Phần 2: Các phương pháp dự báo

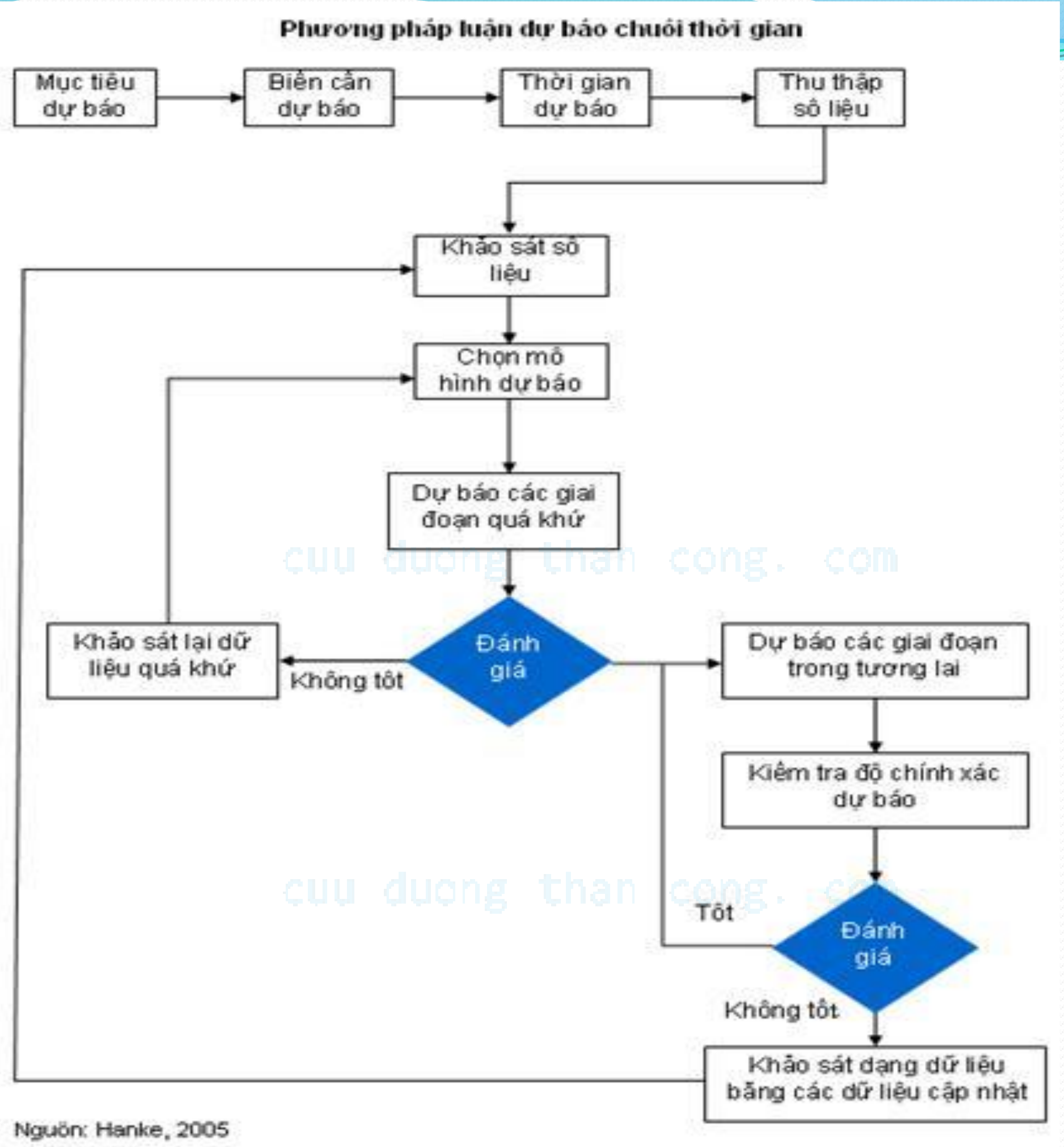
Chương 2. Dự báo bằng phương pháp san mũ

1. Phương pháp ngoại suy xu thế.
2. Phương pháp san mũ :

- *Phương pháp trung bình trượt*

- *San mũ bất biến*

- *San mũ xu thế*



Phương pháp luận dự báo chuỗi thời gian
Mục tiêu dự báo
Biến cần dự báo
Thời gian dự báo
Thu thập số liệu
Khảo sát số liệu
Chọn mô hình dự báo
Dự báo các giai đoạn quá khứ
Đánh giá
Không tốt
Khảo sát lại dữ liệu quá khứ
Dự báo các giai đoạn trong tương lai
Kiểm tra độ chính xác dự báo
Tốt
Đánh giá
Không tốt
Khảo sát dạng dữ liệu bằng các dữ liệu cập nhật
Nguồn: Hanke, 2005

Dự báo theo phương phápNgoại suy xu thế

Khái niệm về ngoại suy

Dự báo giá trị của đối tượng dự báo (chuỗi thời gian) bằng cách mở rộng xu thế của vận động của chuỗi thời gian sang tương lai

- *Ngoại suy là nghiên cứu tiền sử của đối tượng dự báo và chuyển tính quy luật đã phát hiện trong quá khứ, hiện tại sang tương lai với điều kiện:*
  - Đối tượng dự báo phát triển ổn định
  - Những điều kiện chung cho đối tượng dự báo phát triển được duy trì
  - Không có bước nhảy.

Chuỗi thời gian

- *Là tập hợp các giá trị của một biến ngẫu nhiên hay chỉ tiêu thống kê được sắp xếp theo thứ tự thời gian.*
- Dạng tổng quát: $Y_t = f(X_t, W_t, C_t, U_t)$
- Các thành phần cấu thành:
  - Biến động xu thế (X)
  - Biến động thời vụ (W)
  - Biến động theo chu kỳ (C)
  - Biến động ngẫu nhiên (U)

Chuỗi thời gian *(tiếp theo)*

Bước xử lý này là một phần việc của đánh giá trước dự báo.

- Phương pháp xây dựng chuỗi thời gian
  - Đồng nhất về nội dung kinh tế, có thể so sánh,
  - Tuy nhiên, các yêu cầu trên dễ bị vi phạm (do địa giới thay đổi, đối tượng dự báo thay đổi, khoảng thời gian thu thập số liệu khác nhau, khái niệm không thông nhất,…)
  - Do vậy, phải xử lý sơ bộ chuỗi thời gian (loại bỏ sai số): sai số thô, sai số hệ thống và sai số ngẫu nhiên:
  - Các phương pháp xử lý sơ bộ chuỗi thời gian:
    - Phân tích đối chứng kỹ thuật
    - Sử dụng kiểm định thống kê
    - Loại trừ yếu tố ngoài giả thiết

Thành phần biến động xu thế

Xu thế của chuỗi thời gian có thể là tuyến tính hoặc phi tuyến. Trong dự báo bằng phương pháp xu thế, biến độc lập sẽ phải có biến thời gian

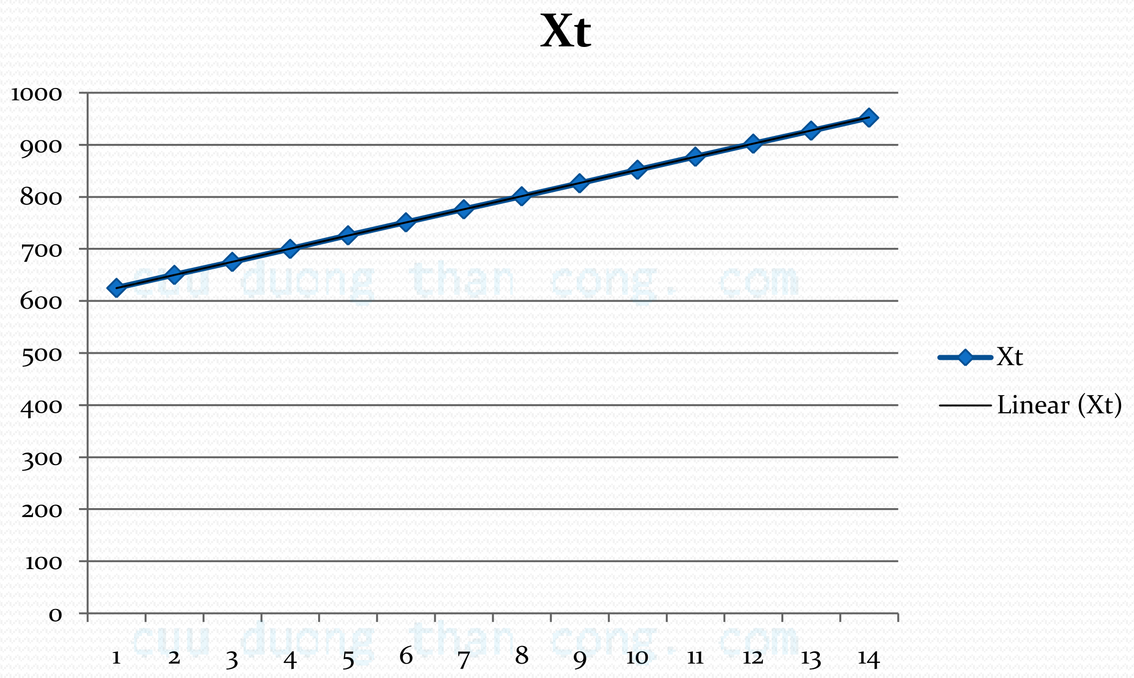
- *Xu thế là một bộ phận của chuỗi thời gian thể hiện khuynh hướng phát triển dài hạn của chuỗi thời gian đó.*
- Cách xác định hàm xu thế
  - Dùng đồ thị
  - Phân tích thống kê
  - Cực tiểu sai số
- Ước lượng hàm xu thế
  - Phương pháp điểm chọn
  - Phương pháp nội suy Newton
  - Phương pháp bình phương nhỏ nhất thông thường

Phân tích số liệu thống kê

- Nếu t và X tăng theo cấp số cộng, xu thế có dạng tuyến tính:
- Nếu t tăng theo cấp số cộng, Xt tăng theo cấp số nhân, xu thế có dạng hàm mũ: $X_t = \alpha + \beta t$
- Nếu logt và logX có quan hệ tuyến tính, xu thế : $X_t = \alpha\beta^t$
- Nếu t tăng đều, sai phân bậc p của Xt là hằng số, xu thế có dạng đa thức bậc p: $X_t = \alpha + \beta_1 t + \beta_2 t^2 + \ldots + \beta_p t^p$
- Nếu t tăng theo cấp số cộng, sai phân bậc nhất của Xt giảm đều, xu thế có dạng Hypebole: $X_t = \alpha + \frac{\beta}{t}$
- Nếu t tăng đều, sai phân bậc nhất của Xt thay đổi dần tới điểm bão hòa, xu thế có dạng Logistic: $X_t = \frac{S}{(1 + e^{-ast-c})}$

Xu thế của chuỗi dân số

Xu thế tuyến tính:



số của hàm xu thế sao cho các tham số ước lượng cho tổng bình phương của các sai số là nhỏ nhất.

Phương pháp OLS

- Với dạng hàm: $X_t = \alpha + \beta_1 t + \beta_2 t^2 + ... + \beta_p t^p$
- Ước lượng tham số sao cho: $Z = \sum_{t=1}^{n}\left(X_t - (\alpha + \sum_{i=1}^{p}\beta_i t^i)\right)^2 --> Min$
- Lấy đạo hàm riêng phần bậc nhất theo các tham số và giải hệ điều kiện cần này:

$$\begin{cases} n\alpha + \beta_1\sum_{t=1}^{n} t + \beta_2\sum_{t=1}^{n} t^2 + ... + \beta_p\sum_{t=1}^{n} t^p = \sum_{t=1}^{n} X_t \\ \alpha\sum_{t=1}^{n} t + \beta_1\sum_{t=1}^{n} t^2 + \beta_2\sum_{t=1}^{n} t^3 ... + \beta_p\sum_{t=1}^{n} t^{p+1} = \sum_{t=1}^{n} X_t t \\ .... \quad .... \quad .... \quad .... \quad .... \quad .... \quad .... \quad ... \\ \alpha\sum_{t=1}^{n} t^p + \beta_1\sum_{t=1}^{n} t^{p+1} + \beta_2\sum_{t=1}^{n} t^{p+2} ... + \beta_p\sum_{t=1}^{n} t^{2p} = \sum_{t=1}^{n} X_t t^p \end{cases}$$

Sai số dự báo và khoảng dự báo

- Khi thành phần ngẫu nhiên tuân theo quy luật phân phối chuẩn

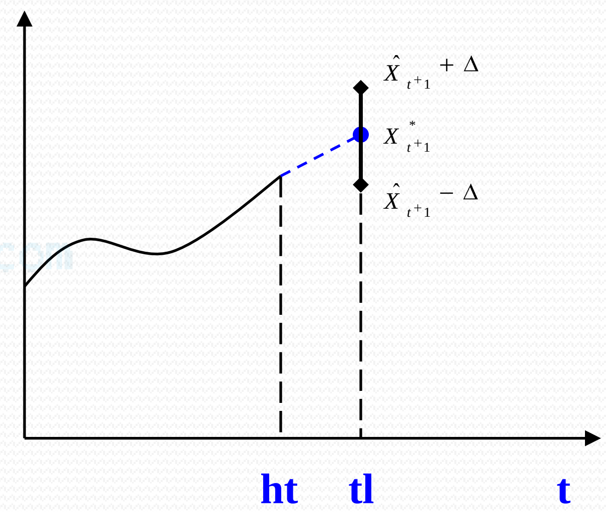
$$\sigma_u = \sqrt{\frac{\sum_{t=1}^{n}(X_t - \hat{X}_t)^2}{n-p-1}}$$

Trong đó: p là bậc của đa thức mô tả xu thế

Sai số cực đại sẽ là: $\Delta = k\sigma_u$ *(k=1, 2, 3)*

Khoảng dự báo sẽ là:

$$\hat{X}_{t+1} - \Delta \le X^*_{t+1} \le \hat{X}_{t+1} + \Delta$$



Sai số dự báo và khoảng dự báo *(tiếp theo)*

- Khi thành phần ngẫu nhiên không tuân theo quy luật phân phối chuẩn, sai số dự báo sẽ được tính

$$\Delta = t_{\alpha}(n).\hat{S}_p$$

trong đó $t_{\alpha}(n)$ là tham số T-student với n bậc tự do và mức ý nghĩa α

$$\hat{S}_p = \sqrt{\frac{S_u^2}{n-p-1}\left(1+\frac{(t_p-\bar{t})^2}{\sum(t_i-\bar{t})^2}\right)}$$

t_p là giá trị ở thời điểm dự báo

$$S_u^2 = \sum_{t=1}^{n} u_t^2 = \sum_{t=1}^{n}(X_t - \hat{X}_t)^2$$

Khoảng dự báo sẽ là: $\hat{X}_{t+1} - \Delta \le X^*_{t+1} \le \hat{X}_{t+1} + \Delta$

Các nhân tố ảnh hưởng tới sai số dự báo

- Tầm xa dự báo
- Độ tin cậy alpha
- Độ dài chuỗi thời gian
- Phương pháp ước lượng tham số
- Chú ý: Khi xu thế là một hàm tuyến tính bậc nhất, ta có:

$$\hat{S}_p = \sigma_u \sqrt{1 + \frac{1}{n} + \frac{3(n + 2l - 1)^2}{n(n^2 - 1)}}$$

Phương pháp san mũ

<u>Đặc điểm:</u>

- *Áp dụng với các chuỗi thời gian đơn biến.*
- *Là phương pháp sử dụng cơ chế thích nghi, được điều chỉnh trên cơ sở các thông tin mới.*
- *Là phương pháp dự báo ngắn hạn.*
- *Áp dụng cho các chuỗi thời gian không thể hiện xu thế hoặc có xu thế dạng đa thức.*

Phương pháp trung bình trượt

Trung bình trượt

- Đây là phương pháp đơn giản nhất trong nhóm các pp san mũ
- Loại bỏ các biến động ngẫu nhiên
- Công thức: $\overline{X}_t = \overline{X}_{t-1} + \dfrac{X_t - X_{t-m}}{m}$ hoặc

$$\overline{X}_t = \frac{X_t + X_{t-1} + X_{t-2} + \ldots + X_{t-(m-1)}}{m}$$

- Nếu chuỗi thời gian không thể hiện xu hướng, và phát triển ổn định thì có thể dùng giá trị trung bình trượt $\overline{X}_t$ cho thời kỳ (t+1).
- Mỗi quan sát trong khoảng trượt m nhận trọng số bằng 1/m vào giá trị dự báo, không phục thuộc vào thời điểm tính.

Cơ sở và công thức của phương pháp san mũ

PP san mũ bất biến là một cách điều chỉnh liên tục giá trị dự báo theo giá trị hiện tại của chuỗi thời gian.

- Dựa trên hai nguyên tắc:
  - Càng xa trong quá khứ, trọng số càng giảm
  - Sai số hiện tại phải được tính tới trong dự báo kế tiếp
- Công thức: Theo nguyên tắc thứ hai

$$\hat{X}_{t+1} = \hat{X}_t + \alpha\varepsilon_t = \hat{X}_t + \alpha(X_t - \hat{X}_t)$$ hoặc

$$\hat{X}_{t+1} = \alpha X_t + (1-\alpha)\hat{X}_t$$

- Bằng phương pháp thế, ta có dạng tổng quát:

$$\hat{X}_{t+1} = \alpha \sum_{i=0}^{t-1} (1-\alpha)^i X_{t-i} + (1+\alpha)^t \hat{X}_t$$

Hệ số san α

Mô hình bất biến

$$0 \leq \alpha \leq 1$$

- Tham số san là trung tâm của phương pháp san mũ
  - Khi $\alpha = 1$ thì giá trị dự báo cho thời kỳ kế tiếp bằng chính giá trị hiện tại
  - Khi $\alpha = 0$ thì giá trị dự báo cho thời kỳ kế tiếp bằng chính giá trị dự báo ở thời kỳ trước đó
  - Với các giá trị tham số san lớn (gần 1) thì trọng số của các quan sát quá khứ càng nhỏ. Tham số san này rất ý nghĩa khi có sự thay đổi lớn cơ bản từ chuỗi thời gian.
  - Với các tham số san nhỏ (gần 0), trọng số của các quan sát quá khứ sẽ lớn hơn các quan sát gần hiện tại. Hệ số san này thích hợp hơn với các chuỗi thời gian có tính ổn định cao.

Đặc điểm và ước lượng tham số

Xu thế san mũ

Phương pháp xu thế san mũ là một trong các phương pháp dự báo cho chuỗi thời gian có thành phần xu thế tuyến tính.

- Dùng dự báo cho các CTG có thành phần biến động xu thế
- Dùng để dự báo trong ngắn hạn
- Các tham số được ước lượng bằng phương pháp cực tiểu hóa tổng bình phương các sai số theo quy luật mũ

$$Z = \alpha \sum_{i=0}^{t-1} (1-\alpha)^i u_{t-i}^2 = \alpha \sum_{i=0}^{t-1} (1-\alpha)^i (X_{t-i} - a + bi)^2$$

- Giải điều kiện cần

$$\frac{\partial Z}{\partial a} = -2\alpha \sum_{i=0}^{t-1} (1-\alpha)^i (X_{t-i} - a + bi) = 0$$

$$\frac{\partial Z}{\partial b} = 2\alpha i \sum_{i=0}^{t-1} (1-\alpha)^i (X_{t-i} - a + bi) = 0$$

Đặc điểm và ước lượng tham số *(tiếp theo)*

Xu thế san mũ

- Cụ thể: $$\begin{cases} a\sum_{i=0}^{t-1}\alpha(1-\alpha)^i - b\sum_{i=0}^{t-1}\alpha(1-\alpha)^i i = \sum_{i=0}^{t-1}\alpha(1-\alpha)^i X_{t-i} \\ a\sum_{i=0}^{t-1}\alpha(1-\alpha)^i i - b\sum_{i=0}^{t-1}\alpha(1-\alpha)^i i^2 = \sum_{i=0}^{t-1}\alpha(1-\alpha)^i X_{t-i}.i \end{cases}$$

- Lưu ý rằng, khi $t \to \infty$

$$\sum_{i=0}^{t-1}\alpha(1-\alpha)^i \to 1$$

$$\sum_{i=0}^{t-1}\alpha(1-\alpha)^i i \to \frac{1-\alpha}{\alpha}$$

$$\sum_{i=0}^{t-1}\alpha(1-\alpha)^i i^2 \to \frac{(1-\alpha)(2-\alpha)}{\alpha^2}$$

Đặc điểm và ước lượng tham số *(tiếp theo)*

- Nếu đặt toán tử cấp 1 và cấp 2 lần lượt là:

$$S_t^1 = \alpha \sum_{i=0}^{t-1} (1-\alpha)^i X_{t-i} \quad \text{và} \quad S_t^2 = \alpha \sum_{i=0}^{t-1} (1-\alpha)^i S_{t-i}^1 \qquad (*)$$

- Khi đó, ta có

$$\begin{cases} a - \dfrac{1-\alpha}{\alpha} b = S_t^1 \\ \dfrac{1-\alpha}{\alpha} a - \dfrac{(1-\alpha)(2-\alpha)}{\alpha^2} b = \dfrac{S_t^2}{\alpha} - S_t^1 \end{cases}$$

- Giải hệ phương trình này, ta được

$$\hat{a} = 2S_t^1 - S_t^2$$

$$\hat{b} = \frac{\alpha}{(1-\alpha)}(S_t^1 - S_t^2)$$

Đặc điểm và ước lượng tham số *(tiếp theo)*

- Để tính được các giá trị tham số, ta phai biết giá trị của tham số san, các toán tử cấp 1 và toán tử cấp 2, và các toán từ này được tính toán theo công thức (*), nếu biết trước toán tử cấp 1 và 2 ở thời điểm t = 0.

$$\begin{cases} S_0^1 = a_0 - \dfrac{1-\alpha}{\alpha} b_0 \\ S_0^2 = a_0 - \dfrac{2(1-\alpha)}{\alpha} b_0 \end{cases}$$

Với a_0 và b_0 là các tham số ước lượng ban đầu của chuỗi thời gian

- Phương trình dự báo có dạng:

$$\hat{X}_{t+l} = \hat{a}_t + \hat{b}_t . l \qquad (l = 1, 2, 3, ...)$$

Mở rộng của phương pháp san mũ

Mở rộng

- **Tham số san động**
  - Chow and Smith sử dụng MFE và MAE cho mục đích thích nghi động của tham số san

$$MFE_t = \beta\varepsilon_t + (1-\beta)MFE_{t-1}$$

$$MAE_t = \beta|\varepsilon_t| + (1-\beta)MAE_{t-1} \quad và \quad \alpha_t = \frac{|MFE_t|}{MAE_t}$$

  - Giá trị tham số san alpha được xác định mới trong mỗi thời kỳ
  - Sự biến thiên của alpha luôn thỏa mãn: $0 \le \alpha \le 1$
  - Để cho tham số san động có sự thay đổi và đảm bảo sự ổn định, Smith tiến hành san mũ với hệ số san động vừa tìm được.

$$\alpha_t = \gamma\alpha_t + (1-\gamma)\alpha_{t-1}$$

Mở rộng của phương pháp san mũ

- **San mũ bậc cao**
  - Brown và Meyer đã xây dựng trong lý thuyết san mũ cho đa thức bậc p. $X_t = \alpha + \beta_1 t + \beta_2 t^2 + ... + \beta_p t^p$
  - Tuy nhiên, do khối lượng tính toán phức tạp, kết quả dự báo không chứng tỏ được sự vượt trội, và khó khăn về giải thích ý nghĩa kinh tế,… nên trong thực tế san mũ bậc cao không được sử dụng nhiều.

Mở rộng của phương pháp san mũ

- **San mũ cho mô hình đa tham số**
  - Nhiều ý kiến cho rằng dự báo san mũ điều chỉnh nhờ vào một tham số san duy nhất là chưa đủ, cần có tham các tham số.
  - Holt (1957) đưa ra mô hình 2 tham số san và dễ dàng chuyển thành 3 tham số, Mô hình có dạng:

$$\hat{X}_{t+1} = \hat{X}_t + \alpha\varepsilon_t + \beta(\varepsilon_t - \varepsilon_{t-1}) \quad \textit{và}$$

$$\hat{X}_{t+1} = \hat{X}_t + \alpha\varepsilon_t + \beta(\varepsilon_t - \varepsilon_{t-1}) + \gamma\sum_{i=0}^{t-1}\varepsilon_{t-i}$$

  - Trong thực tế, chưa chứng minh được tính ưu việt của hàm đa thức với nhiều tham số san, do đó mô hình san mũ bất biến vẫn là mô được sử dụng rộng rãi hơn nhơ tính đơn giản và hiệu quả.

Ưu nhược điểm của phương pháp

- **Ưu điểm**
  - Đơn giản, kết quả tương đối chính xác đáp ứng tốt cho kinh doanh, công tác lập kế hoạch ở cấp vi mô.
  - Dễ chương trình hóa
  - Nhu cầu lưu trữ thấp (Không còn là lợi thế khi công nghệ lưu trữ phát triển)
  - Kết quả dự báo có thể được điều chỉnh cho thích hợp thông qua một hệ số san
  - Các bước tiến hành dự báo khá rõ ràng, dễ áp dụng

  → Ứng dụng nhiều trong dự báo kinh doanh, khối lượng bán hàng,…

Ưu nhược điểm của phương pháp

- **Nhược điểm**
  - Không quan tâm tới ảnh hưởng nhân quả
  - Tham số san alpha không được xác định một cách khoa học khách quan.
  - Hàm mục tiêu được cực tiểu hóa theo quy luật số mũ, do đó nếu những sai số ước lượng là ngẫu nhiên thì việc đánh giá theo trọng số là dư thừa vì mỗi thời kỳ đều nhận một trong số tương tự nhau. Còn nếu các sai số bị ảnh hưởng bởi các nhân tố một cách có hệ thống thì mô hình không phản ánh được, mà cần sử dụng phương pháp hồi quy.