại dần các mức độ đầu và thêm vào đó các mức độ tiếp theo sao cho tổng số lượng mức độ tham gia vào tính số bình quân là không thay đổi.

Giả sử có dãy số thời gian $y_1, y_2, \dots, y_n$.

Nếu tính số bình quân trượt cho ba nhóm mức độ, ta có:

$$\overline{y_2} = \frac{y_1 + y_2 + y_3}{3}$$

$$\overline{y_3} = \frac{y_2 + y_3 + y_4}{3}$$

$$\dots$$

$$\bar{y}_{n-1} = \frac{y_{n-2} + y_{n-1} + y_n}{3}$$

Từ đó ta có dãy số mới gồm các số bình quân trượt: $\overline{y_2}, \overline{y_3}, \dots, \overline{y_{n-1}}$

Tùy vào trong trường hợp cụ thể để chọn số lượng mức độ tham gia vào tính số bình quân trượt. Khi tính số bình quân trượt thì giá trị của nó phải đặt vào vị trí giữa của nhóm các mức độ tham gia để tính toán. Khi áp dụng tính doanh thu của công ty trong khoảng thời gian 2011-2021 (11 năm) với số năm lẻ thì dễ thực hiện hơn nếu số thời gian áp dụng là chẵn thì chúng ta phải tính trượt hai lần.

Vận dụng: Với dãy số thời kỳ theo tháng, quý, năm nhưng không có yếu tố thời vụ.

Ưu điểm: So với mở rộng khoảng cách thời gian thì số lượng các mức độ trong dãy số mất đi ít hơn, khi biểu diễn trên đồ thị sẽ thấy xu hướng rõ ràng hơn.

Hạn chế: Trong trường hợp sử dụng với những hiện tượng có tính chất thời vụ sẽ làm mất đi tính chất thời vụ của hiện tượng.

1.3.3. Phương pháp san bằng mũ

Phương pháp san bằng mũ loại bỏ các biến động ngẫu nhiên giúp làm trơn dãy số thời gian theo mô hình sau đây:

$$S_t = \alpha \cdot y_t + (1 - \alpha) \cdot S_{t-1} \qquad (t \geq 2)$$

Trong đó:

S_t là giá trị san bằng mũ của dãy số thời gian t.

y_t là mức dộ của dãy số ở thời gian t.

S_{t-1} là giá trị san bằng mũ của dãy số thời gian ở thời gian t-1.

α là hệ số san bằng mũ với $0 \leq \alpha \leq 1$.

Việc lựa chọn hệ số san bằng mũ α hoàn toàn là chủ quan, do người phân tích dữ liệu quyết định. Giá trị α càng nhỏ, dãy số mới càng trơn nhẵn giúp ta hình dung được xu hướng tăng lên hay giảm đi của doanh thu qua thời gian nhanh hơn. Ngược lại giá trị α càng cao xu hướng biến động cơ bản của doanh thu sẽ chậm hơn.

1.3.4. Phương pháp xây dựng hàm xu thế

Phương pháp hồi quy trong dãy số thời gian được vận dụng để biểu diễn xu hướng phát triển cơ bản của những hiện tượng có nhiều dao động ngẫu nhiên. Khi đó, người ta xây dựng một hàm số (gọi là phương trình hồi quy) nhằm phản ánh biến động của hiện tượng theo thời gian.

Hàm số này có dạng tổng quát: $\widehat{y_i}$ = f (t_i) và thường được gọi là hàm xu thế.

Trong đó:

t_i là biến thời gian, là thứ tự thời gian theo quy ước, đóng vai trò là biến số độc lập trong phương trình hồi quy.

$$-\infty < t < +\infty$$

$\widehat{y_i}$: Mức độ của doanh thu ở thời gian t tính từ hàm xu thế.

Các dạng hàm xu thế thường được sử dụng:

Hàm xu thế tuyến tính: Sử dụng khi dãy số thời gian có các lượng tăng (giảm) tuyệt đối liên hoàn doanh thu xấp xỉ nhau.

Hàm có dạng: $\widehat{y_i} = b_0 + b_1 t_i$

Các tham số b_0 , b_1 được xác định bằng phương pháp bình phương nhỏ nhất. Theo đó, b_0 và b_1 phải thoả mãn phương trình:

$$\begin{cases} \sum y_i = n \cdot b_0 + b_1 \cdot \sum t_i \\ \sum t_i \cdot y_i = b_0 \cdot \sum t_i + b_1 \Sigma t_i^2 \end{cases}$$

Hàm xu thế Parabol: Được sử dụng khi các sai phân bậc hai của của dãy số xấp xỉ bằng nhau. Hàm có dạng: $\widehat{y_i} = b_0 + b_1 t_i + b_2 {t_i}^2$

Các tham số b_0 , b_1, b_2 được xác định bằng phương pháp bình phương nhỏ nhất và phải thoả mãn hệ phương trình:

$$\begin{cases} \sum y_i = n_0 b_0 + b_i \Sigma t_j + b_\alpha \Sigma t_j^2 \\ \sum t_i \cdot y_i = b_0 \Sigma t_i + b_1 \Sigma t_j^2 + b_2 t_i^3 \\ \sum t_i^2 y_i = b_0 \Sigma t_i^2 + b_1 \Sigma t_i^3 + b_2 t_i^4 \end{cases}$$

Hàm xu thế Hypebol: Được sử dụng khi các mức độ của doanh thu giảm dần theo thời gian.

Hàm có dạng: $\widehat{y_i} = b_0 + \frac{b_1}{t_i}$

Các tham số b_0, b_1 được xác định bằng phương pháp bình phương nhỏ nhất và phải thoả mãn hệ phương trình:

$$\sum y_i = n \cdot b_0 + b_1 \Sigma \frac{1}{t_i}$$
$$\sum \frac{1}{t_i} y_i = b_0 \Sigma \frac{1}{t_i} + b_1 \Sigma \frac{1}{t_i^2}$$

Hàm xu thế mũ: Được sử dụng khi các tốc độ phát triển liên hoàn của doanh thu xấp xỉ nhau.

Hàm có dạng: $\widehat{y_i} = b_0 . b_1^{t_i}$

Hay $ln\, ln\, y_i = ln\, ln\, b_0 + t\, ln\, ln\, b_1$

Các tham số b_0, b_1 được xác định bằng phương pháp bình phương nhỏ nhất và lna, lnb phải thoả mãn hệ phương trình:

$$\begin{cases} \sum \ln y_i = n \ln b_0 + \ln b_1 \Sigma t_i \\ \sum t_i \ln y_i = \ln b_0 \Sigma t_i + \ln b_1 \Sigma t_i^2 \end{cases}$$

Trong thực tế khi có thể xác định được nhiều dạng hàm xu thế biểu diễn cho tập hợp dữ liệu, chúng ta có thể xác đinh dạng hàm phù hợp nhất bằng cách dựa vào sai số chuẩn của mô hình, kí hiệu là SE. Sai số chuẩn của mô hình càng nhỏ, sự phù hợp của dạng hàm càng cao. Sai số chuẩn của mô hình được tính theo công thức sau đây:

$$SE = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n - p}}$$

Trong đó: y_i : là mức độ thực tế của doanh thu ở thời gian t

$\hat{y}_i$: là mức độ ước lượng của doanh thu ở thời gian t

n : là số lượng mức độ của dãy số

p : là số lượng các tham số của hàm xu thế

1.4. Phương pháp biểu hiện biến động thời vụ

Biến động thời vụ của doanh thu là sự biến động của hiện tượng có tính chất lặp đi lặp lại trong từng thời gian nhất định. Nguyên nhân gây ra biến động thời vụ là do ảnh hưởng của khí hậu, thời tiết, các điều kiện về văn hoá xã hội, địa lý – hệ sinh thái, phong tục, tập quán sinh hoạt…

Chỉ số thời vụ (khi dãy số không có xu thế)

$$I_i = \frac{\overline{y_i}}{\overline{y}_o} \; x \; 100$$

Trong đó:

I_i: Chỉ số thời vụ của thời gian thứ i (có thể là tháng, quý,...).

$\overline{y_i}$: Doanh thu bình quân ở thời gian i qua các năm.

$\overline{y}_o$: Doanh thu bình quân chung của cả dãy số.

Nếu I_i> 1 (hoặc 100%) thì sự biến động của doanh thu ở thời gian i tăng, tức đây là thời kỳ bận rộn. Nếu I_i < 1 (hoặc 100%) thì sự biến động của doanh thu ở thời gian i giảm.

Chỉ số thời vụ (khi dãy số có xu thế):

Công thức:

$$I_i = \frac{\sum_{j=1}^{n} \frac{y_{ij}}{\hat{y}_{ij}}}{m} \cdot 100(\%)$$

Trong đó:

y_{ij}: Doanh thu thực tế của thời kỳ thứ i (i = 1,n) thuộc năm j (j = 1,m).

$\hat{y}_{ij}$: Doanh thu lý thuyết của thời kỳ thứ i (i = 1,n) thuộc năm j (j = 1,m) được tính từ hàm xu thế.

m: Số năm nghiên cứu.

Tính chỉ số thời vụ lần lượt theo từng bước sau:

Bước 1: Xác định xu thế (bình quân trượt, hàm xu thế, …). Tính bình quân trượt (với k = m):

- Nếu số liệu theo quý, tính bình quân trượt 4 mức độ (đặt mức độ đầu tại y_3).

- Nếu số liệu theo tháng, tính bình quân trượt 12 mức độ (đặt mức độ đầu tại y_7).

Bước 2: Khử yếu tố xu thế:

- Đối với mô hình cộng: $Y - T = S + I$

- Đối với mô hình nhân: $\frac{y}{T} = S \times I$

Bước 3: Khử yếu tố ngẫu nhiên (tính bình quân):

- Đối với mô hình cộng: Tính bình quân cộng giản đơn.

- Đối với mô hình nhân: Tính bình quân cộng trung tâm (Medial average - Trung bình cộng của các lượng biến loại trừ lượng biến nhỏ nhất và lớn nhất).

Bước 4: Điều chỉnh chỉ số thời vụ:

- Đối với mô hình cộng: Tổng chỉ số thời vụ bằng không (0). Mức độ điều chỉnh bằng tổng chỉ số thời vụ chia cho m.

- Đối với mô hình nhân: Tổng chỉ số thời vụ bằng m. Hệ số điều chỉnh bằng. Tổng chỉ số thời vụ chia cho m.

1.5. Các phương pháp dự đoán dựa vào dãy số thời gian

1.5.1. Một số phương pháp dự đoán thống kê đơn giản

1.5.1.1. Dự đoán dựa vào lượng tăng (giảm) tuyệt đối bình quân

Mô hình dự đoán:

$$\hat{y}_{n+L} = y_n + \bar{\delta} \cdot L$$

Trong đó: $\hat{y}_{n+L}$ là mức độ dự đoán ở thời gian n + L

y_n là mức độ cuối cùng của dãy số thời gian

L là tầm xa dự đoán (L = 1,2,...)

$$\bar{\delta} = \frac{y_n - y_1}{n - 1}$$

Mô hình dự đoán trên sẽ cho kết quả dự đoán tốt khi các lượng tăng (giảm) tuyệt đối liên hoàn của doanh thu xấp xỉ bằng nhau.

1.5.1.2. Dự đoán dựa vào tốc độ phát triển bình quân

Mô hình dự đoán:

$$\hat{y}_{n+L} = y_n \cdot (\bar{t})^L$$

Trong đó:

$\hat{y}_{n+L}$ là mức độ dự đoán ở thời gian n+L

y_n là mức độ cuối cùng của dãy số thời gian

L là tầm xa dự đoán (L = 1,2,..)

$$\bar{t} = \sqrt[n\cdot 1]{\frac{y_n}{y_1}}$$

Mô hình dự đoán này sẽ cho kết quả dự đoán tốt khi các tốc độ phát triển liên hoàn của doanh thu xấp xỉ bằng nhau

1.5.1.3. Dự đoán dựa vào hàm xu thế

Sau khi đã lựa chọn được dạng hàm xu thế phù hợp, chúng ta có thể dự đoán doanh thu tiếp theo của dãy số dựa vào mô hình:

$$\hat{y}_i = f(t_i)$$

1.5.2. Dự đoán dựa vào hàm xu thế và chỉ số thời vụ

Quá trình dự đoán được thực hiện thực hiện theo các bước sau đây:

(1) Xây dựng hàm xu thế phù hợp biểu diễn xu hướng biến động cơ bản của doanh thu.

(2) Tính chỉ số thời vụ.

(3) Tùy vào mô hình kết hợp là mô hình cộng hay nhân để dự đoán doanh thu tiếp theo của dãy số.

Mô hình dự đoán:

$$\hat{y}_t = f(t) \cdot I_i$$

với f(t) là hàm xu thế tìm được, I_i là chỉ số thời vụ đối với doanh thu từng thời kỳ cụ thể, thay giá trị t tương ứng để tìm $\hat{y}_t$.

1.5.3. Dự đoán bằng phương pháp san bằng mũ

❖ **Mô hình san bằng mũ giản đơn**

Mô hình san bằng mũ giản đơn được áp dụng để dự đoán với dãy số thời gian không có xu hướng biến động cơ bản rõ ràng và không có biến động thời vụ.

Theo phương pháp san bằng mũ, giá trị san bằng ở thời gian t là:

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

Giá trị dự đoán ở thời gian t+1 là:

$$\hat{y}_{t+1} = s_t = \alpha \cdot y_z + (1 - \alpha)s_{t-1}$$

Cần lưu ý là khi chưa có giá trị thực tế ở thời gian t+1, nếu chúng ta muốn dự đoán cho các doanh thu ở thời gian t+2 và t+3 thì các doanh thu dự đoán này vẫn là:

$$\hat{y}_{t+2} = s_t \text{ và } \hat{y}_{t+3} = s_t$$

Do đó với mô hình san bằng mũ giản đơn, thông thường chúng ta chỉ dự đoán cho một doanh thu tiếp theo của dãy số.

❖ **Mô hình Holt - Winters**

Mô hình Holt-Winters với dãy số có xu thế và không có biến động thời vụ

Khi dãy số có xu thế và không có biến động thời vụ, ta áp dụng mô hình sau:

$$S_t = \alpha \cdot y_t + (1-\alpha)(S_{t-1} + T_{t-1}) \text{ với } 0 < \alpha < 1$$

$$T_t = \beta(S_t - S_{t-1}) + (1-\beta)T_{1-1} \text{ với } 0 < \beta < 1$$

Trong đó:

$S_t$là giá trị san bằng mũ ở thời gian t.

T_t là xu thế ở thời gian t.

Công thức dự báo:

$$\hat{y}_{t+h} = s_{t+hT_t}$$

Trong đó: h là tầm xa dự đoán.

Mô hình Holt-Winters với dãy số có xu thế và biến động thời vụ

Mô hình cộng: các thành phần được tính theo công thức dưới đây:

$$S_t = \alpha\left(y_t - I_{S_{t-s}}\right) + (1-\alpha)(s_{t-1} + T_{t-1}) \text{ , } 0 < \alpha < 1$$

$$T_t = \beta(S_t - S_{t-1}) + (1-\beta)T_{t-1} \text{ , } 0 < \beta < 1$$

$$I_{S_t} = \gamma(y_t - S_t) + (1-\gamma)I_{S_{t-s}}\text{, } 0 < \gamma < 1$$

Trong đó:

S_t là giá trị san bằng mũ ở thời gian t.

T_t là xu thế ở thời gian t và được tính tương tự trường hợp dãy số chỉ có xu thế.

I_{S_t} là chỉ số thời vụ ở thời gian t và s là độ dài của thời vụ (s=4 nếu thời vụ là quý và s=12 nếu thời vụ là tháng).

$I_{s_{t-s}}$ là chỉ số thời vụ ở thời gian t-s

Công thức dự báo: $\hat{y}_{t+h} = S_t + hT_t + I_{S_{t-s}}$

Trong đó: h là tầm xa dự đoán.

Mô hình nhân: các thành phần được tính theo công thức dưới đây:

$$S_t = \alpha\left(\frac{y_t}{I_{S_{t-s}}}\right) + (1-\alpha)(s_{t-1} + T_{t-1}) \text{ , } 0< \alpha <1$$

$$T_t = \beta(S_t - S_{t-1}) + (1-\beta)T_{t-1} \text{ , } 0< \beta <1$$

$$I_{S_t} = \gamma\left(\frac{y_t}{S_t}\right) + (1-\gamma)I_{S_{t-s}} \text{ , } 0< \gamma <1$$

Công thức dự báo: $\hat{y}_{t+h} = (S_t + hT_t).I_{S_{t-s}}$