



BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC QUY NHƠN

CHƯƠNG 1

TỔNG QUAN

VỀ PHÂN TÍCH DỮ LIỆU VÀ DỰ BÁO



- Hiểu được dự báo là gì và vai trò quan trọng của dự báo
- Hiểu được vì sao dự báo định lượng đang trở nên phổ biến
- Hiểu được các cách phân loại dự báo
- Hiểu được phương pháp luận của dự báo
- Nắm vững quy trình thực hiện và cách thức đánh giá độ chính xác dự báo trên thực tế
- Nắm rõ được các loại dữ liệu sử dụng cho dự báo

TỔNG QUAN VỀ PHÂN TÍCH VÀ DỰ BÁO

1. Dự báo và vai trò của dự báo
2. Phân loại dự báo
3. Các phương pháp dự báo
4. Quy trình thực hiện dự báo
5. Dữ liệu cho dự báo

DỰ BÁO VÀ VAI TRÒ CỦA DỰ BÁO

- Dự báo đã hình thành từ đầu những năm 60 của thế kỉ 20
- Dự báo là một khoa học và nghệ thuật tiên đoán những sự việc sẽ xảy ra trong tương lai, trên cơ sở phân tích khoa học về các dữ liệu đã thu thập được
- Dự báo ngày càng trở nên quan trọng và trở thành bộ phận không thể thiếu trong hầu hết các quyết định của mọi tổ chức
- Dự báo chính xác là căn cứ để các nhà hoạch định các chính sách phát triển kinh tế văn hoá xã hội trong toàn bộ nền kinh tế quốc dân

DỰ BÁO VÀ VAI TRÒ CỦA DỰ BÁO (tt)

- Dự báo chính xác sẽ giảm bớt mức độ rủi ro cho doanh nghiệp nói riêng và toàn bộ nền kinh tế nói chung
- Dự báo giúp các nhà quản trị doanh nghiệp chủ động trong việc đề ra các kế hoạch và các quyết định cần thiết phục vụ cho quá trình sản xuất kinh doanh, đầu tư, quảng bá, quy mô sản xuất, kênh phân phối sản phẩm, nguồn cung cấp tài chính... và chuẩn bị đầy đủ điều kiện cơ sở vật chất, kỹ thuật cho sự phát triển trong tương lai
- Nhờ có dự báo các chính sách kinh tế, các kế hoạch và chương trình phát triển kinh tế được xây dựng có cơ sở khoa học và mang lại hiệu quả kinh tế cao

PHÂN LOẠI DỰ BÁO

- Dựa trên kết quả dự báo
 - + Dự báo điểm và dự báo khoảng
 - + Dự báo ngắn hạn, trung hạn hay dài hạn
- Dựa trên phạm vi dự báo: dự báo kinh tế vi mô và dự báo kinh tế vĩ mô
- Dựa trên phương pháp dự báo
 - + Nhóm phương pháp không chính thức: trực giác cảm tính, kinh nghiệm và khả năng phán đoán
 - + Nhóm phương pháp chính thức: dự báo định tính và dự báo định lượng

CÁC PHƯƠNG PHÁP DỰ BÁO

Định tính

- **Ý kiến chuyên gia:** Tổng hợp ý kiến chuyên gia, khi có sự hỗ trợ của công cụ thống kê
- **Ý kiến bán hàng:** Các đại lý, bán hàng tiến hành dự báo, công ty tổng hợp kết quả
- **Ý kiến khách hàng:** tổng hợp ý kiến khách hàng
- **Phương pháp Delphi:** sự đồng thuận của các bên

Định lượng

Chuỗi thời gian

- Dự báo thô
- Trung bình
- Hàm mũ
- Hàm xu thế
- ARIMA(p,d,q)
- ARCH/GARCH

Hồi quy

- Hồi quy đơn
- Hồi quy bội

CÁC PHƯƠNG PHÁP DỰ BÁO (tt)

Phương pháp Delphi gồm các bước:

1. Những chuyên gia tham gia dự báo được chọn
2. Bảng câu hỏi về các biến dự báo được đưa cho từng thành viên
3. Kết quả được thu thập và lập bảng và tóm tắt
4. Bảng tóm tắt kết quả sẽ gửi lại cho từng chuyên gia xem xét lại
5. Từng chuyên gia sẽ xem xét lại các dự báo của mình trên cơ sở tham khảo ý kiến chung của nhiều chuyên gia khác
6. Lập đi lập lại cho đến khi kết quả không khác biệt nhiều giữa các chuyên gia

CÁC PHƯƠNG PHÁP DỰ BÁO (tt)

- **Ưu và nhược điểm của phương pháp dự báo định tính**

- **Ưu điểm:**

Dễ dàng thực hiện, không cần phải có nền tảng về dự báo, thường được chấp nhận.

- **Nhược điểm:**

Mang tính chủ quan rất cao, không chuẩn, mất nhiều năm để trở thành người có khả năng phán đoán đúng (kinh nghiệm là ông thầy giỏi, tuy nhiên là con dao hai lưỡi)

QUY TRÌNH DỰ BÁO

Theo Wilson và Keating, quy trình dự báo gồm 9 bước:

1. Xác định mục tiêu

- Thống nhất giữa người ra tiến hành dự báo và sử dụng kết quả dự báo (Dùng để làm gì? Tầm quan trọng như thế nào? Giúp ích gì cho quá trình ra quyết định?)
- **Ví dụ:** trên cơ sở dự báo về doanh thu, bộ phận Logistics của dự báo nhu cầu vật tư (người thực hiện), Bộ phận Kế toán sử dụng dự báo này để tính hiệu quả cụ thể của kế hoạch

2. Quyết định đối tượng dự báo

Xác định được đối tượng hay biến dự báo cụ thể (đo bằng gì?). Phạm vi (một sản phẩm hay nhóm sản phẩm hay tổng hợp nhiều ngành sản phẩm, trong khu vực cụ thể hay là toàn vùng, trong nước hay là cả xuất khẩu).

3. Xác định loại dự báo

- Dài hạn, trung hạn hay ngắn hạn và thời đoạn dự báo
- Ví dụ: GDP: năm; Kim ngạch XNK: năm hoặc quý; Doanh số: Quý

4. Nghiên cứu, khảo sát dữ liệu

- Chất lượng dự báo phụ thuộc vào chất lượng của số liệu
- Dữ liệu có thể thu thập từ các nguồn: Nội bộ hoặc bên ngoài; ở dạng sơ cấp hoặc thứ cấp
- Chuyển đổi dữ liệu tương thích với yêu cầu của mục đích nếu có đầy đủ cơ sở

QUY TRÌNH DỰ BÁO (tt)

5. Chọn mô hình: chọn mô hình tùy thuộc vào các tiêu chí sau:

- ✓ Dạng phân bố của dữ liệu
 - ✓ Số lượng quan sát sẵn có
 - ✓ Độ dài của tầng dự báo (ngắn hạn, trung hạn hay dài hạn)
- Nếu dữ liệu phân bố thể hiện tính xu hướng thì có thể áp dụng các phương pháp: trung bình, đường số mũ.
- Nếu số lượng quan sát quá ít: không thể sử dụng mô hình hồi qui (PP này càng nhiều càng tốt). Ngược lại có thể chọn phương pháp đơn giản.

6. Đánh giá mô hình

- Thông qua các thông số như: Sai Số trung bình (Mean Error), Trung bình tuyệt đối sai số (Mean Absolute Error)...
- Sử dụng holdout period: Ví dụ với bộ dữ liệu của 10 năm tính theo quý, ta có thể dùng 8 năm để dự báo cho năm thứ 9 và thứ 10. MH nào dự báo tốt nhất sẽ được chọn để dự báo cho toàn bộ 10 năm. Nếu không MH nào cho kết quả tới một độ chính xác chấp nhận được, quay về bước 5 để chọn mô hình thay thế.
- Đối với các mô hình hồi quy còn có các phương pháp kiểm định và đo lường mức độ giải thích của các biến.

7. Chuẩn bị dự báo

Sau khi qua bước 06 chúng ta tin tưởng rằng mô hình chúng ta chọn sẽ cho ra kết quả chuẩn xác. Có trường hợp chúng ta phải chọn phương pháp kết hợp giữa định tính và định lượng

8. Trình bày dự báo

- Trình bày rõ ràng
- Không cần phải phức tạp hóa kết quả bằng các công thức, hay thuật toán phức tạp
- Rõ ràng dù là bằng văn bản hay truyền đạt.

9. Theo dõi kết quả

- Liên tục theo dõi, so sánh kết quả dự báo với giá trị thực tế
- Thay đổi phương pháp khi cần
- Chấp nhận sai lầm để cải tiến

ĐO LƯỜNG ĐỘ CHÍNH XÁC DỰ BÁO

Đặt: A_t : Giá trị thực ở thời điểm t (quan sát t)

F_t : Giá trị dự báo ở thời điểm t

n : Số lượng quan sát sử dụng trong dự báo

1. Sai số trung bình (mean error)

$$ME = \frac{\sum(A_t - F_t)}{n}$$

2. Sai số tuyệt đối trung bình (Mean absolute error)

$$MAE = \frac{\sum |A_t - F_t|}{n}$$

3. Sai số phần trăm trung bình (mean percentage error)

$$\text{MPE} = \frac{1}{n} \sum \frac{A_t - F_t}{A_t}$$

4. Trung bình sai số phần trăm tuyệt đối (Mean absolute percentage error)

$$\text{MAPE} = \frac{1}{n} \sum \frac{|A_t - F_t|}{A_t}$$

ĐO LƯỜNG ĐỘ CHÍNH XÁC DỰ BÁO (tt)

5. Sai số bình phương trung bình (Mean Squared Error)

$$\text{MSE} = \frac{1}{n} \sum (A_t - F_t)^2$$

6. Căn bậc hai của sai số bình phương trung bình (Root Mean Squared Error)

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum (A_t - F_t)^2}$$

Các tiêu chí 1-6 càng nhỏ thì càng tốt.

7. Hệ số không ngang bằng Theil's U

$$U = \frac{\sqrt{\sum (A_t - F_t)^2}}{\sqrt{\sum (A_t - A_{t-1})^2}}$$

- U càng gần 0 thì mô hình dự báo càng chính xác
- Thực tế giá trị của $U < 0,55$ được đánh giá là rất tốt.

ĐO LƯỜNG ĐỘ CHÍNH XÁC DỰ BÁO (tt)

Một số lưu ý:

1. MAE, MAPE, MSE, RMSE và Theil's U có thể sử dụng để so sánh các mô hình dự báo khác nhau cho cùng một chuỗi dữ liệu
2. Nếu các chuỗi khác nhau về đơn vị đo lường, đơn vị thời gian, dạng dữ liệu (dữ liệu gốc và dữ liệu chuyển hóa logarit) thì chỉ có MAPE và Theil's U có thể dùng để so sánh được
3. Có thể kết hợp đánh giá độ chính xác bằng các đồ thị

ĐO LƯỜNG ĐỘ CHÍNH XÁC DỰ BÁO (tt)

Ví dụ

t	A_t	F_t	e_t	$ e_t $	e_t^2	$ e_t /A_t$	e_t/A_t
1	58	-	-	-	-	-	-
2	54	58	-4	4	16	0.074	-0.074
3	60	54	6	6	36	0.100	0.100
4	55	60	-5	5	25	0.091	-0.091
5	62	55	7	7	49	0.113	0.113
6	62	62	0	0	0	0.000	0.000
7	65	62	3	3	9	0.046	0.046
8	63	65	-2	2	4	0.032	-0.032
9	70	63	7	7	49	0.100	0.100
Tổng cộng			12	34	188	0.556	0.162

A_t : Số khách đến cửa hàng vào thời điểm t

F_t : Số khách dự báo đến cửa hàng vào thời điểm t (theo mô hình dự báo thô giản đơn)

$$e_t = A_t - F_t$$

Ta tính được các chỉ số:

$$MPE = 0.0203; MAE = 4.3$$

$$MAPE = 0.0695; MSE = 23.5$$

$$RMSE = 4.848$$

$$U = 1$$

DỮ LIỆU CHO DỰ BÁO

Theo Hanke & Wichern (2009) dữ liệu cần đạt được 4 tiêu chí sau:

1. Tin cậy và chính xác: nguồn dữ liệu phải tin cậy với độ chính xác đạt yêu cầu.
2. Phù hợp: dữ liệu phải phù hợp với chủ đề.
3. Nhất quán: khi có những thay đổi ở dữ liệu được thu thập thì cần phải điều chỉnh dữ liệu quá khứ để đảm bảo tính nhất quán.
4. Thời gian: không nên quá ít (không đủ thông tin đại diện cho quá khứ) cũng không nên quá nhiều (có thể tích hợp những đặc tính xưa cũ, không đại diện cho hiện trạng)

DỮ LIỆU CHO DỰ BÁO

- **Dữ liệu sơ cấp:** dữ liệu được thu thập lần đầu mà các nguồn khác không sẵn có.

Thu thập: phỏng vấn, gửi bảng câu hỏi, nghiên cứu tài liệu, quan sát trực tiếp, khảo sát qua điện thoại, qua thư điện tử, mạng xã hội hoặc trực tuyến,...

- **Dữ liệu thứ cấp:** dữ liệu được thu thập từ những nguồn sẵn có đã được thu thập cho các mục đích khác nhau.

Thu thập: cơ quan thống kê, cơ quan Chính phủ hay địa phương, báo cáo nội bộ, báo/tạp chí, từ các tổ chức, hiệp hội, trường đại học, viện nghiên cứu,...

DỮ LIỆU CHO DỰ BÁO

1. Dữ liệu chéo: dữ liệu về một hay nhiều biến được thu thập tại cùng một thời điểm ở nhiều đơn vị

Tỉnh	Dân số (nghìn)	Lao động (nghìn)
Quảng Nam	1505.0	891.18
Quảng Ngãi	1233.4	731.69
Bình Định	1487.9	864.50
Phú Yên	874.3	510.68
Khánh Hòa	1240.4	671.70
Kon Tum	555.6	312.80
Gia Lai	1541.8	898.84
Đắk Lắk	1886.9	1101.70

*(Dữ liệu cho năm 2020)
Nguồn: GSO*

DỮ LIỆU CHO DỰ BÁO (tt)

2. Dữ liệu chuỗi thời gian: dữ liệu về một hay nhiều biến được thu thập theo thời gian ở một đơn vị

Năm	Dân số (nghìn)	Lao động (nghìn)
2013	1514.5	914.4
2014	1519.7	889.3
2015	1524.6	897.0
2016	1529.0	893.7
2017	1534.8	907.3
2018	1487.8	891.2
2019	1487.9	864.5
2020	1514.5	914.4

*(Dữ liệu của tỉnh Bình Định)
Nguồn: GSO*

DỮ LIỆU CHO DỰ BÁO (tt)

3. Dữ liệu bảng/mảng: dữ liệu về một hay nhiều biến được thu thập theo thời gian ở nhiều đơn vị

Nguồn: GSO

Năm	Tỉnh	Dân số (nghìn)	Lao động (nghìn)
2017	Bình Định	1534.8	907.3
2018	Bình Định	1487.8	891.2
2019	Bình Định	1487.9	864.5
2020	Bình Định	1514.5	914.4
2017	Quảng Ngãi	1261.6	746.9
2018	Quảng Ngãi	1272.8	746.7
2019	Quảng Ngãi	1231.9	739.2
2020	Quảng Ngãi	1233.4	731.7

Năm	Tỉnh	Dân số (nghìn)	Lao động (nghìn)
2017	Phú Yên	904.4	514.4
2018	Phú Yên	909.5	524.5
2019	Phú Yên	873.2	521.1
2020	Phú Yên	874.3	510.7
2017	Gia Lai	1437.4	879.4
2018	Gia Lai	1458.5	890.0
2019	Gia Lai	1520.2	908.1
2020	Gia Lai	1541.8	898.8



BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC QUY NHƠN

KẾT THÚC CHƯƠNG 1



BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC QUY NHƠN

CHƯƠNG 2

CÁC PHƯƠNG PHÁP PHÂN TÍCH VÀ DỰ BÁO



- Phân biệt được các thành phần của một chuỗi thời gian
- Hiểu được các mô hình và phương pháp dự báo giản đơn
- Xác định được sự phù hợp giữa từng loại dữ liệu với các phương pháp hay mô hình dự báo để áp dụng cho hiệu quả
- Nắm rõ được quy trình thực hiện dự báo các mô hình trên phần mềm Eviews

CÁC PHƯƠNG PHÁP PHÂN TÍCH VÀ DỰ BÁO

1. Các mô hình dự báo thô
2. Các phương pháp dự báo trung bình
3. Phương pháp san mũ giản đơn
4. Phương pháp san mũ Holt
5. Phương pháp san mũ Winters
6. Dự báo với mô hình nhân tính
7. Dự báo với mô hình cộng tính

CÁC MÔ HÌNH DỰ BÁO GIẢN ĐƠN

1. MÔ HÌNH DỰ BÁO THÔ GIẢN ĐƠN

$$\hat{Y}_{t+1} = Y_t$$

- Thường áp dụng khi có quá ít dữ liệu quá khứ (doanh nghiệp mới thành lập)
- 100% trọng số được gán cho giá trị hiện tại của dữ liệu (Y_t) khi dự báo cho giai đoạn $t + 1$

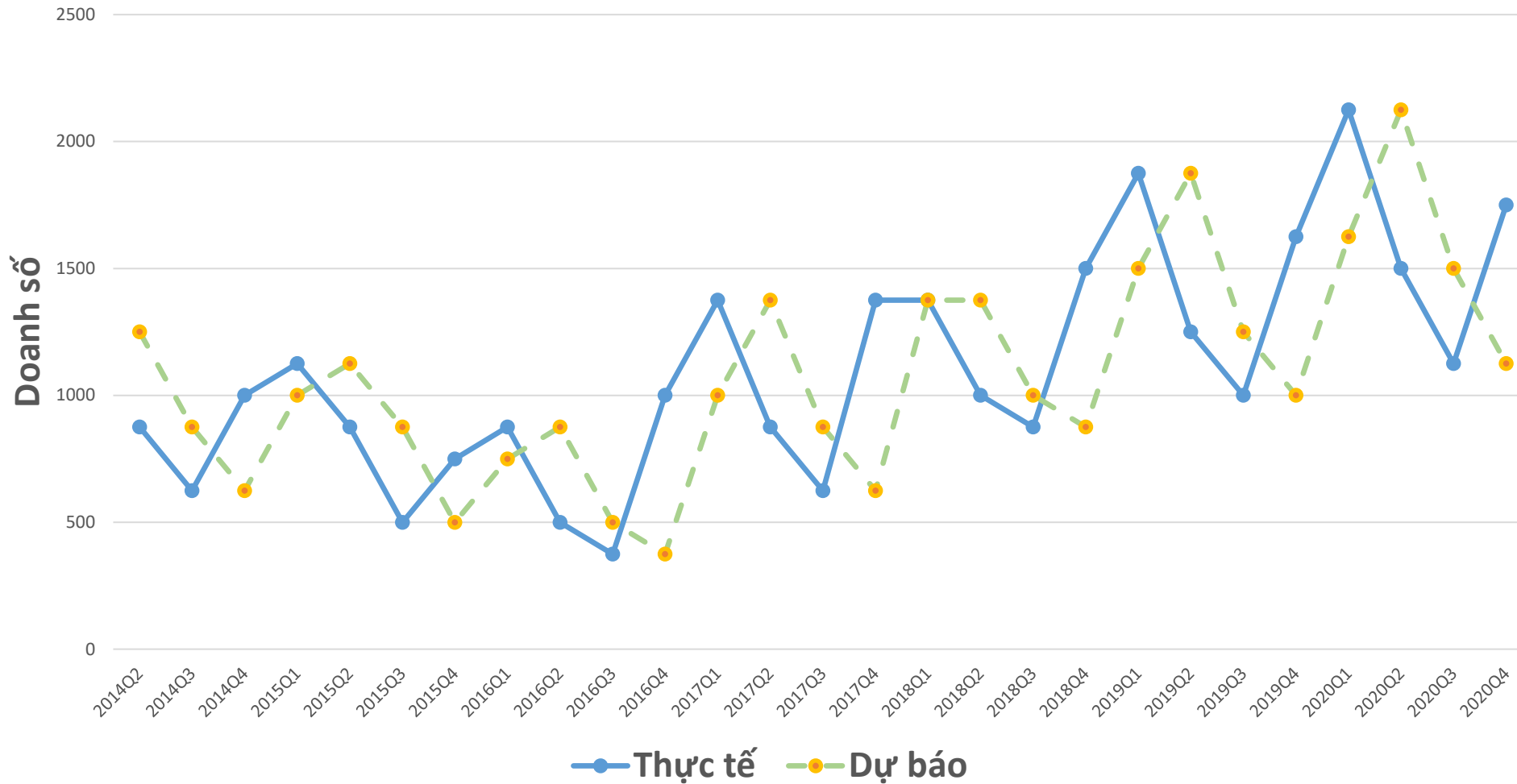
Ví dụ: ta có dữ liệu của công ty ABC giai đoạn 2014 - 2020 như trong bảng dưới đây

CÁC MÔ HÌNH DỰ BÁO GIẢN ĐƠN (tt)

Năm	Quý	t	Sales	Năm	Quý	t	Sales
2014	1	1	1250	2018	3	15	625
	2	2	875		4	16	1375
	3	3	625		1	17	1375
	4	4	1000		2	18	1000
2015	1	5	1125	2019	3	19	875
	2	6	875		4	20	1500
	3	7	500		1	21	1875
	4	8	750		2	22	1250
2016	1	9	875	2020	3	23	1000
	2	10	500		4	24	1625
	3	11	375		1	25	2125
	4	12	1000		2	26	1500
2017	1	13	1375	2020	3	27	1125
	2	14	875		4	28	1750

CÁC MÔ HÌNH DỰ BÁO GIẢN ĐƠN (tt)

Đồ thị giá trị thực tế và dự báo



$MAE = 379.6$

$MAPE = 0.368$

$MPE = 0.072$

$MSE = 181.713$

$RMSE = 426$

$U = 1$

CÁC MÔ HÌNH DỰ BÁO GIẢN ĐƠN (tt)

2. MÔ HÌNH DỰ BÁO THÔ ĐIỀU CHỈNH: điều chỉnh xu thế

$$\hat{Y}_{t+1} = Y_t + (Y_t - Y_{t-1}) \quad \text{HOẶC} \quad \hat{Y}_{t+1} = Y_t \frac{Y_t}{Y_{t-1}}$$

Ví dụ: với dữ liệu công ty ABC ở trên, ta có thể dự báo doanh số của quý 1 năm 2021 như sau

$$\begin{aligned}\hat{Y}_{29} &= \hat{Y}_{28+1} = Y_{28} + (Y_{28} - Y_{27}) \\ &= 1570 + (1570 - 1125) = 2375\end{aligned}$$

MAE	MAPE	MPE	MSE	RMSE	U
509.6	0.484	-0.031	362.981	602	1.41

CÁC MÔ HÌNH DỰ BÁO GIẢN ĐƠN (tt)

2. MÔ HÌNH DỰ BÁO THÔ ĐIỀU CHỈNH: điều chỉnh mùa vụ

Với dữ liệu theo quý, mô hình dự báo thô có thể được điều chỉnh như sau

$$\hat{Y}_{t+1} = Y_{t-3}$$

Ví dụ: với dữ liệu công ty ABC ở trên, ta có thể dự báo doanh số của quý 1 năm 2021 như sau

$$\hat{Y}_{29} = \hat{Y}_{28+1} = Y_{28-3} = Y_{25} = 2125$$

MAE	MAPE	MPE	MSE	RMSE	U
218.8	0.225	-0.053	65.104	255.2	0.599

CÁC MÔ HÌNH DỰ BÁO GIẢN ĐƠN (tt)

2. MÔ HÌNH DỰ BÁO THÔ ĐIỀU CHỈNH: xu thế và mùa vụ

Với dữ liệu vừa có yếu tố xu thế và mùa vụ (theo quý), mô hình dự báo thô có thể được điều chỉnh như sau

$$\hat{Y}_{t+1} = Y_{t-3} + \frac{(Y_t - Y_{t-1}) + \dots + (Y_{t-3} - Y_{t-4})}{4} = \frac{(Y_t - Y_{t-4})}{4}$$

Ví dụ: với dữ liệu công ty ABC ở trên, ta có thể dự báo doanh số của quý 1 năm 2021 như sau

$$\hat{Y}_{29} = \hat{Y}_{28+1} = Y_{25} + \frac{(Y_{28} - Y_{24})}{4} = 2156$$

CÁC MÔ HÌNH DỰ BÁO GIẢN ĐƠN (tt)

SO SÁNH CÁC MÔ HÌNH DỰ BÁO GIẢN ĐƠN

Dự báo	MAE	MAPE	MPE	MSE	RMSE	U
Thô giản đơn	379.6	0.368	0.072	181.713	426	1
Thô điều chỉnh xu thế	509.6	0.484	-0.031	362.981	602	1.41
Thô điều chỉnh mùa vụ	218.8	0.225	-0.053	65.104	255.2	0.599
Thô đ/c xu thế và mùa vụ	180	0.181	-0.047	46.663	216	0.507

CÁC PHƯƠNG PHÁP DỰ BÁO TRUNG BÌNH

1. DỰ BÁO TRUNG BÌNH GIẢN ĐƠN

$$\hat{Y}_{t+1} = \frac{1}{t} \sum_{t=1}^t Y_t$$

- Khi thông tin được cập nhật, công thức trên có thể được điều chỉnh như sau

$$\hat{Y}_{t+1} = \frac{t\hat{Y}_t + Y_{t+1}}{t+1}$$

- Phù hợp với các chuỗi dữ liệu ít biến động

CÁC PHƯƠNG PHÁP DỰ BÁO TRUNG BÌNH (tt)

Tuần	Jean	Tuần	Jean	Tuần	Jean
1	550	11	604	21	620
2	582	12	574	22	598
3	614	13	580	23	570
4	562	14	622	24	500
5	590	15	554	25	520
6	536	16	490	26	490
7	504	17	564	27	542
8	558	18	554	28	564
9	528	19	596	29	604
10	576	20	606	30	570

Bảng bên là dữ liệu số quần Jean tiêu thụ/tuần của một cửa hàng nhãn hiệu thời trang Nino-Max trong 30 tuần.

Ta sẽ dự báo số quần Jean bán theo trung bình giản đơn

CÁC PHƯƠNG PHÁP DỰ BÁO TRUNG BÌNH (tt)

Tuần	TT	DB	Tuần	TT	DB	Tuần	TT	DB
2	582	550.00	12	574	564.00	22	598	569.71
3	614	566.00	13	580	564.83	23	570	571.00
4	562	582.00	14	622	566.00	24	500	570.96
5	590	577.00	15	554	570.00	25	520	568.00
6	536	579.60	16	490	568.93	26	490	566.08
7	504	572.33	17	564	564.00	27	542	563.15
8	558	562.57	18	554	564.00	28	564	562.37
9	528	562.00	19	596	563.44	29	604	562.43
10	576	558.22	20	606	565.16	30	570	563.86
11	604	560.00	21	620	567.20	31		564.07

Bảng bên là kết quả dự báo theo trung bình giản đơn

MAE = 32.15

MAPE = 0.058

MPE = -0.008

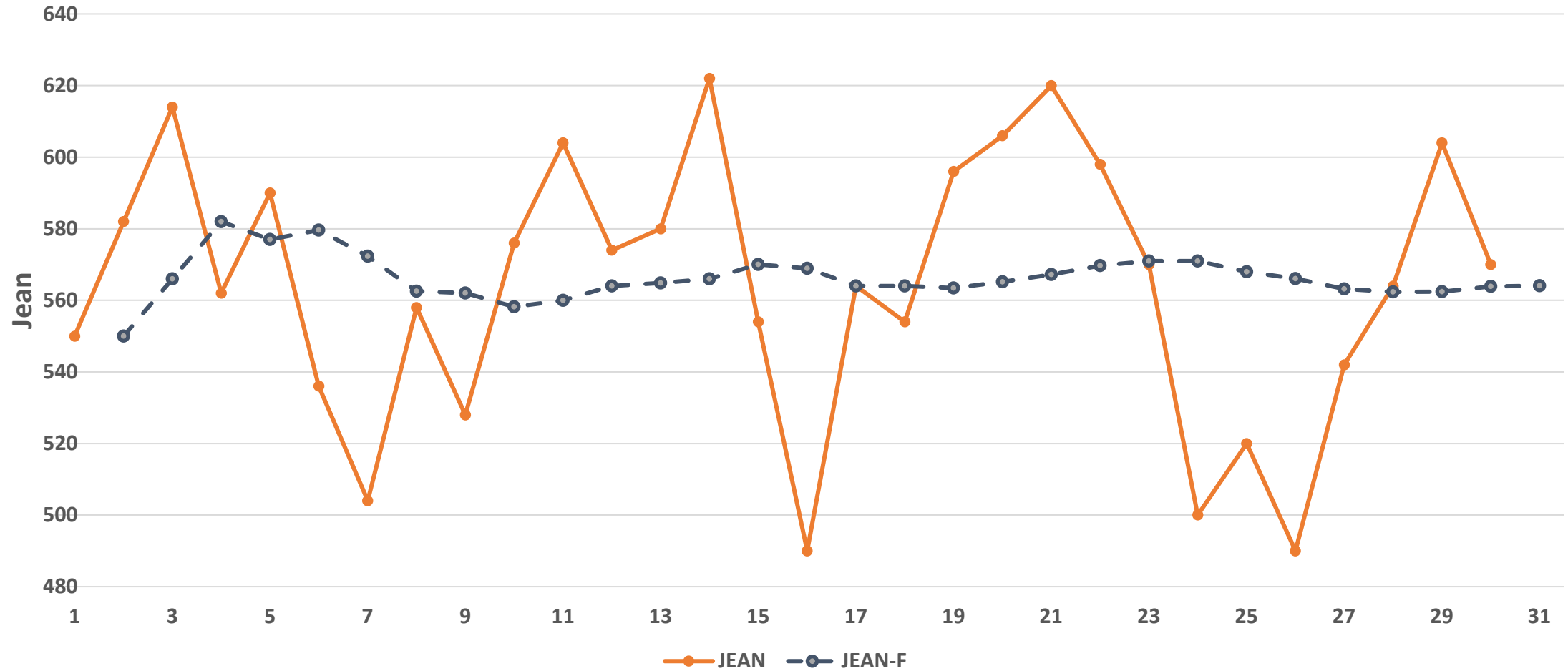
MSE = 15272.7

RMSE = 39.66

U = 0.97

CÁC PHƯƠNG PHÁP DỰ BÁO TRUNG BÌNH (tt)

Dự báo trung bình giản đơn



2. DỰ BÁO TRUNG BÌNH DI ĐỘNG

Với hệ số trượt k , trung bình di động bậc k (ký hiệu MA(k))

$$\hat{Y}_{t+1} = \frac{Y_t + Y_{t-1} + \dots + Y_{t-k+1}}{k}$$

- Khác với dự báo trung bình giản đơn (sử dụng toàn bộ dữ liệu quá khứ làm giá trị dự báo), dự báo trung bình di động chỉ sử dụng một số quan sát gần nhất làm giá trị dự báo.
- **Ví dụ:** dự báo số quần Jean bán/tuần theo MA(5)

CÁC PHƯƠNG PHÁP DỰ BÁO TRUNG BÌNH (tt)

Tuần	TT	DB	Tuần	TT	DB	Tuần	TT	DB
6	536	579.6	15	554	591.2	24	500	598.0
7	504	576.8	16	490	586.8	25	520	578.8
8	558	561.2	17	564	564.0	26	490	561.6
9	528	550.0	18	554	562.0	27	542	535.6
10	576	543.2	19	596	556.8	28	564	524.4
11	604	540.4	20	606	551.6	29	604	523.2
12	574	554.0	21	620	562.0	30	570	544.0
13	580	568.0	22	598	588.0	31		554.0
14	622	572.4	23	570	594.8			

Bảng bên là kết quả dự báo theo trung bình di động

$$\text{MAE} = 45.7$$

$$\text{MAPE} = 0.083$$

$$\text{MPE} = -0.004$$

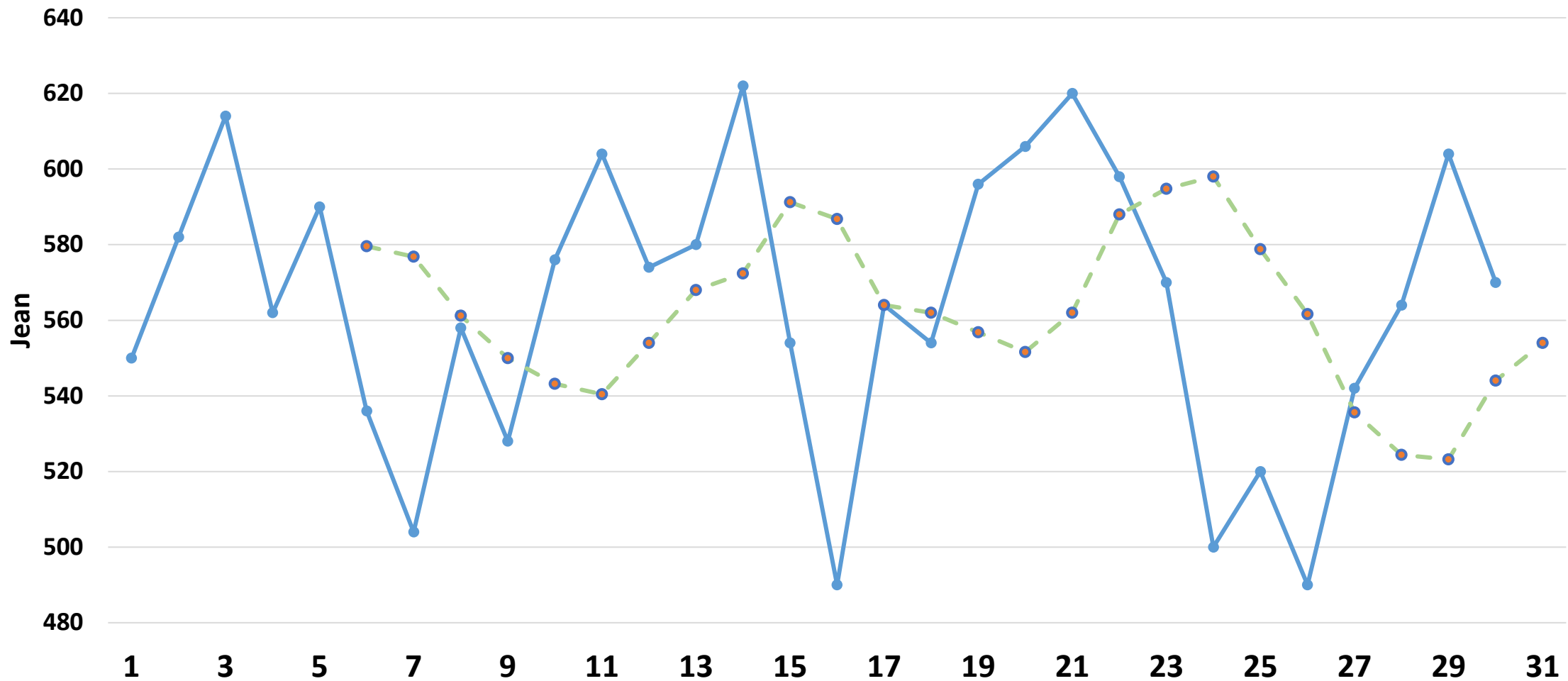
$$\text{MSE} = 2644.5$$

$$\text{RMSE} = 51.4$$

$$U = 1.25$$

CÁC PHƯƠNG PHÁP DỰ BÁO TRUNG BÌNH (tt)

Dự báo trung bình di động MA(5)



PHƯƠNG PHÁP SAN MŨ GIẢN ĐƠN

$$\hat{Y}_{t+1} = \alpha Y_t + (1 - \alpha) \hat{Y}_t$$

α là hệ số san mũ

- Công thức trên có thể triển khai thành:

$$\hat{Y}_{t+1} = \alpha Y_t + (1 - \alpha) \hat{Y}_t = \alpha Y_t + \hat{Y}_t - \alpha \hat{Y}_t$$

$$\hat{Y}_{t+1} = \hat{Y}_t + \alpha(Y_t - \hat{Y}_t) = \hat{Y}_t + \alpha e_t$$

- Với n là số quan sát có sẵn trong mẫu dữ liệu quá khứ:

$$\hat{Y}_{t+1} = \alpha Y_t + \alpha(1 - \alpha)Y_{t-1} + \alpha(1 - \alpha)^2 Y_{t-2} + \cdots + \alpha(1 - \alpha)^n Y_{t-n}$$

PHƯƠNG PHÁP SAN MŨ GIẢN ĐƠN (tt)

Time	Sales	Time	Sales	Time	Sales
2014Q1	500	2016Q3	150	2019Q1	750
2014Q2	350	2016Q4	400	2019Q2	500
2014Q3	250	2017Q1	550	2019Q3	400
2014Q4	400	2017Q2	350	2019Q4	650
2015Q1	450	2017Q3	250	2020Q1	850
2015Q2	350	2017Q4	550	2020Q2	600
2015Q3	200	2018Q1	550	2020Q3	450
2015Q4	300	2018Q2	400	2020Q4	700
2016Q1	350	2018Q3	350		
2016Q2	200	2018Q4	600		

Bảng bên là dữ liệu về doanh thu theo quý của công ty BĐS Hoàng Gia.

Ta dự báo doanh thu theo phương pháp san mũ với α lần lượt là 0.1 và 0.6

PHƯƠNG PHÁP SAN MŨ GIẢN ĐƠN (tt)

Time	Sales	Sales-F	Time	Sales	Sales-F	Time	Sales	Sales-F
2014Q1	500	500.0	2016Q3	150	381.9	2019Q1	750	415.9
2014Q2	350	500.0	2016Q4	400	358.7	2019Q2	500	449.3
2014Q3	250	485.0	2017Q1	550	362.8	2019Q3	400	454.4
2014Q4	400	461.5	2017Q2	350	381.6	2019Q4	650	449.0
2015Q1	450	455.4	2017Q3	250	378.4	2020Q1	850	469.1
2015Q2	350	454.8	2017Q4	550	365.6	2020Q2	600	507.2
2015Q3	200	444.3	2018Q1	550	384.0	2020Q3	450	516.5
2015Q4	300	419.9	2018Q2	400	400.6	2020Q4	700	509.8
2016Q1	350	407.9	2018Q3	350	400.5	2021Q1		528.8
2016Q2	200	402.1	2018Q4	600	395.5			

$\alpha=0.1$

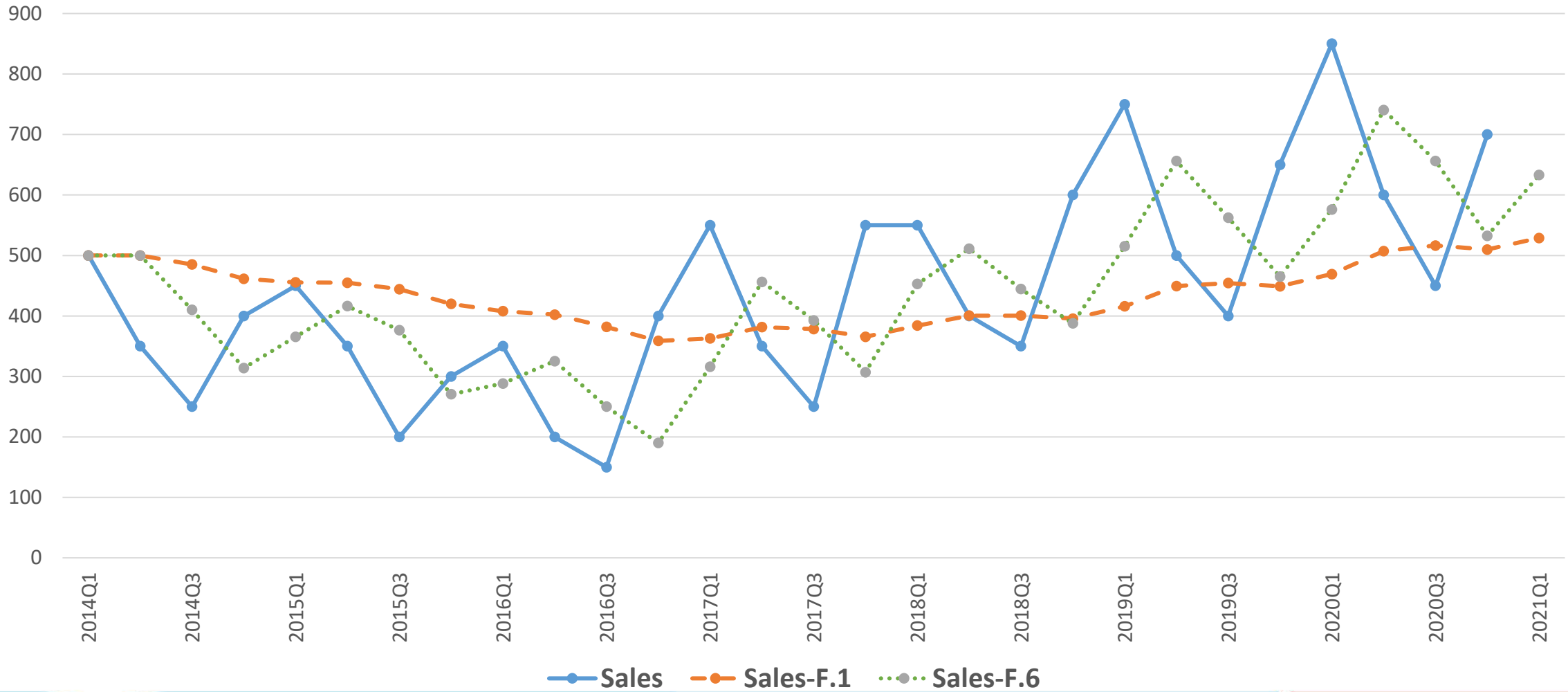
PHƯƠNG PHÁP SAN MŨ GIẢN ĐƠN (tt)

Time	Sales	Sales-F	Time	Sales	Sales-F	Time	Sales	Sales-F
2014Q1	500	500.0	2016Q3	150	250.1	2019Q1	750	515.1
2014Q2	350	500.0	2016Q4	400	190.0	2019Q2	500	656.0
2014Q3	250	410.0	2017Q1	550	316.0	2019Q3	400	562.4
2014Q4	400	314.0	2017Q2	350	456.4	2019Q4	650	465.0
2015Q1	450	365.6	2017Q3	250	392.6	2020Q1	850	576.0
2015Q2	350	416.2	2017Q4	550	307.0	2020Q2	600	740.4
2015Q3	200	376.5	2018Q1	550	452.8	2020Q3	450	656.2
2015Q4	300	270.6	2018Q2	400	511.1	2020Q4	700	532.5
2016Q1	350	288.2	2018Q3	350	444.4	2021Q1		633.0
2016Q2	200	325.3	2018Q4	600	387.8			

$\alpha=0.6$

PHƯƠNG PHÁP SAN MŨ GIẢN ĐƠN (tt)

Dự báo theo phương pháp san mũ giản đơn



PHƯƠNG PHÁP SAN MŨ HOLT

Mô hình san mũ Holt được thể hiện qua ba phương trình dưới đây:

1. Ước lượng giá trị trung bình hiện tại

$$L_t = \alpha Y_t + (1 - \alpha)(L_{t-1} + T_{t-1})$$

2. Ước lượng xu thế (độ dốc)

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1}$$

3. Dự báo p giai đoạn trong tương lai

$$\hat{Y}_{t+p} = L_t + pT_t$$

PHƯƠNG PHÁP SAN MŨ HOLT (tt)

L_t : giá trị san mũ mới (hay giá trị ước lượng trung bình hiện tại)

α : hệ số san mũ của giá trị trung bình ($0 < \alpha < 1$)

β : hệ số san mũ của giá trị xu thế ($0 < \beta < 1$)

T_t : giá trị ước lượng của xu thế

p : thời đoạn dự báo trong tương lai

Y_t : giá trị quan sát hoặc giá trị thực tế vào thời điểm t

\hat{Y}_{t+p} : giá trị báo cho p giai đoạn trong tương lai

PHƯƠNG PHÁP SAN MŨ WINTERS

Mô hình san mũ Holt được thể hiện qua 4 phương trình sau đây:

1. Ước lượng giá trị trung bình hiện tại

$$L_t = \alpha \frac{Y_t}{S_{t-s}} + (1 - \alpha)(L_{t-1} + T_{t-1})$$

2. Ước lượng xu thế (độ dốc): $T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1}$

3. Ước lượng giá trị chỉ số mùa:

$$S_t = \gamma \frac{Y_t}{L_t} + (1 - \gamma)S_{t-s}$$

4. Dự báo p giai đoạn trong tương lai:

$$\hat{Y}_{t+p} = (L_t + pT_t)S_{t-s+p}$$

PHƯƠNG PHÁP SAN MŨ WINTERS

L_t : giá trị san mũ mới (hay giá trị ước lượng trung bình hiện tại)

α : hệ số san mũ của giá trị trung bình ($0 < \alpha < 1$)

β : hệ số san mũ của giá trị xu thế ($0 < \beta < 1$)

γ : hệ số san mũ của chỉ số mùa

T_t : giá trị ước lượng của xu thế

S_t : Giá trị ước lượng của chỉ số mùa

p : thời đoạn dự báo trong tương lai

s : độ dài của yếu tố mùa

DỰ BÁO BẰNG CÁC MÔ HÌNH XU THẾ

- **Xu thế** là sự vận động tăng hay giảm của dữ liệu trong một thời gian dài. Sự vận động này có thể được mô tả bằng một đường thẳng (xu thế tuyến tính) hoặc bởi một vài dạng đường cong toán học (xu thế phi tuyến)
- Có thể nhận dạng xu thế qua việc vẽ đồ thị của biến Y_t theo thời gian t
- Trong mô hình, biến thời gian t thường được sử dụng là biến giải thích

DỰ BÁO BẰNG CÁC MÔ HÌNH XU THẾ (tt)

Dạng hàm xu thế

Phương trình hồi quy tổng thể

Tuyến tính

$$Y_t = \beta_0 + \beta_1 t + u_t$$

Bậc hai

$$Y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + u_t$$

Bậc ba

$$Y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + u_t$$

Lin-log

$$Y_t = \beta_0 + \beta_1 \ln t + u_t$$

Log – lin

$$\ln Y_t = \beta_0 + \beta_1 t + u_t$$

Nghịch đảo

$$Y_t = \beta_0 + \beta_1 \left(\frac{1}{t}\right) + u_t$$

Tăng trưởng mũ

$$Y_t = e^{\beta_0 + \beta_1 t + u_t}$$

DỰ BÁO BẰNG PHƯƠNG PHÁP PHÂN TÍCH

1. CÁC THÀNH PHẦN CỦA CHUỖI THỜI GIAN

- **Xu thế (Trend):** là thành phần thể hiện sự tăng (hoặc giảm) ẩn bên trong của một chuỗi thời gian. Ký hiệu Tr hoặc T.
- **Chu kỳ (Cyclical):** là một chuỗi những sự dao động giống như hình song và sự dao động này sẽ lặp lại sau một thời kỳ thường đại hơn 1 năm. Ký hiệu Cl hoặc C.
- **Mùa vụ (Seasonal):** sự dao động mùa liên quan đến kiểu thay đổi khá ổn định xuất hiện hàng năm và kiểu thay đổi đó được lặp lại ở năm sau và các năm sau nữa. Ký hiệu Sn hay S.
- **Ngẫu nhiên/bất thường (Irregular):** bao gồm những thay đổi ngẫu nhiên hay không dự đoán được. Ký hiệu Ir hay I.

DỰ BÁO BẰNG PHƯƠNG PHÁP PHÂN TÍCH (tt)

Lưu ý:

- Thành phần chu kỳ cần có một chuỗi dữ liệu lưu trữ ít nhất là trên 30 năm.
- Các dao động bất thường thì không thể nào dự đoán được.
- Phương pháp dự báo phân tích chủ yếu đề cập hai thành phần là xu thế và mùa vụ.
- Xem xét hai thành phần trên liên quan như thế nào với chuỗi dữ liệu gốc: Mô hình nhân tính và Mô hình cộng tính.

DỰ BÁO BẰNG PHƯƠNG PHÁP PHÂN TÍCH (tt)

- Mô hình nhân tính:

$$Y_t = Tr_t \cdot Cl_t \cdot Sn_t \cdot Ir_t$$

Phù hợp khi sự biến thiên của chuỗi thời gian tăng dần theo thứ tự của thời gian

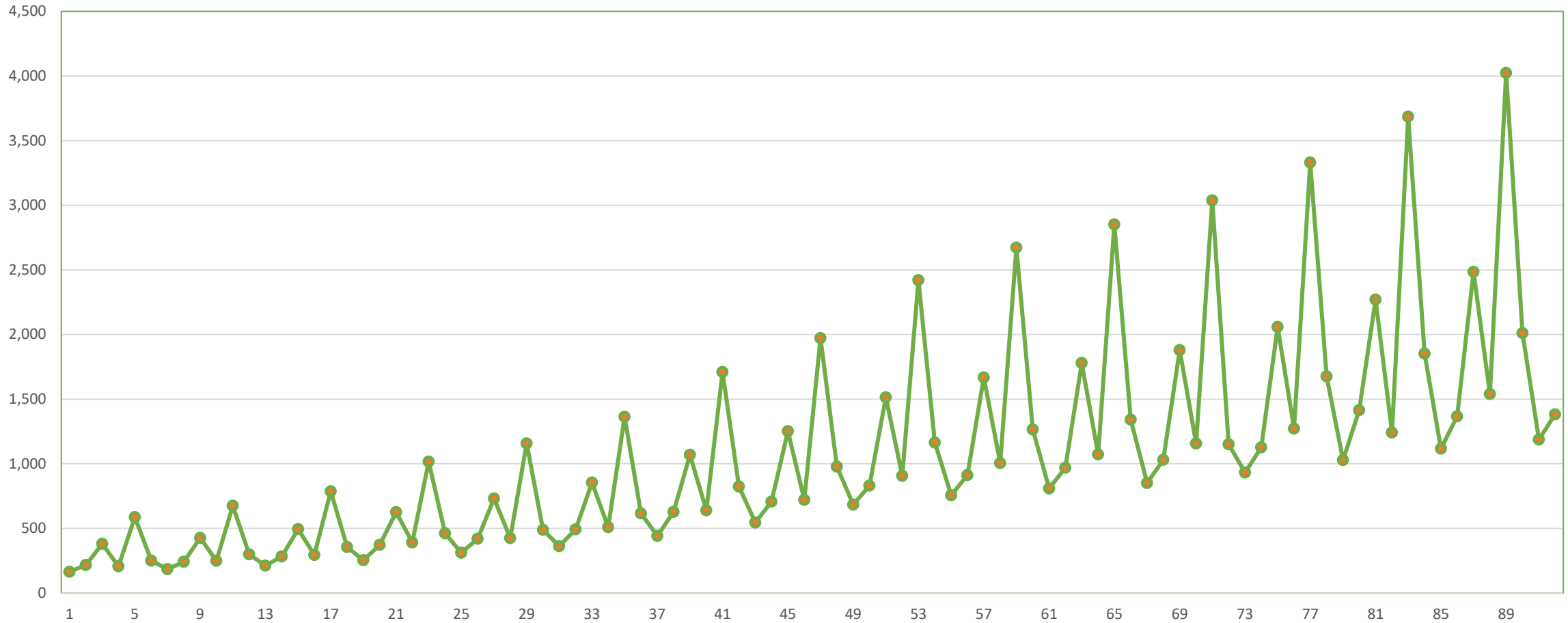
- Mô hình cộng tính:

$$Y_t = Tr_t + Cl_t + Sn_t + Ir_t$$

Có hiệu quả khi chuỗi dữ liệu phân tích có sự biến thiên xấp xỉ đều nhau suốt độ dài của chuỗi thời gian

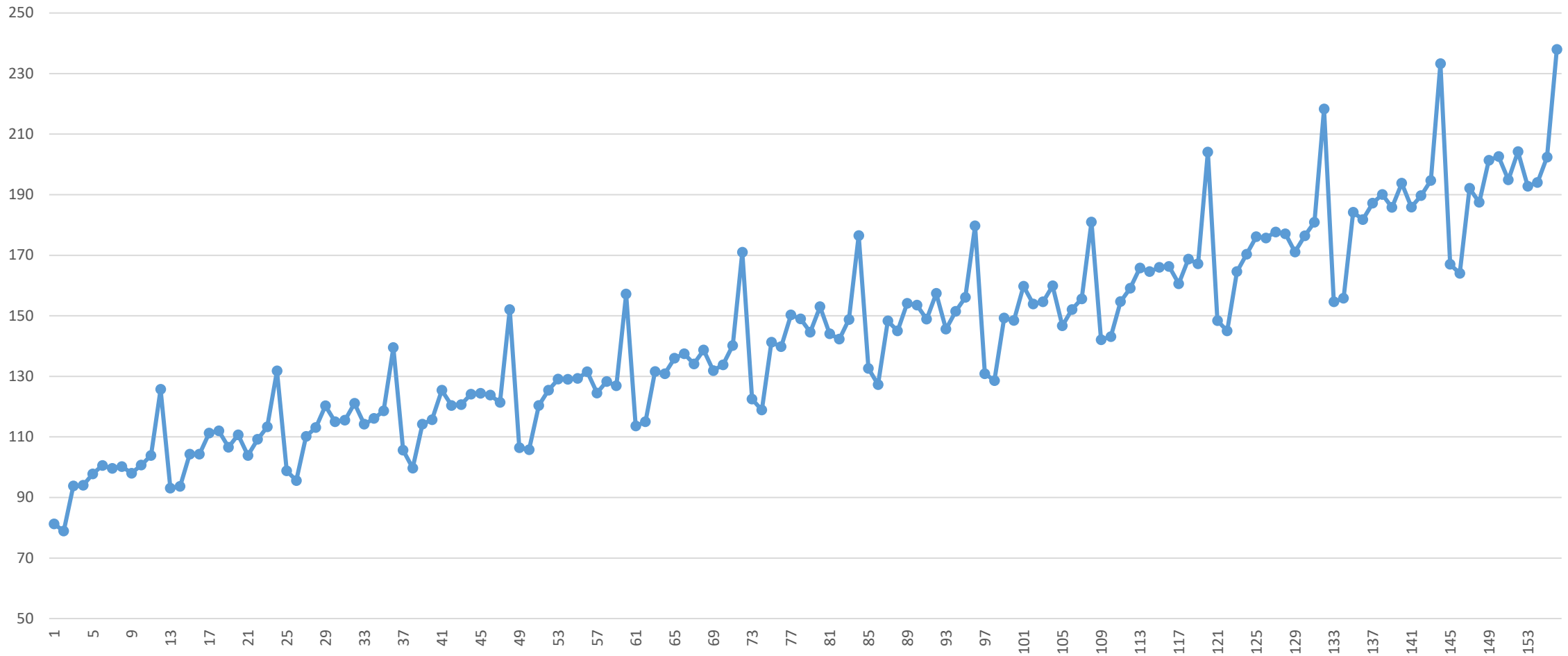
DỰ BÁO BẰNG PHƯƠNG PHÁP PHÂN TÍCH (tt)

Đặc trưng mô hình nhân tính



DỰ BÁO BẰNG PHƯƠNG PHÁP PHÂN TÍCH (tt)

Đặc trưng mô hình cộng tính



DỰ BÁO BẰNG PHƯƠNG PHÁP PHÂN TÍCH (tt)

2. DỮ LIỆU ĐƯỢC ĐIỀU CHỈNH YẾU TỐ MÙA

2.1. Tỷ lệ trung bình di động – mô hình nhân tính

B1: Tính toán trung bình trung tâm

- $CMA_t = (0.5Y_{t+6} + \dots + Y_t + \dots + 0.5Y_{t-6})/12$ (SL theo tháng)
- $CMA_t = (0.5Y_{t+2} + \dots + Y_t + \dots + 0.5Y_{t-1})/4$ (SL theo quý)

Trong mô hình dự báo nhân tính: $CMA_t = Tr_t \cdot Cl_t$

B2: Tính tỷ lệ: $\tau_t = Y_t / CMA_t$

Với mô hình nhân tính: $\tau_t = Sn_t \cdot Ir_t$

DỰ BÁO BẰNG PHƯƠNG PHÁP PHÂN TÍCH (tt)

B3: Tính toán các chỉ số mùa vụ

- Với chuỗi dữ liệu theo tháng, chỉ số mùa i_m cho tháng m bằng trung bình của τ_t với các quan sát chỉ cho tháng m (mỗi năm có 1 tháng m)
- Với chuỗi dữ liệu theo quý, chỉ số mùa i_q cho quý q bằng trung bình của τ_t với các quan sát chỉ cho quý q (mỗi năm có 1 quý q)

B4: Điều chỉnh các chỉ số mùa để tích của chúng bằng 1. Điều này được thực hiện bằng cách tính các nhân tố mùa:

DỰ BÁO BẰNG PHƯƠNG PHÁP PHÂN TÍCH (tt)

- $S_n = \frac{i_m}{\sqrt[12]{i_1 \cdot i_2 \dots i_{12}}}$ nếu dữ liệu theo tháng
- $S_n = \frac{i_q}{\sqrt[4]{i_1 \cdot i_2 \cdot i_3 \cdot i_4}}$ nếu dữ liệu theo tháng

B5: Chuỗi dữ liệu đã hiệu chỉnh yếu tố mùa được xác định bởi

$$\frac{Y_t}{S_{n_t}} = Tr_t \cdot Cl_t \cdot Ir_t$$

Nếu không có yếu tố chu kỳ và yếu tố ngẫu nhiên bị triệt tiêu khi xét trung bình để tìm chỉ số mùa τ_t ($Cl_t = 1, Ir_t = 1$), do vậy chuỗi Y_t/S_{n_t} chỉ còn lại yếu tố xu thế.

DỰ BÁO BẰNG PHƯƠNG PHÁP PHÂN TÍCH (tt)

2.2. Chênh lệch so với trung bình di động – mô hình nhân tính

B1: Tính toán trung bình trung tâm

- $CMA_t = (0.5Y_{t+6} + \dots + Y_t + \dots + 0.5Y_{t-6})/12$

(nếu số liệu theo tháng)

- $CMA_t = (0.5Y_{t+2} + \dots + Y_t + \dots + 0.5Y_{t-1})/4$

(nếu số liệu theo quý)

B2: Tính sự khác biệt:

$$d_t = Y_t - CMA_t$$

DỰ BÁO BẰNG PHƯƠNG PHÁP PHÂN TÍCH (tt)

B3: Tính toán các chỉ số mùa vụ

- Với chuỗi dữ liệu theo tháng, chỉ số mùa i_m cho tháng m bằng trung bình của d_t với các quan sát chỉ cho tháng m (mỗi năm có 1 tháng m)
- Với chuỗi dữ liệu theo quý, chỉ số mùa i_q cho quý q bằng trung bình của d_t , với các quan sát chỉ cho quý q (mỗi năm có 1 quý q)

B4: Điều chỉnh các chỉ số mùa để tổng của chúng bằng 0. Điều này được thực hiện bởi thiết lập: $S_n = i_t - \bar{i}$

Với \bar{i} là trung bình của tất cả các chỉ số mùa.

DỰ BÁO BẰNG PHƯƠNG PHÁP PHÂN TÍCH (tt)

B5: Chuỗi dữ liệu đã điều chỉnh yếu tố mùa có được bằng cách lấy

$$Y_t - Sn_t = Tr_t + Cl_t + Ir_t$$

Nếu không có yếu tố chu kỳ, trong khi yếu tố ngẫu nhiên bị triệt tiêu khi xét trung bình để tìm chỉ số mùa ($Cl_t = 0, Ir_t = 0$), do vậy chuỗi $Y_t - Sn_t$ chỉ còn lại yếu tố xu thế.

DỰ BÁO BẰNG PHƯƠNG PHÁP PHÂN TÍCH (tt)

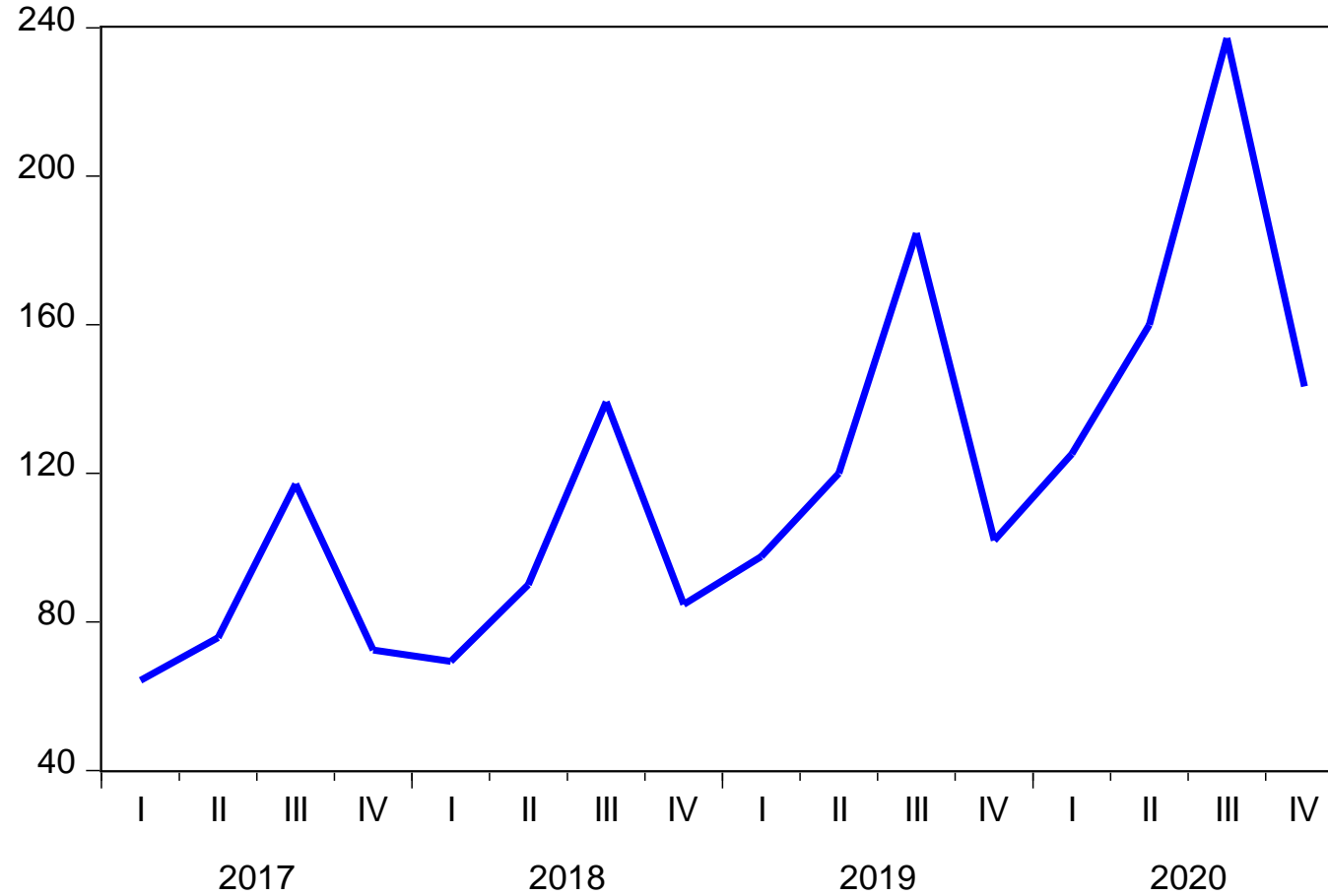
2.3. Dự báo với mô hình nhân tính

Ta có dữ liệu về doanh thu của một công ty Y (tỷ VNĐ). Ta thực hiện dự báo doanh thu cho các quý năm 2021

Năm	Quý	Y	Năm	Quý	Y
2017	1	64.2	2019	1	97.6
2017	2	75.7	2019	2	120
2017	3	117.1	2019	3	184.7
2017	4	72.4	2019	4	101.9
2018	1	69.4	2020	1	125.2
2018	2	90	2020	2	160
2018	3	139.3	2020	3	237.2
2018	4	84.7	2020	4	143.4

DỰ BÁO BẰNG PHƯƠNG PHÁP PHÂN TÍCH (tt)

Đồ thị doanh số theo quý
Y



DỰ BÁO BẰNG PHƯƠNG PHÁP PHÂN TÍCH (tt)

Kết quả điều chỉnh yếu tố mùa

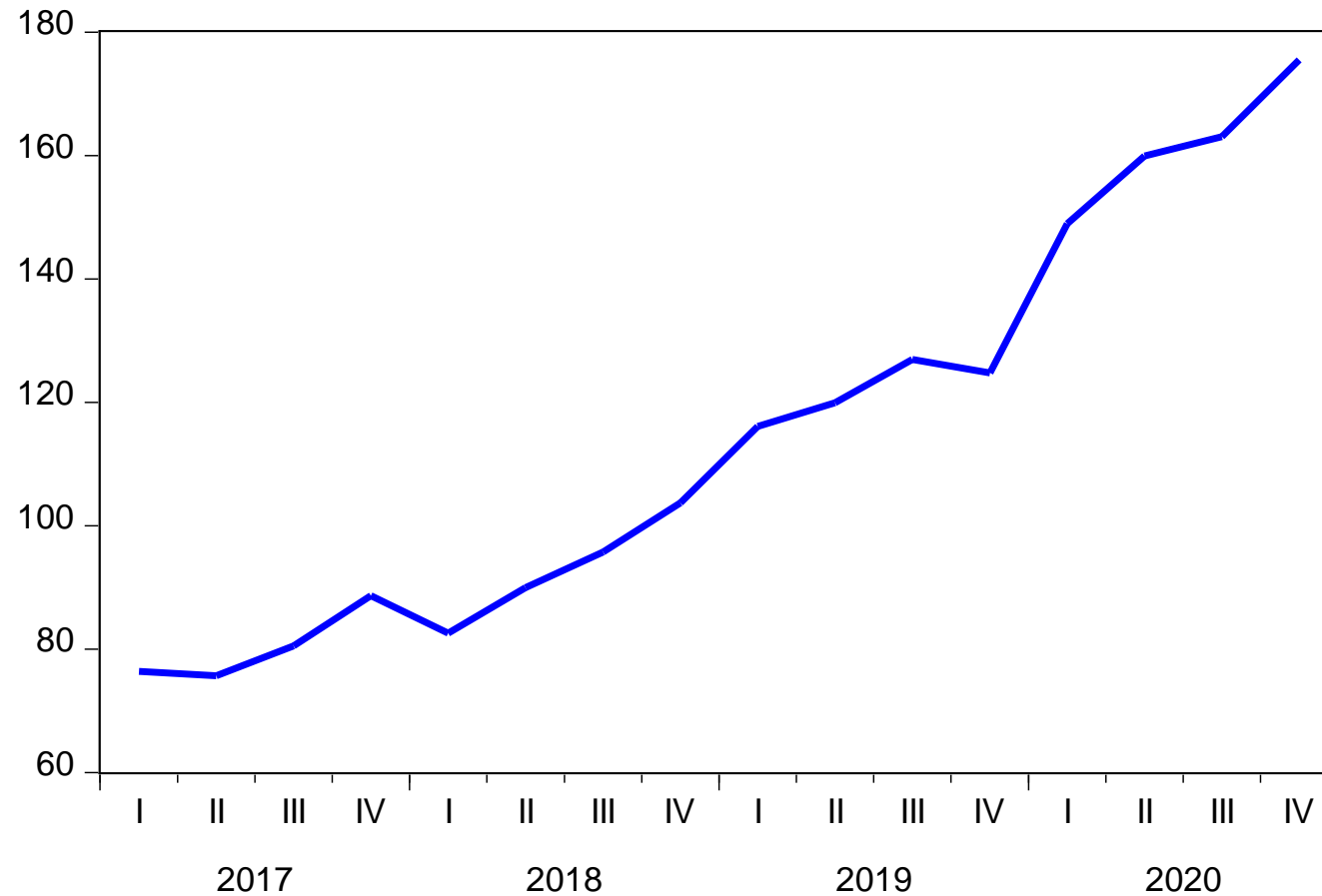
Quý	Y	SN	YSA	Quý	Y	SN	YSA
2017Q1	64.20	0.84	76.36	2019Q1	97.60	0.84	116.09
2017Q2	75.70	1.00	75.66	2019Q2	120.00	1.00	119.93
2017Q3	117.10	1.45	80.48	2019Q3	184.70	1.45	126.94
2017Q4	72.40	0.82	88.61	2019Q4	101.90	0.82	124.72
2018Q1	69.40	0.84	82.55	2020Q1	125.20	0.84	148.92
2018Q2	90.00	1.00	89.95	2020Q2	160.00	1.00	159.91
2018Q3	139.30	1.45	95.74	2020Q3	237.20	1.45	163.03
2018Q4	84.70	0.82	103.67	2020Q4	143.40	0.82	175.51

SN là nhân tố mùa, YSA là chuỗi dữ liệu đã được điều chỉnh yếu tố mùa

DỰ BÁO BẰNG PHƯƠNG PHÁP PHÂN TÍCH (tt)

Đồ thị doanh số theo quý sau khi điều chỉnh yếu tố mùa

YSA



DỰ BÁO BẰNG PHƯƠNG PHÁP PHÂN TÍCH (tt)

Dependent Variable: LOG(YSA)

Method: Least Squares

Date: 11/24/21 Time: 20:53

Sample: 2017Q1 2020Q4

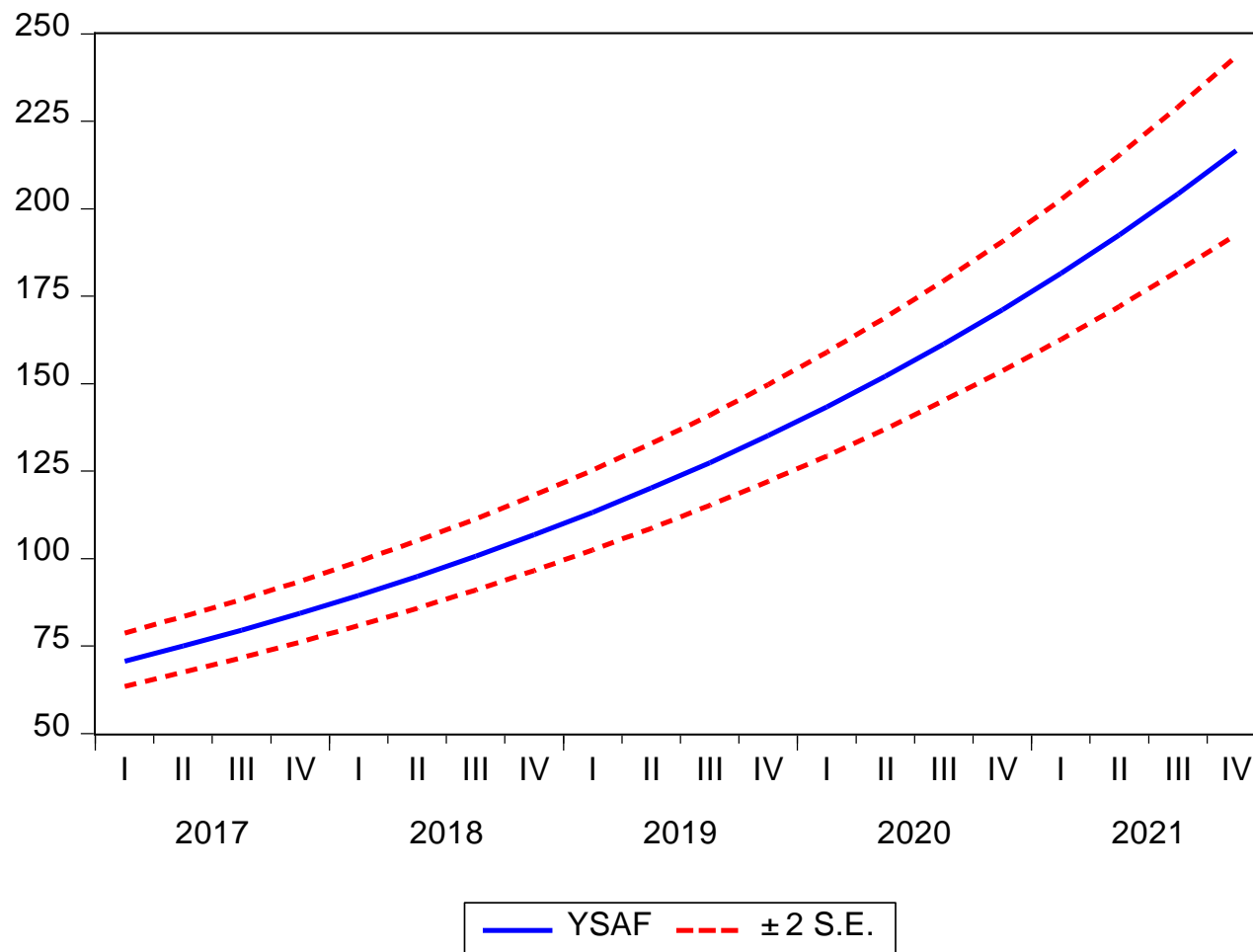
Included observations: 16

Kết quả ước lượng yếu tố xu thế

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4.198907	0.025540	164.4053	0.0000
T	0.058953	0.002641	22.32002	0.0000
R-squared	0.972666	Mean dependent var		4.700011
Adjusted R-squared	0.970714	S.D. dependent var		0.284591
S.E. of regression	0.048703	Akaike info criterion		-3.089692
Sum squared resid	0.033207	Schwarz criterion		-2.993118
Log likelihood	26.71753	Hannan-Quinn criter.		-3.084746
F-statistic	498.1832	Durbin-Watson stat		1.492009
Prob(F-statistic)	0.000000			

DỰ BÁO BẰNG PHƯƠNG PHÁP PHÂN TÍCH (tt)

Kết quả dự báo các quý năm 2021 cho YSA



Forecast: YSAF

Actual: YSA

Forecast sample: 2017Q1 2021Q4

Included observations: 20

Root Mean Squared Error 4.938460

Mean Absolute Error 4.062936

Mean Abs. Percent Error 3.748212

Theil Inequality Coefficient 0.020854

Bias Proportion 0.001655

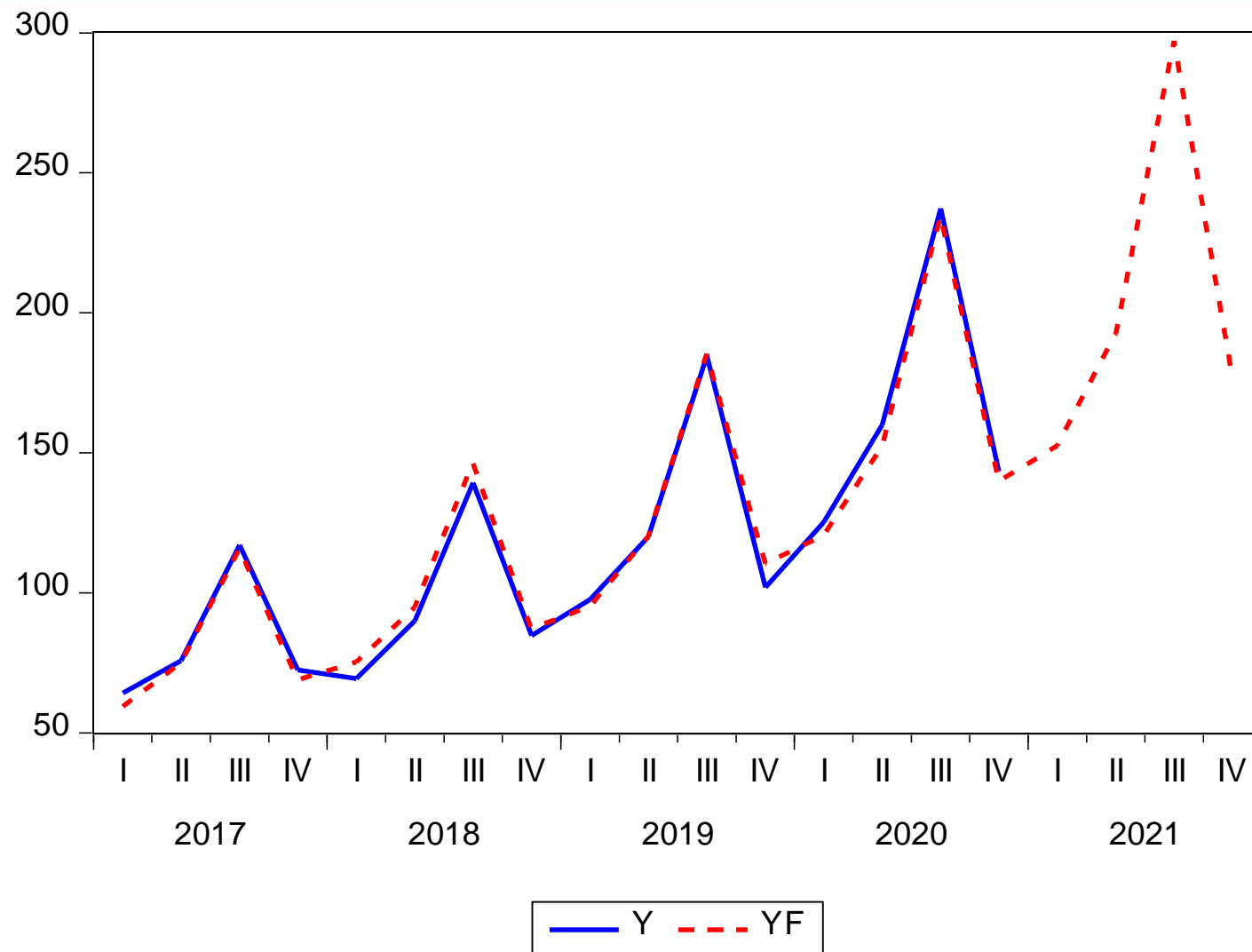
Variance Proportion 0.074397

Covariance Proportion 0.923948

Theil U2 Coefficient 0.529458

Symmetric MAPE 3.729159

DỰ BÁO BẰNG PHƯƠNG PHÁP PHÂN TÍCH (tt)



Kết quả dự báo các quý năm 2021 cho doanh số Y

Quý	YF
2021Q1	152.57
2021Q2	192.60
2021Q3	297.09
2021Q4	176.95



BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC QUY NHƠN

KẾT THÚC CHƯƠNG 2



BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC QUY NHƠN

CHƯƠNG 3

PHƯƠNG PHÁP PHÂN TÍCH HỒI QUY



Goal 1

Hiểu rõ các vấn đề cơ bản về phân tích hồi quy

Goal 2

Áp dụng được mô hình hồi quy để dự báo nhằm hỗ trợ cho việc ra quyết định

Goal 3

Sử dụng được phần mềm Eviews để ước lượng, kiểm định và dự báo từ các mô hình hồi quy

PHƯƠNG PHÁP PHÂN TÍCH HỒI QUY

1. Mô hình hồi quy đơn
2. Mô hình hồi quy bội
3. Suy diễn thống kê và dự báo
 - Ước lượng khoảng tin cậy
 - Kiểm định giả thuyết thống kê
 - Dự báo từ mô hình hồi quy

MÔ HÌNH HỒI QUY TUYẾN TÍNH

Mô hình hồi quy tuyến tính k biến

$$Y_t = \beta_1 + \beta_2 X_{2t} + \cdots + \beta_k X_{kt} + u_t$$

Với giả thiết $E(u_t | X_{2t}, \dots, X_{kt}) = 0$. Khi đó

$$E(Y_t | X_{2t}, \dots, X_{kt}) = \beta_1 + \beta_2 X_{2t} + \cdots + \beta_k X_{kt}$$

được gọi là ***hàm hồi quy tổng thể***

Y_t : biến phụ thuộc (biến được giải thích)

X_{2t}, \dots, X_{kt} : biến độc lập (biến giải thích)

$\beta_1, \beta_2, \dots, \beta_k$: các hệ số hồi quy; u_t : sai số ngẫu nhiên (hay nhiễu)

MÔ HÌNH HỒI QUY TUYẾN TÍNH (tt)

$$E(Y_t | X_{2t}, \dots, X_{kt}) = \beta_1 + \beta_2 X_{2t} + \dots + \beta_k X_{kt}$$

- + β_1 : hệ số chặn: giá trị trung bình của biến phụ thuộc khi tất cả các biến độc lập bằng 0
- + β_j ($j = 2, \dots, k$): hệ số hồi quy riêng của biến X_j , khi X_j tăng (giảm) 1 đơn vị thì trung bình biến phụ thuộc thay đổi β_j đơn vị trong điều kiện các yếu tố khác không đổi
- Trong thực tế, các hệ số hồi quy thường không biết nên ta ước lượng chúng bằng phương pháp bình phương nhỏ nhất thông thường (OLS – Ordinary Least Squares) thông qua một mẫu

MÔ HÌNH HỒI QUY TUYẾN TÍNH (tt)

Phương pháp OLS để ước lượng các hệ số hồi quy từ mẫu kích thước n đối với mô hình hồi quy tuyến tính hai biến

$$\hat{\beta}_2 = \frac{n\sum Y_t X_t - \sum X_t \sum Y_t}{n\sum X_t^2 - (\sum X_t)^2} = \frac{\sum x_t y_t}{\sum x_t^2}$$

$$\hat{\beta}_1 = \frac{\sum X_t^2 \sum Y_t - \sum X_t \sum (X_t Y_t)}{n\sum X_t^2 - (\sum X_t)^2} = \bar{Y} - \hat{\beta}_2 \bar{X}$$

Với $\bar{X} = \frac{1}{n} \sum_{t=1}^n X_t$; $\bar{Y} = \frac{1}{n} \sum_{t=1}^n Y_t$; $x_t = X_t - \bar{X}$; $y_t = Y_t - \bar{Y}$

MÔ HÌNH HỒI QUY TUYẾN TÍNH (tt)

Các giả thiết của mô hình:

1. Mô hình có dạng tuyến tính và được xác định đúng
2. $E(u_t) = 0$
3. $u_t \sim N(\mu, \sigma^2)$
4. X_t và u_t không tương quan: $\text{cov}(X_t, u_t) = 0$
5. Không có đa cộng tuyến hoàn hảo giữa các biến độc lập
6. Phương sai không đổi: $\text{var}(u_t) = \sigma^2$
7. Không có tương quan chuỗi: $\text{cov}(u_s, u_t) = 0, \forall t \neq s$

MÔ HÌNH HỒI QUY TUYẾN TÍNH (tt)

Đặc điểm của các ước lượng OLS khi thỏa mãn các giả thiết:

1. Các ước lượng OLS đạt tiêu chuẩn BLUE (Best linear unbiased estimator)
2. Mức độ chính xác của các ước lượng OLS có thể đánh giá qua phương sai của các hệ số hồi quy $var(\hat{\beta}_j)$

Với MH 2 biến: $E(Y_t|X_{2t}) = \beta_1 + \beta_2 X_{2t}$

$$var(\hat{\beta}_1) = \frac{\sum X_t^2}{n \sum X_t^2} \sigma^2, var(\hat{\beta}_2) = \frac{\sigma^2}{\sum X_t^2}$$

Ước lượng của σ^2 là $\hat{\sigma}^2 = \frac{\sum (Y_t - \hat{Y}_t)^2}{n-2}$

MÔ HÌNH HỒI QUY TUYẾN TÍNH (tt)

Với MH 3 biến: $E(Y_t|X_{2t}, X_{3t}) = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t}$

$$\text{var}(\hat{\beta}_1) = \left[\frac{1}{n} + \frac{\bar{X}_2^2 \sum x_{3t}^2 + \bar{X}_3^2 \sum x_{2t}^2 - 2\bar{X}_2 \bar{X}_3 \sum x_{2t} x_{3t}}{(\sum x_{2t}^2)(\sum x_{3t}^2) - (\sum x_{2t} x_{3t})^2} \right] \sigma^2,$$

$$\text{var}(\hat{\beta}_2) = \frac{\sigma^2}{\sum x_{2t}^2 (1 - r_{23}^2)},$$

$$\text{var}(\hat{\beta}_3) = \frac{\sigma^2}{\sum x_{3t}^2 (1 - r_{23}^2)}$$

với ước lượng không chệch của σ^2 là

$$\hat{\sigma}^2 = \frac{\sum (Y_t - \hat{Y}_t)^2}{n - 3}$$

MÔ HÌNH HỒI QUY TUYẾN TÍNH (tt)

Nhận xét:

- Trường hợp hồi quy nhiều biến, phương sai của các hệ số hồi quy phụ thuộc vào phương sai hạng nhiễu, mẫu và mối tương quan giữa các biến giải thích (hệ số tương quan của biến X_i và X_j là r_{ij} và $-1 \leq r_{ij} \leq 1$)
- Trong phân tích hồi quy thường sử dụng sai số chuẩn của các hệ số, ký hiệu $se(\hat{\beta}_j)$

$$se(\hat{\beta}_j) = \sqrt{var(\hat{\beta}_j)}$$

MÔ HÌNH HỒI QUY TUYẾN TÍNH (tt)

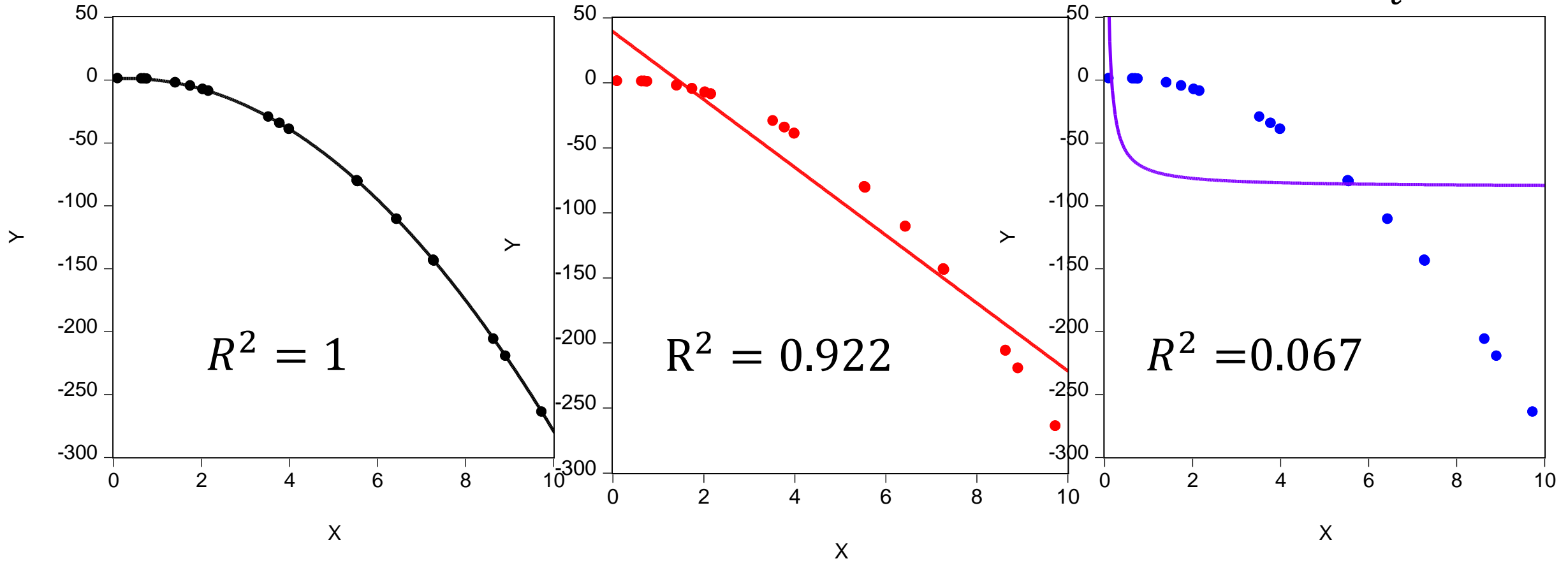
Đánh giá mức độ phù hợp của mô hình:

- Hệ số xác định R^2 ($0 \leq R^2 \leq 1$):

- $R^2 = 1$: MH giải thích 100% sự biến thiên của Y
- $R^2 = 0$: MH không giải thích được gì cho sự biến thiên của Y
- R^2 càng gần 1 thì MH có độ phù hợp càng cao (với dữ liệu mẫu)
- R^2 càng gần 0 thì MH có độ phù hợp càng thấp (với dữ liệu mẫu)

MÔ HÌNH HỒI QUY TUYẾN TÍNH (tt)

$$Y_t = \beta_1 + \beta_2 X_t + \beta_3 X_t^2 + u_t \quad Y_t = \beta_1 + \beta_2 X_t + u_t \quad Y_t = \beta_1 + \beta_2 \left(\frac{1}{X_t}\right) + u_t$$



MÔ HÌNH HỒI QUY TUYẾN TÍNH (tt)

Lưu ý:

- R^2 là hàm tăng theo số biến giải thích
- Khi so sánh mức độ phù hợp giữa các mô hình (cùng dạng biến phụ thuộc) khác nhau về số biến độc lập ta sử dụng hệ số xác định hiệu chỉnh

$$\bar{R}^2 = 1 - \frac{\frac{\sum \hat{u}_t^2}{n - k}}{\frac{\sum y_t^2}{n - 1}} = 1 - (1 - R^2) \frac{n - 1}{n - k}$$

k là số hệ số ước lượng trong mô hình

MÔ HÌNH HỒI QUY TUYẾN TÍNH (tt)

Lưu ý:

- Vấn đề hồi quy giả mạo dẫn đến R^2 rất cao
- Tương quan mạnh giữa các biến giải thích dẫn đến R^2 rất cao
- Tương quan không nhất thiết hàm ý quan hệ nhân quả
- Hệ số R^2 không có nghĩa chọn lựa sai biến giải thích X_t
- Các giá trị R^2 từ các phương trình với biến phụ thuộc khác nhau không thể so sánh được

MÔ HÌNH HỒI QUY TUYẾN TÍNH (tt)

Các tiêu chí lựa chọn mô hình:

$$1. AIC = \left(\frac{RSS}{n}\right) \hat{u}^{2k/n}; \quad \ln(AIC) = \ln\left(\frac{RSS}{n}\right) + \frac{2k}{n}$$

$$2. SBC = \left(\frac{RSS}{n}\right) \hat{u}^{k/n}; \quad \ln(SBC) = \ln\left(\frac{RSS}{n}\right) + \frac{k}{n}$$

$$3. FPE = \left(\frac{RSS}{n}\right) \frac{n+k}{n-k}$$

$$4. HQC = \left(\frac{RSS}{n}\right) (\ln(n))^{2k/n}$$

Ta chọn MH với các tiêu chí trên sao cho chúng có giá trị nhỏ nhất; AIC và SBC thường được ưu tiên với chuỗi thời gian

HỆ SỐ HỒI QUY CHUẨN HÓA

- Để xác định biến giải thích nào có ảnh hưởng nhiều hơn đến biến phụ thuộc, ta cần phải dựa vào các hệ số hồi quy chuẩn hóa
- Hệ số hồi quy chuẩn hóa cho biết tầm quan trọng tương đối của các biến giải thích trong một mô hình hồi quy
- Để ước lượng hệ số hồi quy chuẩn hóa, ta cần chuyển hóa mỗi biến (cả biến phụ thuộc) sang dạng biến chuẩn hóa, sau đó dùng phương pháp OLS để ước lượng
- Ta chuyển hóa mô hình hồi quy tuyến tính k biến

$$Y_t = \beta_1 + \beta_2 X_{2t} + \dots + \beta_k X_{kt} + u_t$$

HỆ SỐ HỒI QUY CHUẨN HÓA (tt)

Mô hình trên được chuyển hóa như sau:

$$\frac{Y_t - \bar{Y}}{S_Y} = \beta_2^* \frac{X_{2t} - \bar{X}_2}{S_{X_2}} + \dots + \beta_k^* \frac{X_{kt} - \bar{X}_k}{S_{X_k}} + u_t$$

Quan hệ giữa hệ số hồi quy chuẩn hóa với các hệ số hồi quy riêng

$$\beta_j^* = \beta_j \frac{S_{X_j}}{S_Y} \quad (*)$$

Giả sử một hệ số hồi quy chuẩn hóa $\hat{\beta}_j^* = 0.7$ sẽ cho biết một sự thay đổi bằng 1 độ lệch chuẩn của biến X_j sẽ dẫn đến sự thay đổi 0.7 độ lệch chuẩn trong biến phụ thuộc

HỆ SỐ HỒI QUY CHUẨN HÓA (tt)

Trong Eviews, ta thực hiện các bước như sau:

- **Bước 1:** Ước lượng mô hình hồi quy theo OLS (giả sử đó là mô hình tốt nhất)
- **Bước 2:** Tính các độ lệch chuẩn của tất cả các biến
$$\text{scalar sy}=\text{scalar}(Y)$$
$$\text{scalar sx}=\text{scalar}(X)$$
- **Bước 3:** Tính hệ số hồi quy chuẩn hóa theo công thức (*) ở trên

ỨNG DỤNG DỰ BÁO

Dự báo trung bình:

Xét hàm hồi quy hai biến: $E(Y_t|X_t) = \beta_1 + \beta_2 X_t$

- Dự báo điểm của $E(Y_t|X_t)$ tại điểm X^0 là

$$\hat{Y}_0 = \hat{\beta}_1 + \hat{\beta}_2 X^0$$

- Khoảng tin cậy cho giá trị dự báo trung bình tại X^0

$$\left(\hat{Y}_0 - t_{\frac{\alpha}{2}}^{(n-2)} \cdot se(\hat{Y}_0) \leq E(Y|X_0) \leq \hat{Y}_0 + t_{\frac{\alpha}{2}}^{(n-2)} \cdot se(\hat{Y}_0) \right)$$

$$\text{Trong đó: } se(\hat{Y}_0) = \sqrt{var(\hat{Y}_0)} = \sqrt{\hat{\sigma}^2 \left[\frac{1}{n} + \frac{(X_0 - \bar{X})^2}{\sum(X - \bar{X})^2} \right]}$$

ỨNG DỤNG DỰ BÁO (tt)

Dự báo riêng biệt:

- Khoảng tin cậy cho giá trị dự báo riêng biệt tại X_0

$$\left(\hat{Y}_0 - t_{\alpha/2}^{(n-2)} \cdot se(Y_0 - \hat{Y}_0) \leq Y_0 \leq \hat{Y}_0 + t_{\alpha/2}^{(n-2)} \cdot se(Y_0 - \hat{Y}_0) \right)$$

Trong đó: $se(Y_0 - \hat{Y}_0) = \sqrt{var(Y_0 - \hat{Y}_0)} = \sqrt{\hat{\sigma}^2 \left[1 + \frac{1}{n} + \frac{(X_0 - \bar{X})^2}{\sum(X - \bar{X})^2} \right]}$

$$var(Y_0 - \hat{Y}_0) = \hat{\sigma}^2 + var(\hat{Y}_0)$$

* Các KTC cho dự báo trung bình và riêng biệt đều được tính toán dễ dàng từ Eviews trong cả trường hợp nhiều biến

ỨNG DỤNG DỰ BÁO (tt)

Ví dụ

Khảo sát số lượng sản phẩm A bán ra mỗi tuần Y (sản phẩm/tuần) theo giá bán X_1 (trăm nghìn đồng) và chi phí quảng cáo X_2 (triệu đồng/tuần).

Ta sử dụng mô hình tuyến tính sau để dự báo số lượng sản phẩm bán ra mỗi tuần nếu bán với giá $X_1 = 8.5$ trăm nghìn đồng và quảng cáo với chi phí $X_2 = 4$ triệu đồng/tuần:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + u$$

ỨNG DỤNG DỰ BÁO (tt)

Dependent Variable: Y
Method: Least Squares
Date: 11/24/21 Time: 22:20
Sample: 1 15
Included observations: 15

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	306.5262	114.2539	2.682851	0.0199
X1	-24.97509	10.83213	-2.305650	0.0398
X2	74.13096	25.96732	2.854779	0.0145

R-squared	0.521478	Mean dependent var	399.3333
Adjusted R-squared	0.441724	S.D. dependent var	63.52352
S.E. of regression	47.46341	Akaike info criterion	10.73465
Sum squared resid	27033.31	Schwarz criterion	10.87626
Log likelihood	-77.50989	Hannan-Quinn criter.	10.73314
F-statistic	6.538607	Durbin-Watson stat	1.683120
Prob(F-statistic)	0.012006		

Bảng bên là
kết quả
ước lượng
mô hình

ỨNG DỤNG DỰ BÁO (tt)

Ta thực hiện một số kiểm định cần thiết đối với mô hình

Ramsey RESET Test

Equation: UNTITLED

Specification: Y C X1 X2

Omitted Variables: Powers of fitted values from 2 to 3

	Value	df	Probability
F-statistic	0.198642	(2, 10)	0.8230
Likelihood ratio	0.584393	2	0.7466

Heteroskedasticity Test: Breusch-Pagan-Godfrey

F-statistic	0.566245	Prob. F(2,12)	0.5821
Obs*R-squared	1.293536	Prob. Chi-Square(2)	0.5237
Scaled explained SS	0.707522	Prob. Chi-Square(2)	0.7020

ỨNG DỤNG DỰ BÁO (tt)

Variance Inflation Factors

Date: 11/25/21 Time: 08:47

Sample: 1 16

Included observations: 15

Variable	Coefficient Variance	Uncentered VIF	Centered VIF
C	13053.95	86.91913	NA
X1	117.3349	35.17065	1.000927
X2	674.3016	55.37427	1.000927

ỨNG DỤNG DỰ BÁO (tt)

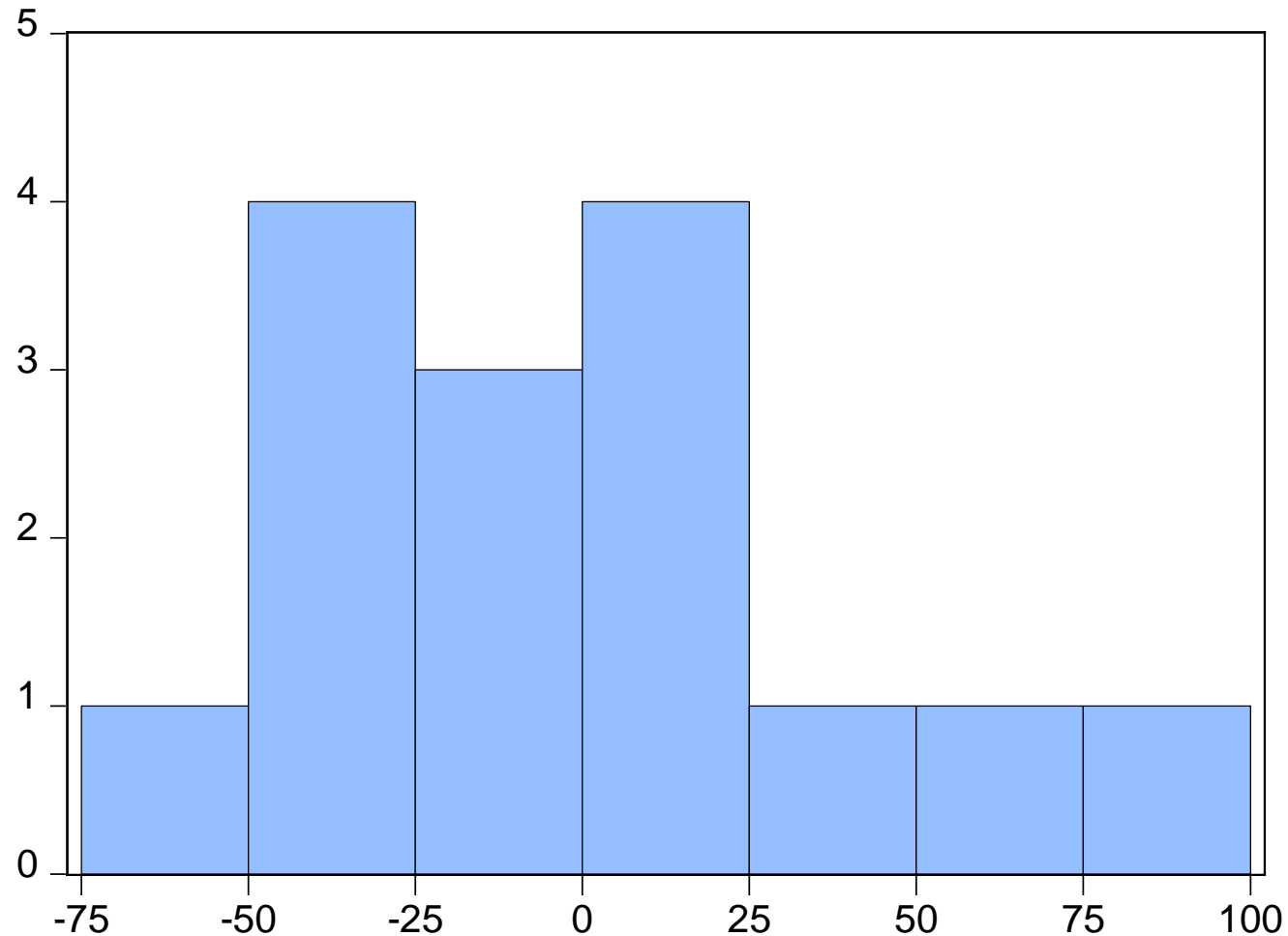
Date: 11/25/21 Time: 08:51

Sample: 1 16

Included observations: 15

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.064	0.064	0.0755	0.783
		2	-0.189	-0.194	0.7740	0.679
		3	-0.281	-0.265	2.4548	0.484
		4	-0.232	-0.270	3.7079	0.447
		5	0.159	0.066	4.3487	0.500
		6	0.199	0.041	5.4756	0.484
		7	0.103	0.028	5.8162	0.561
		8	-0.137	-0.111	6.5007	0.591
		9	-0.183	-0.063	7.9220	0.542
		10	-0.061	-0.040	8.1121	0.618
		11	0.073	-0.017	8.4498	0.673
		12	0.044	-0.124	8.6149	0.735

ỨNG DỤNG DỰ BÁO (tt)



Series: Residuals
Sample 1 15
Observations 15

Mean	1.23e-14
Median	-9.088457
Maximum	96.15482
Minimum	-63.79536
Std. Dev.	43.94258
Skewness	0.594834
Kurtosis	2.709273

Jarque-Bera	0.937396
Probability	0.625817

ỨNG DỤNG DỰ BÁO (tt)

- Khoảng tin cậy 95% của hệ số β_1 : $(-48.58 < \beta_1 < -1.37)$
- Khoảng tin cậy 95% của hệ số β_2 : $(17.55 < \beta_2 < 130.71)$

Như vậy, trong điều kiện các yếu tố khác không đổi:

- Nếu giá bán tăng 1 trăm nghìn đồng thì số lượng sản phẩm A bán sẽ giảm trong khoảng $(1.37; 48.58)$ sản phẩm.
- Nếu chi phí quảng cáo tăng thêm 1 triệu đồng/tuần thì sẽ bán thêm được khoảng 18 đến 131 sản phẩm.

ỨNG DỤNG DỰ BÁO (tt)

Hàm hồi quy mẫu: $\hat{Y} = 306.52 - 24.97X_1 + 74.13X_2$

Nếu bán với giá 8.5 trăm nghìn và chi phí quảng cáo hết 4 triệu đồng/tuần thì:

- Dự báo điểm cho số sản phẩm bán là khoảng 391 sản phẩm.

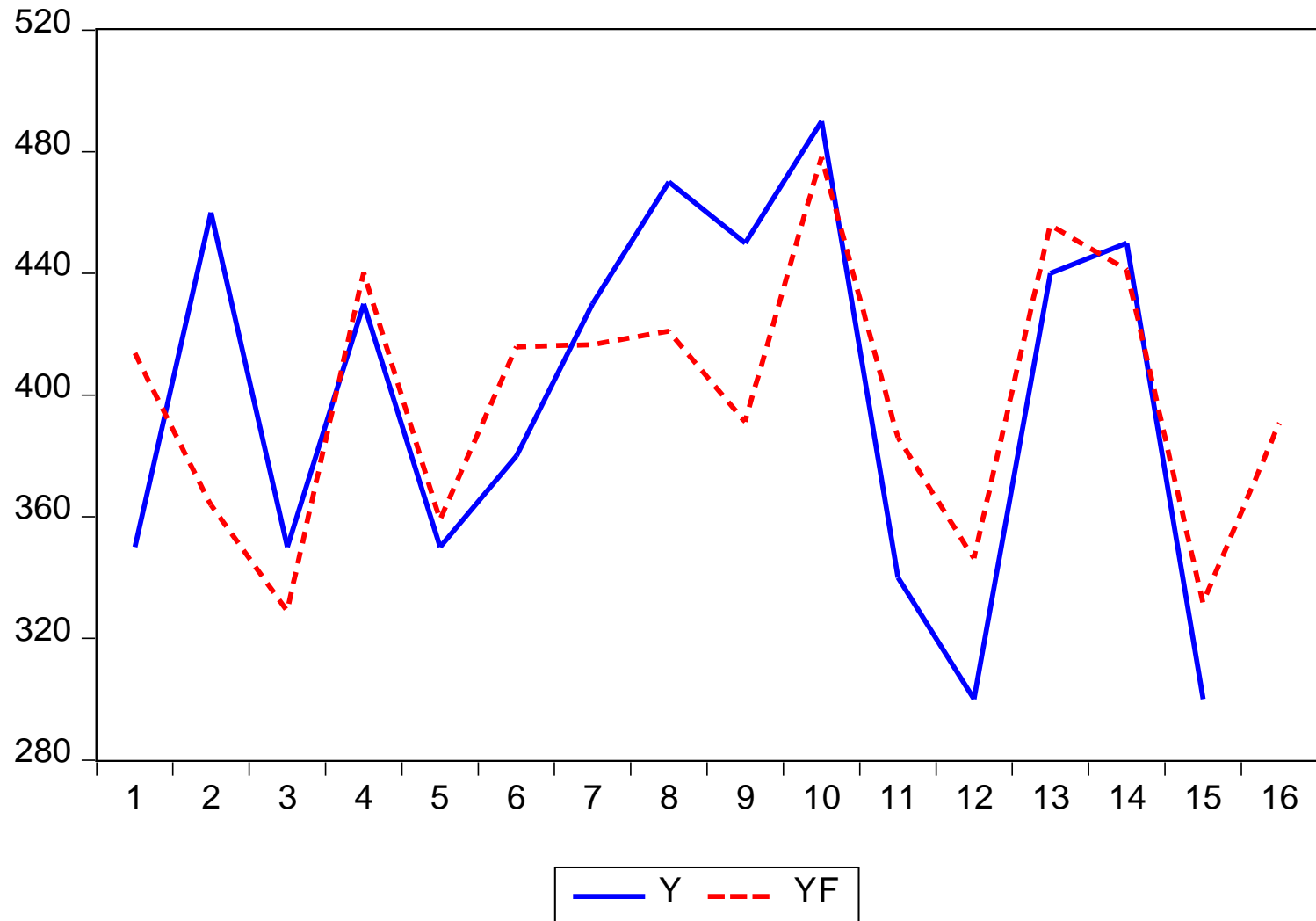
- Dự báo khoảng trung bình cho số sản phẩm bán được là:

$$(333 < E(Y|X_1 = 8.5, X_2 = 4) < 449)$$

- Dự báo khoảng riêng biệt cho số sản phẩm bán được là:

$$(274 < Y_0 < 508)$$

ỨNG DỤNG DỰ BÁO (tt)



Đồ thị đánh giá kết quả dự báo



BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC QUY NHƠN

KẾT THÚC CHƯƠNG 3



BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC QUY NHƠN

CHƯƠNG 4

CÁC MÔ HÌNH DỰ BÁO

THEO PHƯƠNG PHÁP BOX-JENKINS

DỰ BÁO THEO PHƯƠNG PHÁP BOX-JENKINS

1

Goal 1

Hiểu được khái niệm tính dừng và có thể kiểm định tính dừng của chuỗi thời gian

2

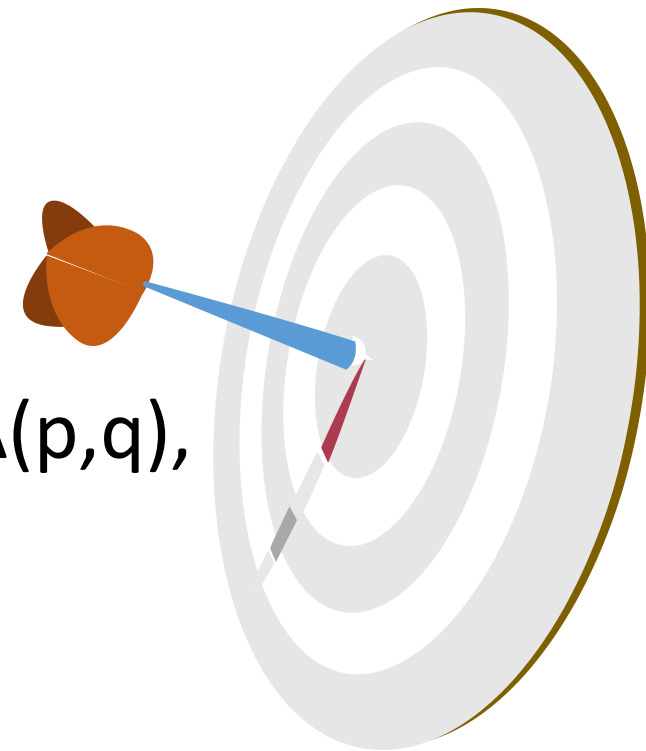
Goal 2

Hiểu rõ về các mô hình: $AR(p)$, $MA(q)$, $ARMA(p,q)$, $ARIMA(p,d,q)$

3

Goal 3

Nắm được quy trình thực hiện dự báo bằng phương pháp Box-Jenkins



DỰ BÁO THEO PHƯƠNG PHÁP BOX - JENKINS

1. Tính dừng của chuỗi thời gian
2. Mô hình tự hồi quy AR
3. Mô hình trung bình động MA
4. Mô hình ARMA(p,q)
5. Tịnh hóa dữ liệu
6. Mô hình ARIMA cho dữ liệu có tính mùa vụ
7. Các bước cơ bản của phương pháp ARIMA

TÍNH DỪNG CỦA CHUỖI THỜI GIAN

Một chuỗi thời gian dừng có các đặc điểm sau đây:

1. Dữ liệu dao động xung quanh một giá trị trung bình cố định trong dài hạn
2. Dữ liệu có giá trị phương sai xác định không thay đổi theo thời gian
3. Dữ liệu có một giản đồ tự tương quan với các hệ số tự tương quan sẽ giảm dần khi độ trễ tăng lên

TÍNH DỪNG CỦA CHUỖI THỜI GIAN (tt)

Theo ngôn ngữ thống kê, các đặc điểm trên được thể hiện bởi: (Y_t)

1. $E(Y_t)$ là một hằng số cho tất cả các thời điểm t

$$E(Y_t) = \mu, \forall t$$

2. $Var(Y_t)$ là một hằng số cho tất cả các thời điểm t

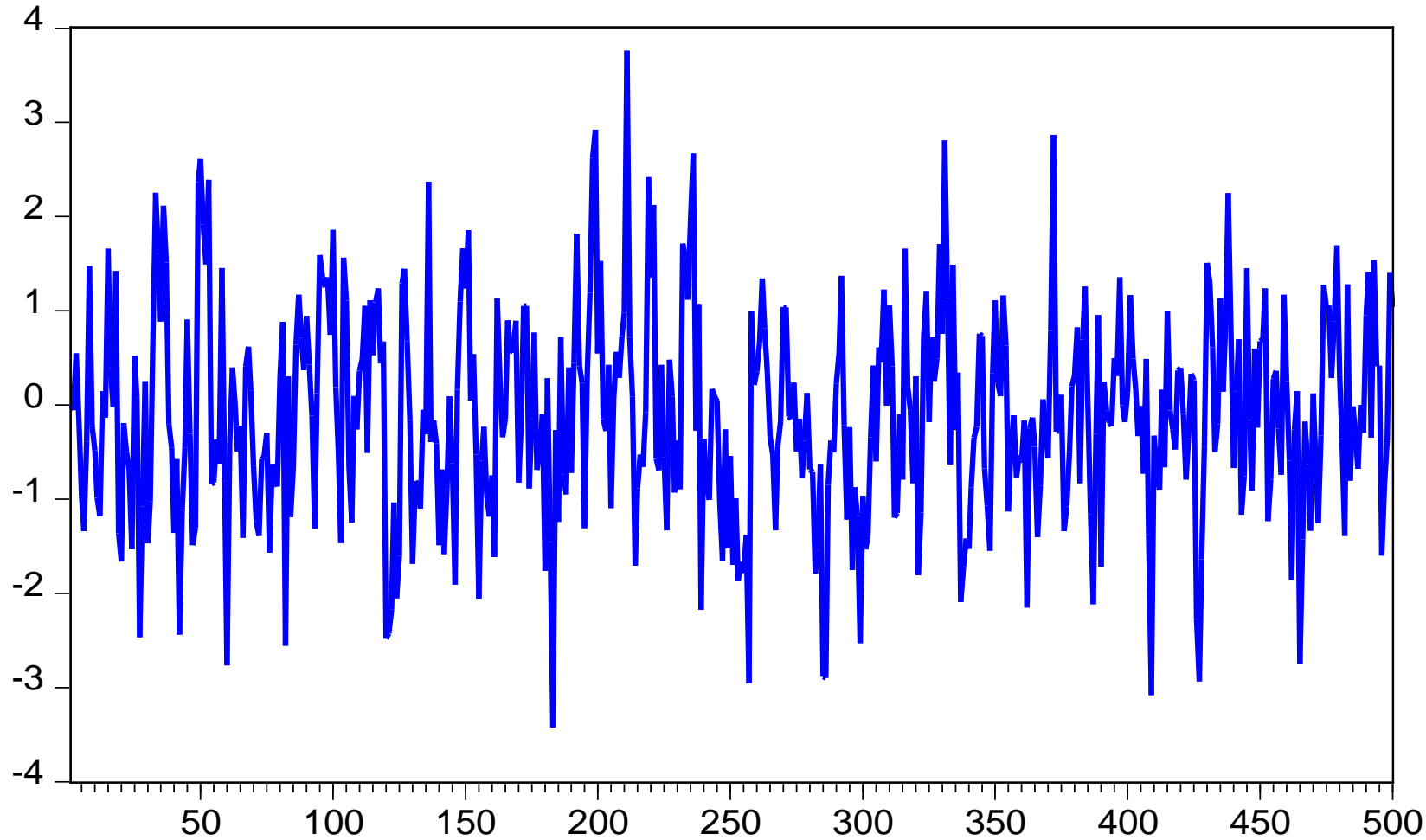
$$Var(Y_t) = E(Y_t - \mu)^2 = \sigma^2, \forall t$$

3. Hiệp phương sai giữa hai giai đoạn chỉ phụ thuộc vào khoảng cách giữa hai giai đoạn

$$cov(Y_t, Y_{t+k}) = E[(Y_t - \mu)(Y_{t+k} - \mu)] = \gamma_k, \forall t$$

TÍNH DỪNG CỦA CHUỖI THỜI GIAN (tt)

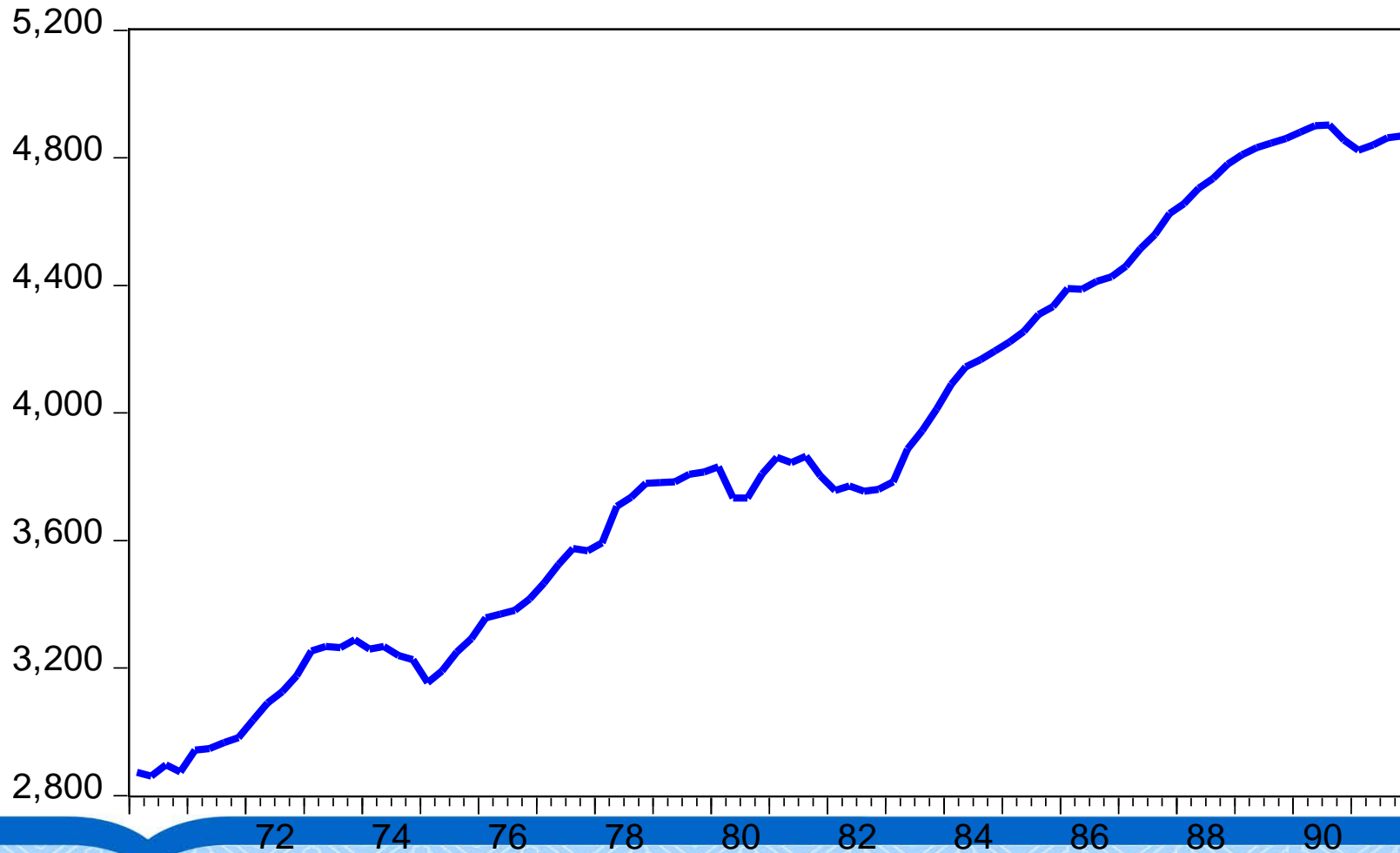
YT



Hình bên là ví dụ minh họa cho một chuỗi dừng với trung bình bằng 0

TÍNH DỪNG CỦA CHUỖI THỜI GIAN (tt)

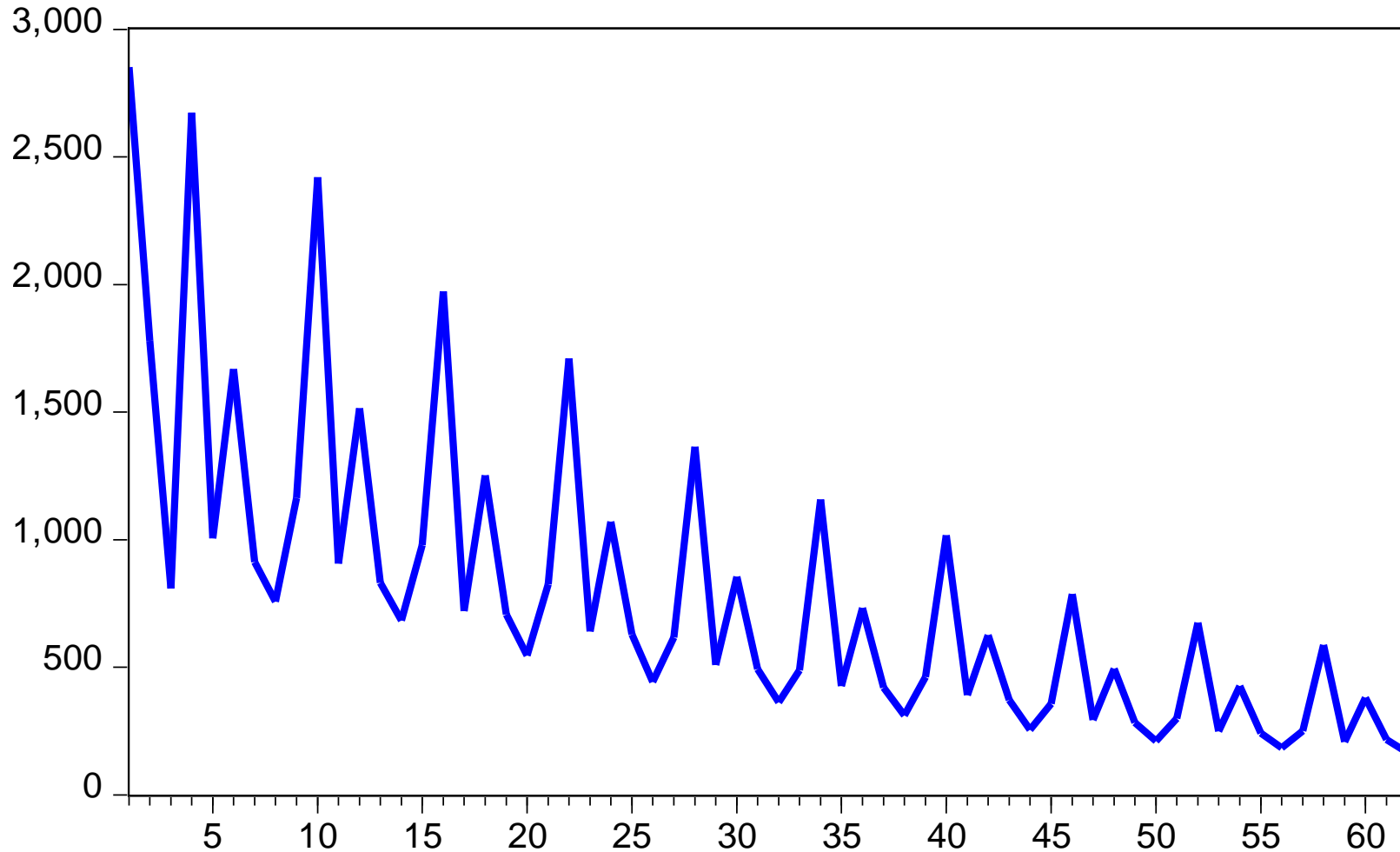
YT



Hình bên là ví dụ minh họa cho một chuỗi không dừng khi trung bình thay đổi

TÍNH DỪNG CỦA CHUỖI THỜI GIAN (tt)

YT



Hình bên là ví dụ minh họa cho một chuỗi không dừng khi cả trung bình và phương sai thay đổi

TÍNH DỪNG CỦA CHUỖI THỜI GIAN (tt)

Kiểm định tính dừng – Giải đồ tự tương quan

Hệ số tự tương quan bậc k được xác định bởi công thức

$$\rho_k = \frac{\sum_{t=k+1}^n (Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t=1}^n (Y_t - \bar{Y})^2} \quad (1)$$

Chia cả tử và mẫu trong (1) cho n ta có

$$\rho_k = \frac{\text{cov}(Y_t, Y_{t-1})}{\text{var}(Y_t)} \quad (2)$$

Các phương trình (1) và (2) được gọi là hàm tự tương quan, ký hiệu AFC.

TÍNH DỪNG CỦA CHUỖI THỜI GIAN (tt)

Kiểm định tính dừng – Giảns đầ tự tương quan (tt)

Thống kê t

Cặp giả thuyết $H_0: \rho_k = 0$ (chuỗi dừng)

$H_1: \rho_k \neq 0$

Với α xác định, ta xây dựng khoảng tin cậy cho ρ_k , trong đó $se(\rho_k) = 1/\sqrt{n}$.

Nếu ρ_k nằm ngoài khoảng tin cậy tìm được thì ta bác bỏ giả thuyết H_0

TÍNH DỪNG CỦA CHUỖI THỜI GIAN (tt)

Kiểm định tính dừng – Giải đồ tự tương quan (tt)

Thống kê Q của Ljung – Box

Cặp giả thuyết $H_0: \rho_k = 0$ (chuỗi dừng)

$$H_1: \rho_k \neq 0$$

Giá trị thống kê Q : $Q = n \sum_{k=1}^m \rho_k^2$

Với cỡ mẫu lớn, Q có phân phối χ^2 với bậc tự do bằng số độ trễ.

Với α cho trước, nếu giá trị Q tính toán lớn hơn giá trị tra tới hạn của χ^2 thì ta bác bỏ giả thuyết H_0 .

TÍNH DỪNG CỦA CHUỖI THỜI GIAN (tt)

Kiểm định tính dừng – Giải đồ tự tương quan (tt)

Date: 11/21/21 Time: 17:02

Sample: 1970Q1 1991Q4

Included observations: 88

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
██████████	██████████	1 0.969	0.969	85.462	0.000
██████████	█	2 0.935	-0.058	166.02	0.000
██████████	█	3 0.901	-0.020	241.72	0.000
██████████	█	4 0.866	-0.045	312.39	0.000
██████████	█	5 0.830	-0.024	378.10	0.000
██████████	█	6 0.791	-0.062	438.57	0.000
██████████	█	7 0.752	-0.029	493.85	0.000
██████████	█	8 0.713	-0.024	544.11	0.000
██████████	█	9 0.675	0.009	589.77	0.000
██████████	█	10 0.638	-0.010	631.12	0.000

Ví dụ về giải
đồ tự tương
quan của một
chuỗi không
dừng





















TÍNH DỪNG CỦA CHUỖI THỜI GIAN (tt)

Kiểm định tính dừng – Giải đồ tự tương quan (tt)

Date: 11/21/21 Time: 16:22

Sample: 1970Q1 1991Q4

Included observations: 87

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.316	0.316	9.0136	0.003
		2	0.186	0.095	12.165	0.002
		3	0.049	-0.038	12.389	0.006
		4	0.051	0.033	12.631	0.013
		5	-0.007	-0.032	12.636	0.027
		6	-0.019	-0.020	12.672	0.049
		7	-0.073	-0.062	13.188	0.068
		8	-0.189	-0.180	21.380	0.006
		9	-0.067	0.128	21.820	0.009
		10	0.019	0.100	21.855	0.016

Ví dụ về
giải đồ tự
tương quan
của một
chuỗi dừng

TÍNH DỪNG CỦA CHUỖI THỜI GIAN (tt)

Kiểm định tính dừng – Kiểm định nghiệm đơn vị

Giả sử có phương trình tự hồi quy sau:

$$Y_t = \rho Y_{t-1} + u_t \quad (-1 \leq \rho \leq 1)$$

hay
$$\Delta Y_t = \delta Y_{t-1} + u_t \quad (\delta = \rho - 1)$$

Ta có các giả thuyết: $H_0: \rho = 1$ (Y_t là chuỗi không dừng)

$$H_0: \rho < 1 \text{ (} Y_t \text{ là chuỗi dừng)}$$

Hay tương đương: $H_0: \delta = 0$ (Y_t là chuỗi không dừng)

$$H_0: \delta < 0 \text{ (} Y_t \text{ là chuỗi dừng)}$$

TÍNH DỪNG CỦA CHUỖI THỜI GIAN (tt)

Kiểm định tính dừng – Kiểm định nghiệm đơn vị (tt)

Theo Dickey và Fuller, giá trị t ước lượng của hệ số Y_{t-1} có phân phối xác suất τ

Giá trị thống kê:
$$\tau_0 = \frac{\hat{\delta}}{se(\hat{\delta})}$$

Nếu $|\tau_0|$ lớn hơn giá trị thống kê τ tra bảng Dickey-Fuller thì ta bác bỏ giả thuyết H_0 , tức là Y_t là một chuỗi dừng

TÍNH DỪNG CỦA CHUỖI THỜI GIAN (tt)

Bảng dưới đây là kết quả kiểm định cho chuỗi có giả đồ tương quan trường hợp không dừng ở trên

Null Hypothesis: GDP has a unit root

Exogenous: Constant, Linear Trend

Lag Length: 1 (Automatic - based on SIC, maxlag=10)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-2.215287	0.4749
Test critical values:		
1% level	-4.068290	
5% level	-3.462912	
10% level	-3.157836	

TÍNH DỪNG CỦA CHUỖI THỜI GIAN (tt)

Bảng dưới đây là kết quả kiểm định cho chuỗi có giả đồ tương quan trường hợp chuỗi dừng ở trên

Null Hypothesis: $D(\text{GDP})$ has a unit root

Exogenous: Constant, Linear Trend

Lag Length: 0 (Automatic - based on SIC, maxlag=10)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-6.588446	0.0000
Test critical values:		
1% level	-4.068290	
5% level	-3.462912	
10% level	-3.157836	

MÔ HÌNH TỰ HỒI QUY - AR

- Mô hình tự hồi quy bậc 1 – AR(1)

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + u_t \quad (-1 < \phi_1 < 1)$$

- Mô hình tự hồi quy bậc 2 – AR(2)

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + u_t$$
$$(-1 < \phi_1, \phi_2 < 1; \phi_1 + \phi_2 < 1)$$

- Mô hình tự hồi quy bậc p – AR(p)

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + u_t$$
$$(\sum_{i=1}^p \phi_i < 1)$$

MÔ HÌNH TỰ HỒI QUY – AR (tt)

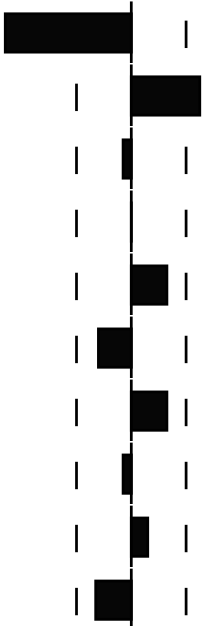
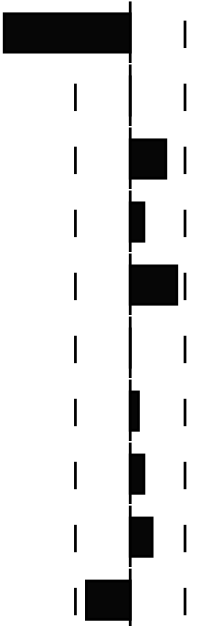
- Xác định độ trễ p dựa trên giản đồ tự tương quan theo cách sau: ACF sẽ có xu hướng bằng 0 ngay lập tức, trong khi đó hệ số tự tương quan riêng phần PACF sẽ có xu hướng khác 0 một cách có ý nghĩa thống kê cho đến độ trễ p và sẽ bằng 0 ngay sau độ trễ p đó.
- $PACF_k$ đo lường mức độ quan hệ giữa Y_t và Y_{t-k} khi các ảnh hưởng của các độ trễ từ 1 đến $k - 1$ đã được loại trừ.

MÔ HÌNH TỰ HỒI QUY – AR (tt)

Date: 11/21/21 Time: 21:11

Sample: 1 75

Included observations: 75

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.528	-0.528	21.796	0.000
		2	0.282	0.003	28.068	0.000
		3	-0.038	0.155	28.186	0.000
		4	0.008	0.065	28.191	0.000
		5	0.144	0.189	29.908	0.000
		6	-0.137	0.002	31.488	0.000
		7	0.147	0.026	33.316	0.000
		8	-0.036	0.060	33.428	0.000
		9	0.068	0.084	33.831	0.000
		10	-0.150	-0.184	35.841	0.000

Giản đồ tự
tương quan
ở bên cho
thấy mô
hình AR(1)
là thích hợp

MÔ HÌNH TỰ HỒI QUY – AR (tt)

Dependent Variable: Y

Method: Least Squares

Date: 11/21/21 Time: 21:14

Sample (adjusted): 2 75

Included observations: 74 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	115.2651	7.570277	15.22601	0.0000
Y(-1)	-0.528921	0.098905	-5.347738	0.0000
R-squared	0.284282	Mean dependent var		75.44595
Adjusted R-squared	0.274342	S.D. dependent var		13.79636
S.E. of regression	11.75251	Akaike info criterion		7.792666
Sum squared resid	9944.745	Schwarz criterion		7.854938
Log likelihood	-286.3286	Hannan-Quinn criter.		7.817507
F-statistic	28.59830	Durbin-Watson stat		1.990986
Prob(F-statistic)	0.000001			

Bảng bên
là kết quả
ước
lượng mô
hình
AR(1)

MÔ HÌNH TRUNG BÌNH DI ĐỘNG - MA

- Mô hình MA(1)

$$Y_t = \mu + u_t + \theta_1 u_{t-1}$$

- Mô hình MA(q)

$$Y_t = \mu + u_t + \sum_{j=1}^q \theta_j u_{t-j}$$

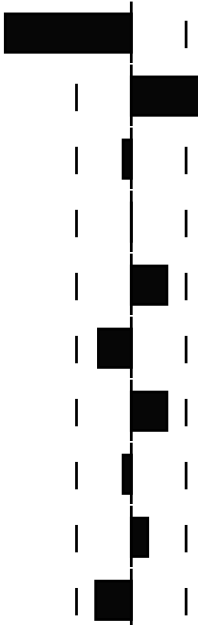
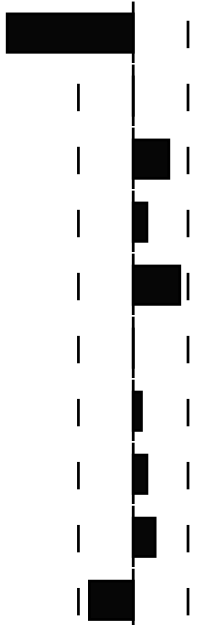
- Mô hình này dự báo Y_t trên cơ sở một kết hợp tuyến tính của các giá trị sai số quá khứ
- Ta xác định độ trễ q từ giản đồ tự tương quan: AFC sẽ có xu hướng khác 0 một cách có ý nghĩa thống kê cho đến độ trễ q và sẽ bằng 0 ngay sau độ trễ q đó. PACF sẽ có xu hướng bằng 0 ngay lập tức

MÔ HÌNH TRUNG BÌNH DI ĐỘNG – MA (tt)

Date: 11/21/21 Time: 21:11

Sample: 1 75

Included observations: 75

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.528	-0.528	21.796	0.000
		2	0.282	0.003	28.068	0.000
		3	-0.038	0.155	28.186	0.000
		4	0.008	0.065	28.191	0.000
		5	0.144	0.189	29.908	0.000
		6	-0.137	0.002	31.488	0.000
		7	0.147	0.026	33.316	0.000
		8	-0.036	0.060	33.428	0.000
		9	0.068	0.084	33.831	0.000
		10	-0.150	-0.184	35.841	0.000

Giả đồ tự
tương quan
ở bên cho
thấy mô
hình MA(2)
là thích hợp

MÔ HÌNH TRUNG BÌNH DI ĐỘNG – MA (tt)

Dependent Variable: Y

Method: ARMA Conditional Least Squares (BFGS / Marquardt steps)

Date: 11/21/21 Time: 22:31

Sample: 1 75

Included observations: 75

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	75.42992	1.062521	70.99144	0.0000
MA(1)	-0.571435	0.109267	-5.229693	0.0000
MA(2)	0.363662	0.113213	3.212189	0.0020
R-squared	0.311809	Mean dependent var		75.24000
Adjusted R-squared	0.292693	S.D. dependent var		13.81841
S.E. of regression	11.62150	Akaike info criterion		7.782768
Sum squared resid	9724.262	Schwarz criterion		7.875468
Log likelihood	-288.8538	Hannan-Quinn criter.		7.819782
F-statistic	16.31106	Durbin-Watson stat		1.939570
Prob(F-statistic)	0.000001			
Inverted MA Roots	.29+.53i	.29-.53i		

Bảng bên là
kết quả ước
lượng mô
hình MA(2)

MÔ HÌNH ARMA

- Kết hợp mô hình AR(p) với mô hình MA(q) ta thu được mô hình ARMA(p,q) như sau:

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2} + \dots + \theta_q u_{t-q}$$

Dạng rút gọn của mô hình ARMA(p,q) như sau:

$$Y_t = \phi_0 + \sum_{i=1}^p \phi_i Y_{t-i} + u_t + \sum_{j=1}^q \theta_j u_{t-j}$$





















Việc xác định các độ trễ p và q thích hợp được thực hiện như ở các phần trên

MÔ HÌNH ARMA (tt)

Date: 11/21/21 Time: 22:39

Sample: 1955Q1 2012Q4

Included observations: 228

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.814	0.814	152.90	0.000
		2	0.634	-0.082	246.24	0.000
		3	0.444	-0.143	292.14	0.000
		4	0.195	-0.311	301.03	0.000
		5	0.108	0.316	303.75	0.000
		6	0.032	-0.053	303.99	0.000
		7	-0.049	-0.130	304.56	0.000
		8	-0.075	-0.100	305.91	0.000
		9	-0.077	0.225	307.35	0.000
		10	-0.080	-0.047	308.90	0.000

Với giản đồ
tương
quan ở
bên, mô
hình phù
hợp là
ARMA(1,3)

MÔ HÌNH ARMA (tt)

Dependent Variable: Y

Date: 11/21/21 Time: 22:46

Sample (adjusted): 1956Q2 2012Q4

Included observations: 227 after adjustments

Convergence achieved after 30 iterations

MA Backcast: 1955Q3 1956Q1

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.024452	0.002886	8.472157	0.0000
AR(1)	0.070404	0.068607	1.026191	0.3059
MA(1)	0.981165	0.017391	56.41717	0.0000
MA(2)	0.982492	0.015598	62.98925	0.0000
MA(3)	0.977187	0.012989	75.23157	0.0000
R-squared	0.806703	Mean dependent var		0.025021
Adjusted R-squared	0.803221	S.D. dependent var		0.023192
S.E. of regression	0.010288	Akaike info criterion		-6.293927
Sum squared resid	0.023496	Schwarz criterion		-6.218488
Log likelihood	719.3608	Hannan-Quinn criter.		-6.263486
F-statistic	231.6235	Durbin-Watson stat		2.037055
Prob(F-statistic)	0.000000			

Bảng bên là
kết quả ước
lượng mô
hình
ARMA(1,3)

MÔ HÌNH ARIMA

Các giai đoạn lựa chọn mô hình ARIMA:

Giai đoạn 1: Nhận dạng

- + Bước 1: Thống kê mô tả để kiểm tra sự bất thường của dữ liệu (outlier, thiếu dữ liệu, thay đổi cấu trúc)
- + Bước 2: Kiểm tra tính dừng của dữ liệu. Nếu không dừng thì lấy sai phân bậc 1.
- + Bước 3: Khi đã có chuỗi dừng thì xác định p và q

Giai đoạn 2: Ước lượng - Ước lượng từng mô hình có thể có và lựa chọn mô hình tốt nhất

Giai đoạn 3: Phân tích chuẩn đoán

- Vẽ đồ thị phần dư theo thời gian hoặc đồ thị tần suất
- Kiểm tra tính ngẫu nhiên của phần dư bằng giản đồ tự tương quan
- Quan sát và so sánh đồ thị giá trị dự báo với giá trị thực tế
- Các kiểm định thống kê khác
- Kiểm tra sai số dự báo

MÔ HÌNH ARIMA (tt)

Quy trình 6 bước của Box - Jenkins:

Bước 1: Kiểm tra tính dừng của chuỗi dữ liệu. Nếu chuỗi dừng, chuyển sang bước 3, nếu không dừng chuyển sang bước 2

Bước 2: Lấy sai phân bậc nhất, tính các hàm ACF và PACF với các chuỗi sai phân bậc nhất

Bước 3: Dựa vào giản đồ ACF và PACF để xác định các mô hình ARMA(p,q) có thể có

Bước 4: Ước lượng các mô hình xác định ở bước 3

Bước 5: Với mỗi mô hình ước lượng ta cần phải:

MÔ HÌNH ARIMA (tt)

Quy trình 6 bước của Box - Jenkins: (tt)

- Kiểm tra hệ số của biến trễ dài nhất có ý nghĩa thống kê hay không. Nếu không có ta nên giảm bậc p và/hoặc q
- Xem xét ACF và PACF của sai số. Nếu mô hình có vừa đủ các tham số thì ACF và PACF của sai số sẽ không có ý nghĩa thống kê
- Kiểm tra các tiêu chí AIC, SBC cùng với R^2 hiệu chỉnh của các mô hình để xem mô hình nào là phù hợp nhất

Bước 6: Nếu mô hình lựa chọn chưa hợp lý và cần có những điều chỉnh thì quay lại bước 4 và thực hiện các bước tiếp theo cho đến khi tìm được mô hình đúng

MÔ HÌNH ARIMA (tt)

Quy trình 6 bước của Box - Jenkins: (tt)

Ngoài ra, chúng ta cần thực hiện các kiểm định sự vi phạm các giả định hồi quy truyền thống đối với các mô hình, bao gồm:

- Kiểm định tự tương quan
- Kiểm định phương sai sai số ngẫu nhiên thay đổi
- Kiểm định thiếu biến/dạng hàm đúng
- Kiểm định sai số ngẫu nhiên có phân phối chuẩn

Các kiểm định trên đều được thực hiện dễ dàng từ phần mềm Eviews

MÔ HÌNH ARIMA (tt)

VÍ DỤ

Sử dụng chuỗi dữ liệu về GDP của một quốc gia theo quý từ 1999Q1 đến 2020Q4. Ta thực hiện các bước của Box – Jenkins để xây dựng mô hình ARIMA phù hợp cho việc dự báo. Ta sử dụng dữ liệu từ 1999Q1 đến 2019Q4 để xây dựng mô hình, dữ liệu còn lại để so sánh kết quả dự báo.

Bước 1: Kiểm tra tính dừng của chuỗi GDP cho thấy GDP là chuỗi không dừng

Các giản đồ tương quan và bảng kiểm định nghiệm đơn vị dưới đây

MÔ HÌNH ARIMA (tt)

Date: 11/23/21 Time: 10:53

Sample: 1999Q1 2020Q4

Included observations: 88

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
██████████	██████████	1 0.969	0.969	85.462	0.000
██████████	█	2 0.935	-0.058	166.02	0.000
██████████	█	3 0.901	-0.020	241.72	0.000
██████████	█	4 0.866	-0.045	312.39	0.000
██████████	█	5 0.830	-0.024	378.10	0.000
██████████	█	6 0.791	-0.062	438.57	0.000
██████████	█	7 0.752	-0.029	493.85	0.000
██████████	█	8 0.713	-0.024	544.11	0.000
██████████	█	9 0.675	0.009	589.77	0.000
██████████	█	10 0.638	-0.010	631.12	0.000

Hình bên là
giản đồ
tương quan
của chuỗi gốc
GDP

MÔ HÌNH ARIMA (tt)

Kết quả KĐ tính dừng của chuỗi gốc GDP bằng KĐ nghiệm đơn vị

Null Hypothesis: GDP has a unit root

Exogenous: Constant, Linear Trend

Lag Length: 1 (Automatic - based on SIC, maxlag=10)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-2.215287	0.4749
Test critical values:		
1% level	-4.068290	
5% level	-3.462912	
10% level	-3.157836	

*MacKinnon (1996) one-sided p-values.

MÔ HÌNH ARIMA (tt)

Bước 2: Tạo biến sai phân bậc 1 của chuỗi GDP (dGDP). Kết quả kiểm định cho thấy dGDP là chuỗi dừng

Các giản đồ tương quan và bảng kiểm định nghiệm đơn vị dưới đây

Bước 3: Xác định các độ trễ p và q cho mô hình ARMA với chuỗi dGDP. Kết quả cho thấy các độ trễ khả dĩ của p và q đều là 1, 8 và 12. Như vậy ta sẽ có các mô hình ARIMA($p,1,q$), trong đó

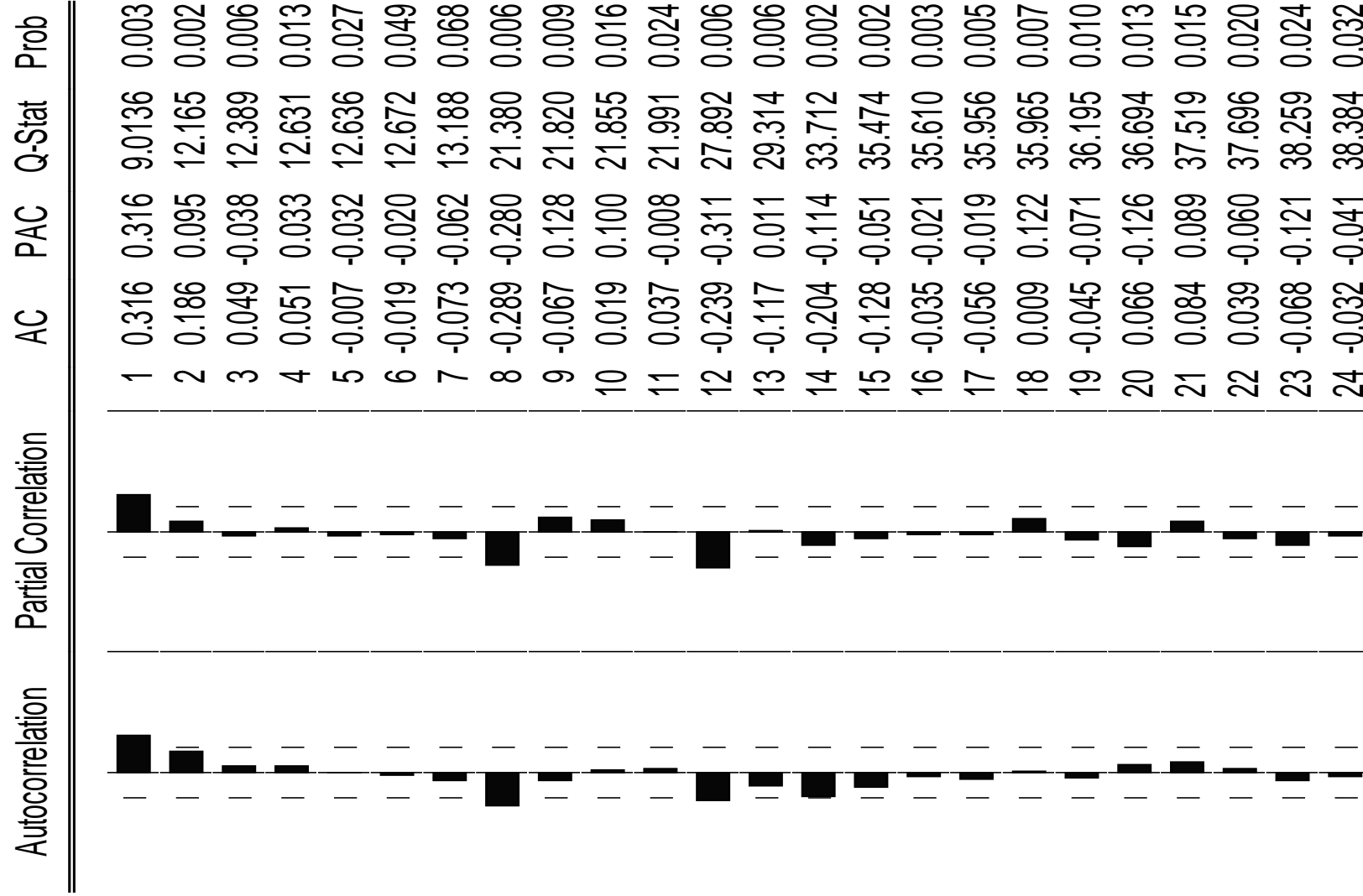
$$p \in \{1,8,12\}, q \in \{1,8,12\}$$

Date: 11/23/21 Time: 15:09

Sample: 1999Q1 2020Q4

Included observations: 87

MÔ HÌNH ARIMA (tt)



Hình bên là
giản đồ
tương quan
của chuỗi sai
phân bậc
nhất của
chuỗi GDP
(dGDP)

MÔ HÌNH ARIMA (tt)

Kết quả KĐ tính dừng của chuỗi dGDP bằng KĐ nghiệm đơn vị

Null Hypothesis: DGDP has a unit root

Exogenous: Constant, Linear Trend

Lag Length: 0 (Automatic - based on SIC, maxlag=24)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-6.588446	0.0000
Test critical values:		
1% level	-4.068290	
5% level	-3.462912	
10% level	-3.157836	

*MacKinnon (1996) one-sided p-values.

MÔ HÌNH ARIMA (tt)

Bước 4: Thực hiện ước lượng các mô hình ở bước 3

Bước 5: Lựa chọn mô hình tốt nhất từ các mô hình đã thực hiện ở bước 4.

Sử dụng các tiêu chuẩn AIC, SBC, HQC so sánh giữa các mô hình kết hợp kiểm định hệ số hồi quy khác không và các kiểm định mô hình truyền thống như dạng hàm đúng, sai số ngẫu nhiên phân phối chuẩn, tự tương quan, phương sai thay đổi.

Kết quả cho thấy mô hình với AR(1) MA(8) MA(12) là phù hợp nhất. Kết quả ước lượng mô hình này như sau:

MÔ HÌNH ARIMA (tt)

Dependent Variable: D(GDP)

Method: ARMA Maximum Likelihood (OPG - BHHH)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	23.75756	2.224168	10.68155	0.0000
AR(1)	0.362748	0.111834	3.243624	0.0017
MA(8)	-0.404994	0.101349	-3.996046	0.0001
MA(12)	-0.362782	0.135681	-2.673782	0.0091
SIGMASQ	887.2184	143.5167	6.181985	0.0000
R-squared	0.316292	Mean dependent var		23.88313
Adjusted R-squared	0.281230	S.D. dependent var		36.24199
S.E. of regression	30.72607	Akaike info criterion		9.796948
Sum squared resid	73639.12	Schwarz criterion		9.942661
Log likelihood	-401.5733	Hannan-Quinn criter.		9.855488
F-statistic	9.020947	Durbin-Watson stat		1.996538
Prob(F-statistic)	0.000005			

MÔ HÌNH ARIMA (tt)

Một số kiểm định mô hình

Ramsey RESET Test

Equation: ARIMA1_812

Specification: D(GDP) C AR(1) MA(8) MA(12)

Omitted Variables: Powers of fitted values from 2 to 4

	Value	df	Probability
F-statistic	0.850700	(3, 75)	0.4706
Likelihood ratio	5.202316	3	0.1576

Heteroskedasticity Test: ARCH

F-statistic	0.748934	Prob. F(4,74)	0.5619
Obs*R-squared	3.073717	Prob. Chi-Square(4)	0.5456

MÔ HÌNH ARIMA (tt)

Date: 11/23/21 Time: 16:49

Sample: 1999Q1 2019Q4

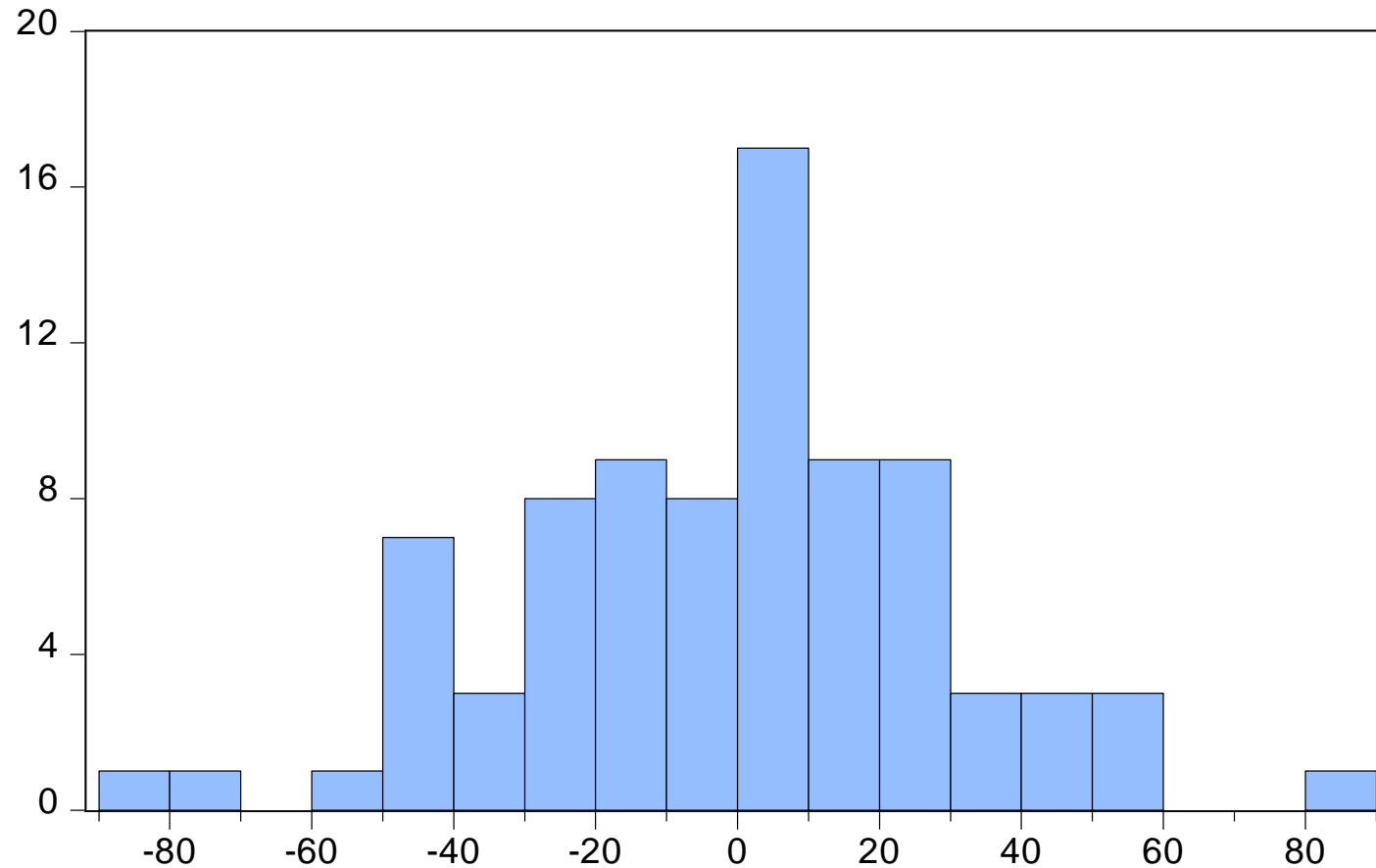
Included observations: 83

Q-statistic probabilities adjusted for 3 ARMA terms

	Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
1	1	1	-0.021	-0.021	0.0368	
2	0.093	0.093	0.093	0.093	0.7910	
3	-0.016	-0.012	-0.016	-0.012	0.8127	
4	-0.170	-0.181	-0.170	-0.181	3.3887	0.066
5	-0.031	-0.037	-0.031	-0.037	3.4772	0.176
6	-0.047	-0.014	-0.047	-0.014	3.6829	0.298
7	-0.041	-0.043	-0.041	-0.043	3.8409	0.428
8	-0.043	-0.073	-0.043	-0.073	4.0117	0.548
9	0.021	0.014	0.021	0.014	4.0522	0.670
10	0.134	0.140	0.134	0.140	5.7818	0.565
11	0.140	0.134	0.140	0.134	7.6928	0.464
12	-0.056	-0.107	-0.056	-0.107	8.0050	0.534
13	0.044	0.014	0.044	0.014	8.1960	0.610
14	-0.197	-0.140	-0.197	-0.140	12.167	0.351
15	-0.132	-0.112	-0.132	-0.112	13.964	0.303
16	-0.047	-0.043	-0.047	-0.043	14.200	0.360
17	-0.096	-0.063	-0.096	-0.063	15.190	0.365
18	0.044	0.016	0.044	0.016	15.398	0.423
19	-0.022	-0.046	-0.022	-0.046	15.453	0.492
20	0.066	0.007	0.066	0.007	15.944	0.528
21	0.105	0.057	0.105	0.057	17.198	0.510
22	0.009	-0.019	0.009	-0.019	17.207	0.576
23	-0.035	-0.078	-0.035	-0.078	17.351	0.630
24	-0.055	-0.033	-0.055	-0.033	17.718	0.667

MÔ HÌNH ARIMA (tt)

Kiểm định phần dư có phân phối chuẩn



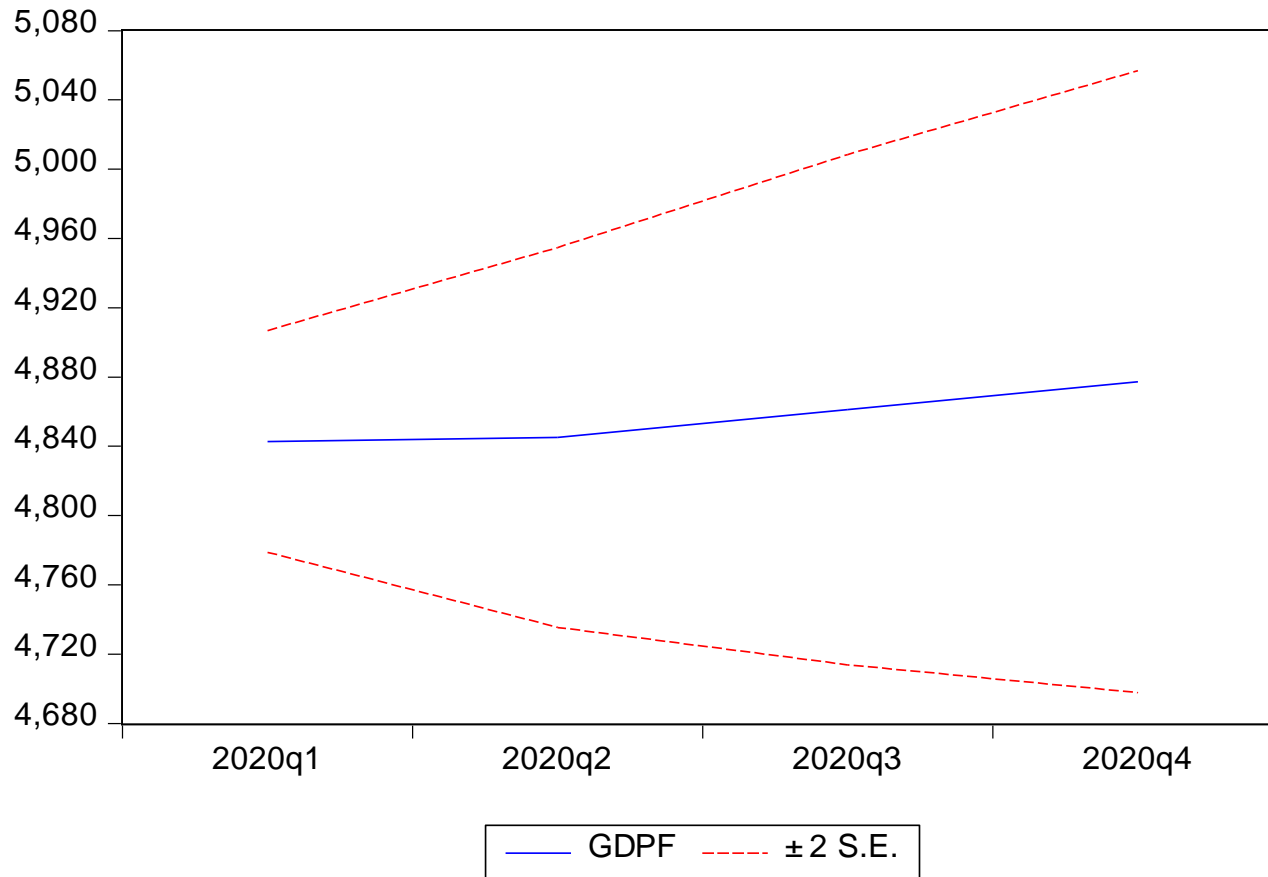
Series: Residuals
Sample 1999Q2 2019Q4
Observations 83

Mean	-0.674460
Median	1.725865
Maximum	85.08337
Minimum	-87.08154
Std. Dev.	29.95960
Skewness	-0.124600
Kurtosis	3.458114

Jarque-Bera	0.940562
Probability	0.624827

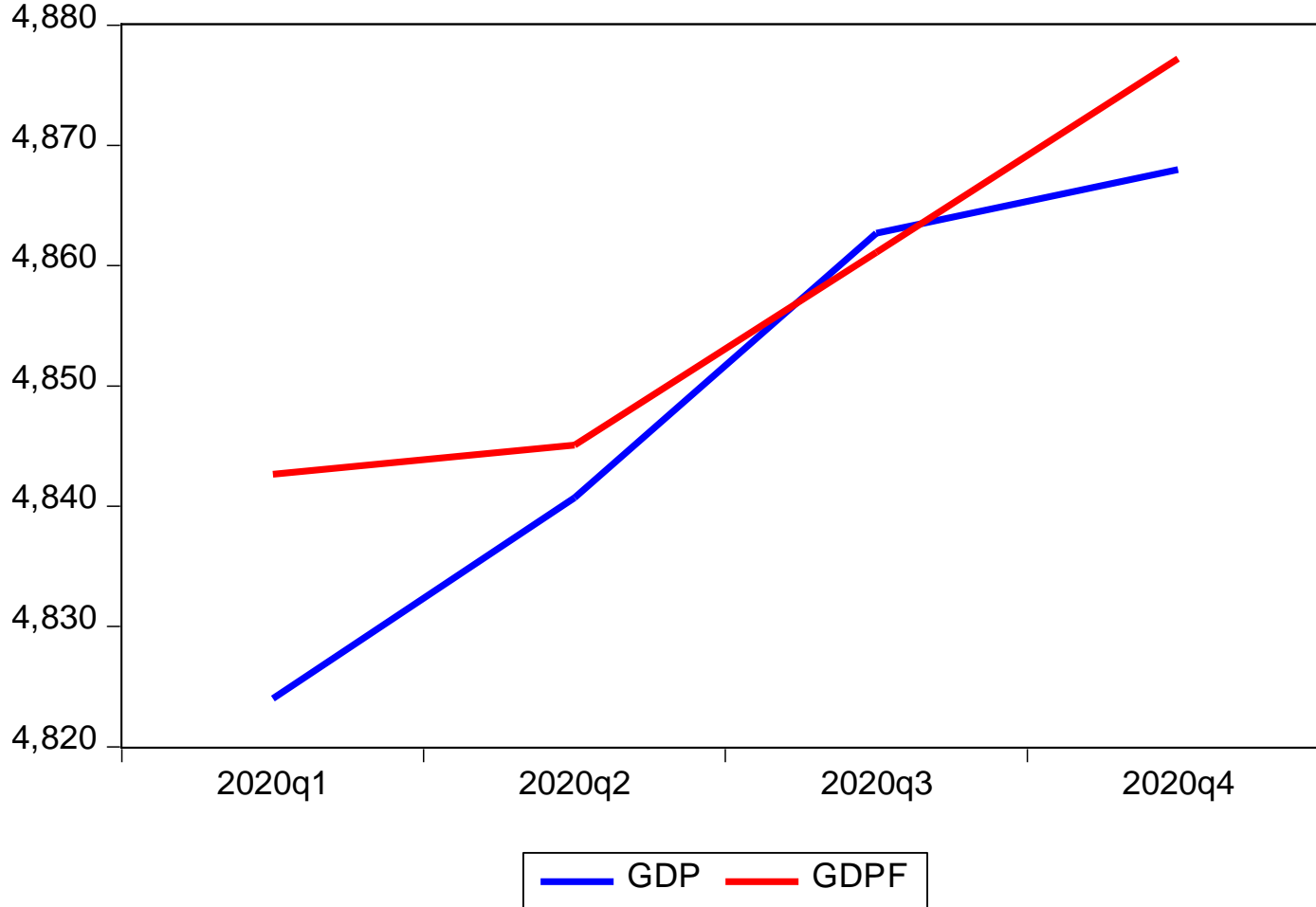
MÔ HÌNH ARIMA (tt)

Bước 6: Thực hiện dự báo cho 4 thời kỳ 2020Q1 – 2020Q4



Forecast: GDPF	
Actual: GDP	
Forecast sample: 2020Q1 2020Q4	
Included observations: 4	
Root Mean Squared Error	10.65296
Mean Absolute Error	8.449882
Mean Abs. Percent Error	0.174589
Theil Inequality Coefficient	0.001098
Bias Proportion	0.516430
Variance Proportion	0.122697
Covariance Proportion	0.360873
Theil U2 Coefficient	0.364815
Symmetric MAPE	0.174349

MÔ HÌNH ARIMA (tt)



Time	GDP	GDPF
2020Q1	4824.0	4842.6
2020Q2	4840.7	4845.1
2020Q3	4862.7	4861.1
2020Q4	4868.0	4877.2

There cannot be an exact or perfect ARIMA model because it is more “of an art than of science”



BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC QUY NHƠN

KẾT THÚC CHƯƠNG 4

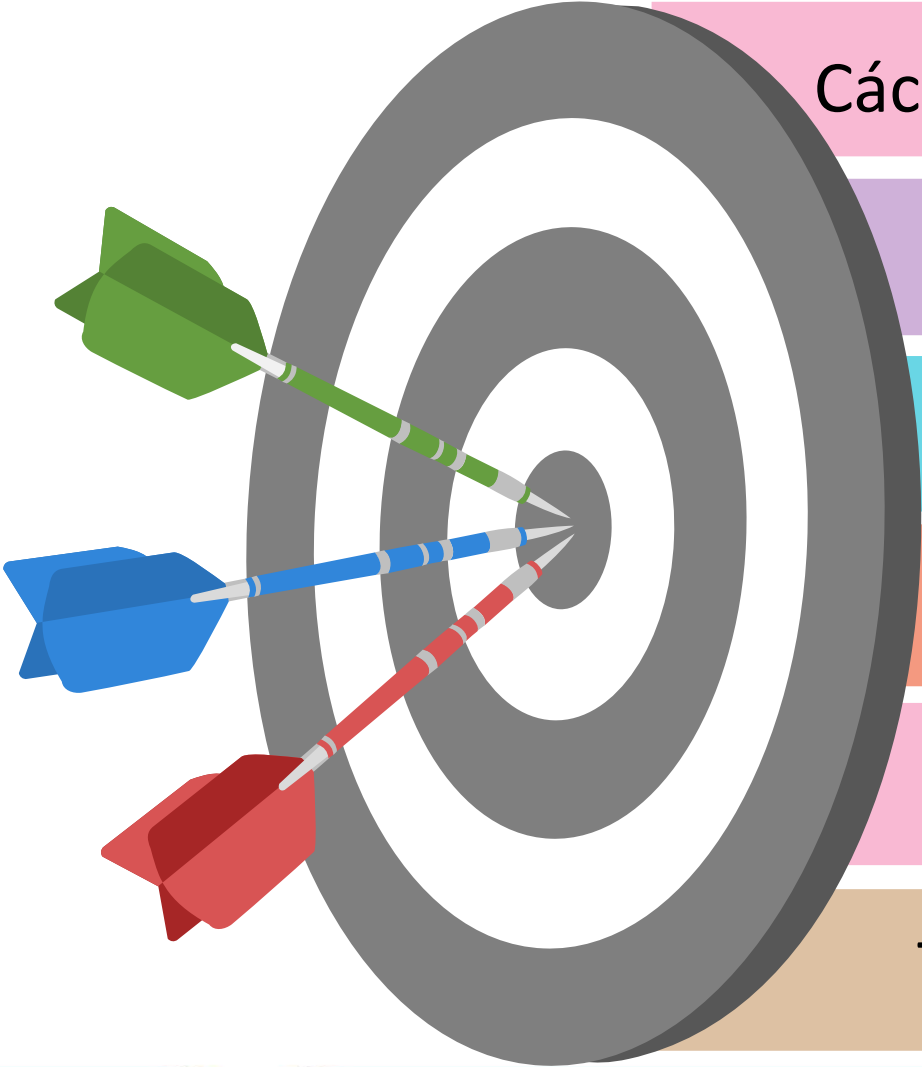


BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC QUY NHƠN

CHƯƠNG 5

KIỂM SOÁT VÀ QUẢN LÝ

QUY TRÌNH DỰ BÁO



Các nhân tố then chốt quyết định kết quả dự báo

Quy trình dự báo chuẩn mực

Lựa chọn được các PP dự báo thích hợp

Cách thức giám sát tốt quy trình dự báo

Xây dựng khung quy trình dự báo

Trách nhiệm thực hiện dự báo

NHÂN TỐ QUYẾT ĐỊNH KẾT QUẢ DỰ BÁO

Nhân tố 1: Xác định đối tượng dự báo

- Dự báo nên được hiểu là một quy trình quản trị thay vì là một lập trình máy tính
- Cần phân biệt điều gì cần và điều gì không cần dự báo

Nhân tố 2: Dự báo sự khác biệt giữa nhu cầu và khả năng

- Dự báo dựa trên khả năng của công ty có thể cung cấp hàng hóa và dịch vụ hơn là nhu cầu thực tế của khách hàng?
- Các công ty nên kết hợp chặt chẽ giữa dự báo khả năng cung cấp và dự báo nhu cầu để có quyết định tốt nhất cho tương lai

NHÂN TỐ QUYẾT ĐỊNH KẾT QUẢ DỰ BÁO

Nhân tố 3: Dự báo cần trao đổi, hợp tác và cộng tác

- Tương tác thông tin từ các thành viên khác nhau trong các bộ phận chức năng khác nhau sẽ góp phần cải thiện đáng kể kết quả dự báo chung của đơn vị thực hiện dự báo.
- Cần tạo được niềm tin giữa các bộ phận chức năng và thiết lập một cơ chế để kết nối họ trong thực hiện dự báo

Nhân tố 4: Dự báo cần loại bỏ những “ốc đảo”

- Tránh trường hợp mỗi bộ phận trong cùng đơn vị tự có xu hướng thực hiện các dự báo cho riêng mình

NHÂN TỐ QUYẾT ĐỊNH KẾT QUẢ DỰ BÁO

- Các dự báo độc lập trong cùng đơn vị có thể dẫn đến kết quả mâu thuẫn với nhau và không thể sử dụng kết quả dự báo
- Cần thiết lập một quy trình dự báo duy nhất dưới sự hỗ trợ của “hạ tầng dự báo” và cơ sở dữ liệu thống nhất

Nhân tố 5: Sử dụng các phương pháp dự báo hiệu quả

- Các phương pháp phải được hiểu và sử dụng một cách sáng suốt phù hợp với mỗi điều kiện môi trường của đơn vị thực hiện dự báo
- Có thể kết hợp cả phương pháp dự báo định tính và định lượng

NHÂN TỐ QUYẾT ĐỊNH KẾT QUẢ DỰ BÁO

- Có thể sử dụng ý kiến định tính từ các bộ phận chức năng để điều chỉnh kết quả dự báo định lượng ban đầu

Nhân tố 6: Làm cho dự báo trở nên quan trọng

- Mỗi doanh nghiệp cần phải có chính sách thể hiện dự báo thực sự quan trọng cho thành công của doanh nghiệp

- Cần phải làm cho cả người sử dụng lẫn người thực hiện quen với toàn bộ quy trình dự báo

- Doanh nghiệp cần phải thiết lập một khung quản lý quy trình dự báo thống nhất

NHÂN TỐ QUYẾT ĐỊNH KẾT QUẢ DỰ BÁO (tt)

Nhân tố 7: Quy trình và kết quả dự báo cần đo lường và đánh giá

- Cần có hệ thống các tiêu chí và thang đo thành quả của dự báo để có thể đánh giá mức độ hoàn thành công việc của mỗi cá nhân.
- Việc đánh giá kết quả dự báo giúp đơn vị thực hiện dự báo biết được lý do tại sao một dự báo thành công hoặc thất bại để có những biện pháp cải thiện quy trình dự báo.

ĐÁNH GIÁ LẠI QUY TRÌNH DỰ BÁO

- Trong suốt quy trình dự báo đòi hỏi phải có sự trao đổi qua lại liên tục giữa người làm dự báo và những người sử dụng kết quả dự báo.
- Quy trình dự báo luôn luôn phải được quản lý và giám sát chặt chẽ.
- Trong suốt quy trình thực hiện dự báo cần phải kết hợp hài hòa giữa phương pháp dự báo định lượng và các phương pháp dự báo định tính.
- Ý kiến chuyên gia luôn cần thiết trong việc xem xét và lựa chọn các phương pháp dự báo thích hợp.

ĐÁNH GIÁ LẠI QUY TRÌNH DỰ BÁO (tt)

- Một số câu hỏi điển hình để kiểm soát xem việc quản lý quy trình dự báo có được thực hiện tốt không:

1. Tại sao cần dự báo
2. Ai sử dụng kết quả dự báo và cụ thể là họ cần gì?
3. Dự báo chi tiết hay tổng hợp đến mức nào và độ dài thời gian dự báo là bao nhiêu?
4. Dữ liệu sẵn có là gì và dữ liệu đó có đủ để thực hiện dự báo mong muốn hay không?
5. Chi phí dự báo là bao nhiêu?

ĐÁNH GIÁ LẠI QUY TRÌNH DỰ BÁO (tt)

7. Mức độ chính xác mong muốn dự báo là bao nhiêu?
8. Kết quả dự báo có kịp cho quá trình ra quyết định hay không?
9. Người làm dự báo có hiểu rõ kết quả dự báo sẽ được sử dụng như thế nào trong tổ chức hay không?
10. Đã có sẵn quy trình phản hồi để đánh giá dự báo sau khi được thực hiện và để điều chỉnh quy trình dự báo thích hợp hơn hay chưa?

LỰA CHỌN CÁC PHƯƠNG PHÁP DỰ BÁO THÍCH HỢP

Phương pháp dự báo	Dạng dữ liệu	Lượng dữ liệu	Độ dài dự báo
<i>Phương pháp định tính</i>			
- Tổng hợp lực lượng bán hàng	Bất kỳ	Ít	Ngắn, trung hạn
- Khảo sát khách hàng	Bất kỳ	Không	Trung, dài hạn
- Ý kiến ban quản lý	Bất kỳ	Ít	Bất kỳ
- Ý kiến chuyên gia	Bất kỳ	Ít	Dài hạn
<i>Dự báo thô</i>	Dừng*	1 hoặc 2	Rất ngắn
<i>Bình quân di động</i>	Dừng*	Ít nhất bằng hệ số trượt	Rất ngắn

LỰA CHỌN CÁC PHƯƠNG PHÁP DỰ BÁO THÍCH HỢP (tt)

Phương pháp dự báo	Dạng dữ liệu	Lượng dữ liệu	Độ dài dự báo
<p>San mũ</p> <ul style="list-style-type: none"> - Giản đơn - Holt - Winter 	<p>Dừng*</p> <p>Xu thế tuyến tính</p> <p>Xu thế mùa vụ</p>	<p>5-10</p> <p>10-15</p> <p>Ít nhất là 4-5 quan sát/mùa vụ</p>	<p>Ngắn hạn</p> <p>Ngắn, trung hạn</p> <p>Ngắn, trung hạn</p>
<p>Hồi quy</p> <ul style="list-style-type: none"> - Xu thế 	<p>Tuyến tính/phi tuyến, có hoặc không có yếu tố mùa</p>	<p>Tối thiểu 10/4 hoặc 5 quan sát/mùa nếu có yếu tố mùa</p>	<p>Ngắn, trung hạn</p>

LỰA CHỌN CÁC PHƯƠNG PHÁP DỰ BÁO THÍCH HỢP (tt)

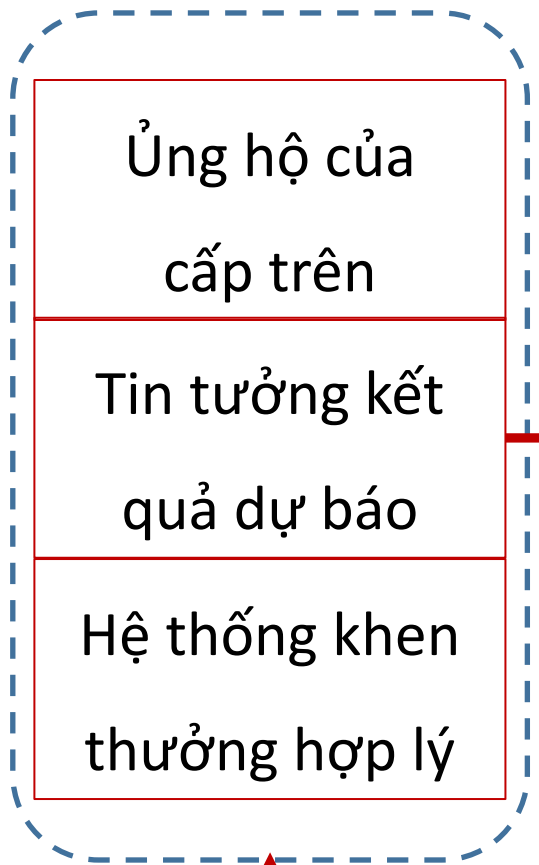
Phương pháp dự báo	Dạng dữ liệu	Lượng dữ liệu	Độ dài dự báo
Hồi quy - Nhân quả	Tất cả các dạng dữ liệu	Tối thiểu quan sát/biến giải thích	Ngắn, trung và dài hạn
Phân tích chuỗi thời gian	Xu thế, mùa vụ và chu kỳ	Đủ lớn để có thể phát hiện đỉnh và đáy trong chu kỳ	Ngắn, trung và dài hạn
ARIMA/SARIMA	Dừng*	Tối thiểu 50	Ngắn, trung và dài hạn

* Kể cả các chuỗi sau khi đã biến đổi thành chuỗi dừng

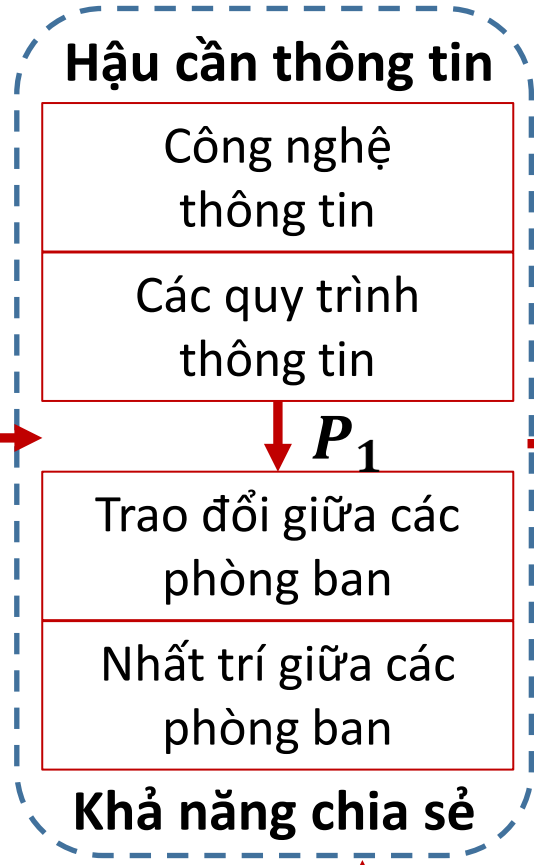
Nguồn: Holton Wilson & Barry Keating, 2007, tr.436

XÂY DỰNG KHUNG QUẢN LÝ QUY TRÌNH DỰ BÁO

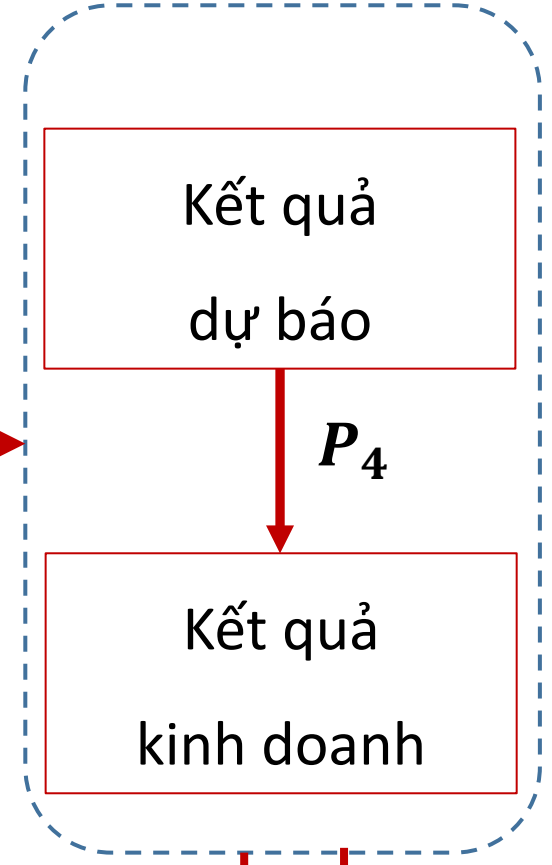
Môi trường dự báo



Năng lực dự báo



Kết quả dự báo



P_2

P_1

P_3

P_4

P_6

P_5

Đo lường dự báo

XÂY DỰNG KHUNG QUẢN LÝ QUY TRÌNH DỰ BÁO

P_1 : Hậu cần thông tin có mối quan hệ tích cực với khả năng chia sẻ giữa các bộ phận chức năng

P_2 : Môi trường dự báo càng tích cực thì năng lực dự báo của doanh nghiệp càng được phát huy

P_3 : Năng lực dự báo có ảnh hưởng tích cực đến kết quả dự báo

P_4 : Mức độ chính xác của dự báo có mối quan hệ tích cực với kết quả hoạt động của doanh nghiệp

P_5 & P_6 : Kết quả dự báo (và kết quả hoạt động kinh doanh) có ảnh hưởng tích cực lên năng lực dự báo và môi trường dự báo

GIÁM SÁT KẾT QUẢ DỰ BÁO

- Các dữ liệu cũ nên được loại ra và cập nhật dữ liệu gần đây vào cơ sở dữ liệu nhằm phục vụ cho dự báo.
- Khi dữ liệu cập nhật, các tham số được sử dụng trong mô hình dự báo cần được tính toán lại, hay chính là cập nhật mô hình dự báo khi cơ sở dữ liệu thay đổi.
- Mô hình dự báo với các tham số mới được xem xét để kiểm tra độ chính xác. Nếu độ chính xác thích hợp thì mô hình tiếp tục được sử dụng cho đến lần cập nhật dữ liệu mới. Ngược lại, nên xem xét lại dạng dữ liệu và chọn mô hình khác thích hợp hơn.

MỘT SỐ LƯU Ý KHÁC CHO QUY TRÌNH DỰ BÁO

- Trách nhiệm dự báo
- Chi phí dự báo
- Phát triển dự báo định lượng
- Không quên vai trò của dự báo định tính



BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC QUY NHƠN

KẾT THÚC CHƯƠNG 5