เฉลย
การพยากรณ์สถิติ

เฉลยบทที่ 1

1. จงให้ความหมายของการพยากรณ์

ตอบ การประมาณค่า ณ เวลาใดเวลาหนึ่งในอนาคต

2. การพยากรณ์เชิงปริมาณคืออะไร

ตอบ การพยากรณ์ซึ่งข้อมูลในอดีตจะถูกนำมาใช้ในการสร้างตัวแบบ (Model) เพื่อที่จะพยากรณ์ค่าในอนาคต

3. การพยากรณ์เชิงคุณภาพคืออะไร จงอธิบาย

ตอบ การพยากรณ์ซึ่งเน้นหนักไปที่ข้อมูลนำเข้าซึ่งมีลักษณะเป็นข้อมูลเชิงคุณภาพหรือเชิงอัตวิสัย อันได้แก่ ความคิดเห็น การตัดสินใจ หรือความคิดเห็นของผู้เชี่ยวชาญเพื่อพยากรณ์ค่าในอนาคต

4. จงอธิบายความแตกต่างของการพยากรณ์เชิงคุณภาพและเชิงปริมาณ

ตอบ สิ่งที่ใช้เป็นพื้นฐานในการพยากรณ์

5. การพยากรณ์โดยใช้เทคนิค Delphi มีวัตถุประสงค์เพื่อปิดช่องโหว่ของการพยากรณ์เชิงอัตวิสัยในแง่ใด จงอธิบาย

ตอบ การพยากรณ์เชิงคุณภาพนั้นจะขึ้นอยู่กับความเห็นของกลุ่มคณะบุคคล (Panel) เป็นหลัก ซึ่งอาจทำให้เกิดอคติขึ้นมาได้จากผู้เชี่ยวชาญที่มีอิทธิพล เทคนิค Delphi จะเน้นที่ผู้เชี่ยวชาญแต่ละท่านสามารถแสดงความคิดเห็นได้อย่างเป็นอิสระซึ่งกันและกัน อีกประการหนึ่งก็คือ ผู้เชี่ยวชาญแต่ละคนจะต้องผ่านสอบถามมากกว่า 1 ครั้ง เพื่อให้ได้คำตอบที่มีความแม่นยำและหนักแน่นที่สุด

6. เกณฑ์ที่ใช้ในการเลือกเทคนิคในการพยากรณ์มีอะไรบ้าง จงอธิบาย

ตอบ

- ปัจจัยในเรื่องของเวลา ได้แก่ ระยะเวลาในการพยากรณ์ในอนาคต เช่น ระยะสั้น หรือ ระยะยาว
- งบประมาณที่มี
- ระดับของความแม่นยำที่คาดหวัง
- ข้อมูลพื้นฐานที่มีอยู่
- ความซับซ้อนของตัวแบบพยากรณ์

7. ขั้นตอนในการพยากรณ์มีอะไรบ้าง จงอธิบาย

ตอบ

- ทำการวิเคราะห์ว่าอะไรคือสิ่งที่ต้องการทำการพยากรณ์ โดยเริ่มต้นจากการระบุปัญหาที่ชัดเจน
- ทำการระบุตัวแปรอิสระ
- ทำการวิเคราะห์เพื่อหาค่าระยะเวลาในอนาคตที่ต้องการจะพยากรณ์ ว่าเป็นระยะสั้นหรือ ระยะยาว
- เลือกหลักการพยากรณ์ว่าจะเป็นการพยากรณ์เชิงปริมาณหรือการพยากรณ์เชิงคุณภาพ แล้วจึงเลือกตัวแบบพยากรณ์ที่เหมาะสมต่อไป
- ทำการเก็บข้อมูลและทำการพยากรณ์
- ทำการประเมินผลจากการพยากรณ์ในแง่ของความแม่นยำ
- เลือกตัวแบบในการพยากรณ์ใหม่เพื่อปรับปรุงค่าความแม่นยำในการพยากรณ์

8. ค่าความคลาดเคลื่อนจากการพยากรณ์คืออะไร และมีความสัมพันธ์อย่างไรกับการเลือกเทคนิคในการพยากรณ์

ตอบ หากเลือกตัวแบบที่เหมาะสมในการพยากรณ์แล้ว ค่าความคลาดเคลื่อนจากการพยากรณ์จะต้องมีลักษณะเป็นแบบสุ่ม (Random) และมีค่าเฉลี่ยเท่ากับศูนย์

เฉลยบทที่ 2

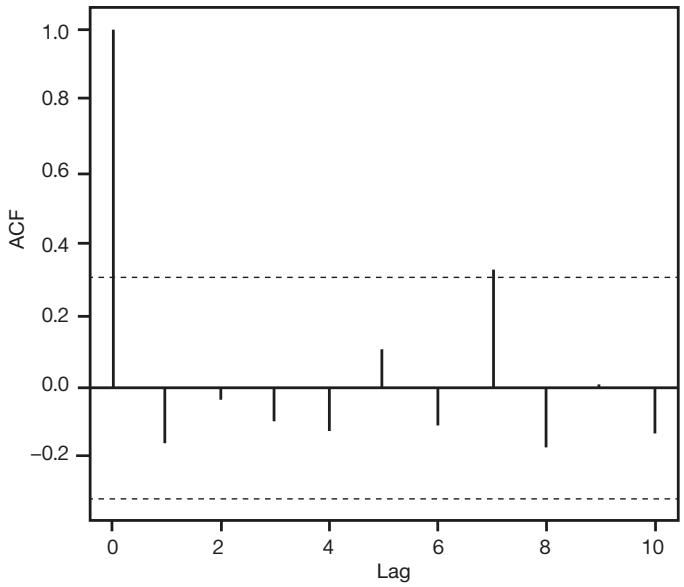
1. จากข้อมูลส่วนเหลือที่ได้จากการพยากรณ์ดังต่อไปนี้

ลำดับ	ส่วนเหลือ	ลำดับ	ส่วนเหลือ	ลำดับ	ส่วนเหลือ	ลำดับ	ส่วนเหลือ	ลำดับ	ส่วนเหลือ
1	0.7693	9	0.2637	17	0.6600	25	0.4488	33	0.5849
2	0.4249	10	0.7372	18	0.4543	26	0.5742	34	0.6568
3	0.7157	11	0.5455	19	0.0933	27	0.9278	35	0.0051
4	0.2046	12	0.7463	20	0.4491	28	0.2329	36	0.8882
5	0.3708	13	0.1156	21	0.2293	29	0.7059	37	0.7467
6	0.7270	14	0.2151	22	0.7474	30	0.7881	38	0.7432
7	0.4312	15	0.7380	23	0.9669	31	0.7656	39	0.4672
8	0.9137	16	0.3735	24	0.7297	32	0.3595	40	0.1234

1.1 จงคำนวณฟังก์ชันสหสัมพันธ์ในตัวของจำนวนช่วงเวลาที่ช้ากว่ากัน (Lag) ที่ 1, 2, 3, 4, 5

ตอบ

```
> acf(ex,lag.max=10,plot=F)

Autocorrelations of series 'ex', by lag

     0      1      2      3      4      5      6      7      8      9     10
 1.000 -0.154 -0.036 -0.097 -0.124  0.115 -0.108  0.329 -0.170  0.013 -0.131
```



1.2 จงคำนวณค่าสถิติ Ljung–Box Q

ตอบ

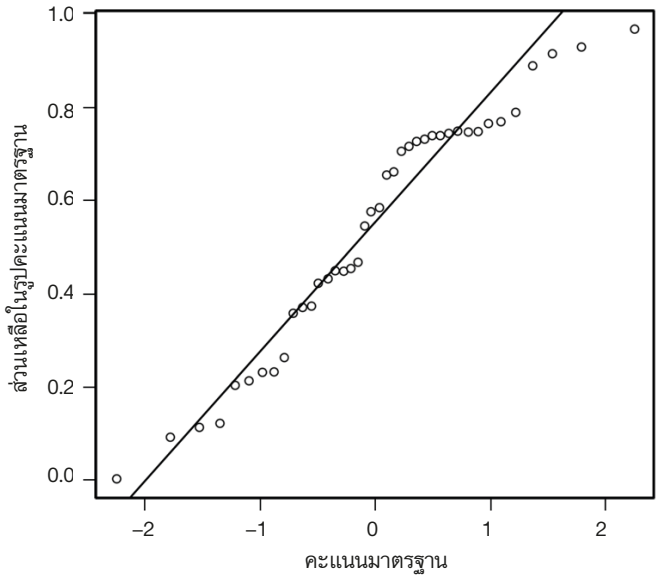
```
> Box.test(ex,lag=10,type="Ljung-Box")

        Box-Ljung test

data:  ex
X-squared = 11.4363, df = 10, p-value = 0.3245
```

1.3 จากผลที่ได้รับในข้อที่ 1.1 และ 1.2 จงสรุปว่าส่วนเหลือดังกล่าวมีลักษณะเป็นสัญญาณรบกวนขาวหรือไม่

ตอบ จากผลทั้งในข้อที่ 1.1 และ 1.2 ยืนยันตรงกันว่าข้อมูลส่วนเหลือดังกล่าวมีลักษณะเป็นสัญญาณรบกวนขาว เนื่องจากค่าสหสัมพันธ์ในตัวเท่ากับศูนย์ ซึ่งเป็นคุณสมบัติหนึ่งของสัญญาณรบกวนขาว แต่อย่างไรก็ดี ต้องตรวจสอบค่าความแปรปรวนและค่าเฉลี่ยของส่วนเหลือด้วย

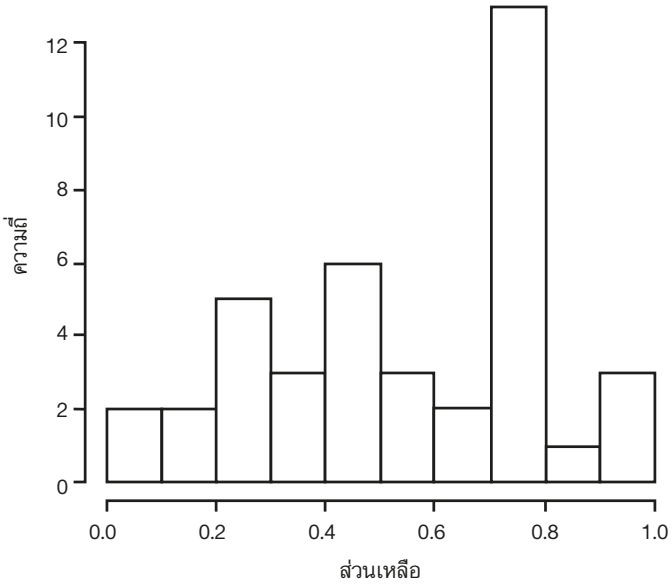
1.4 สร้างกราฟความน่าจะเป็นปกติ

ตอบ



1.5 สร้างแผนภาพฮิสโตแกรม

ตอบ



1.6 จากผลที่ได้รับในข้อที่ 1.4 และ 1.5 จงสรุปว่าส่วนเหลือดังกล่าวมีการแจกแจงความน่าจะเป็นปกติหรือไม่

ตอบ จากกราฟความน่าจะเป็นปกติ (ข้อมูลไม่มีการจัดเรียงกันเป็นเส้นตรง) และแผนภาพฮิสโตแกรม (การกระจายตัวมีลักษณะไม่สมมาตร มีลักษณะเบ้ซ้าย) จะเห็นได้ว่าข้อมูลดังกล่าวมีลักษณะไม่เป็นไปตามการแจกแจงความน่าจะเป็นปกติ

เฉลยบทที่ 3

1. บริษัทแห่งหนึ่งได้ทำการประมาณการรายรับในรูปของ GDP (Gross Domestic Product) ของประเทศ และได้ทำการสร้างสมการถดถอยออกมาดังต่อไปนี้ $\hat{y} = 0.78 + 0.6x$ จากสมการดังกล่าว

1.1 จงอธิบายรูปแบบความสัมพันธ์ระหว่างค่า GDP (x) และรายรับ (y)

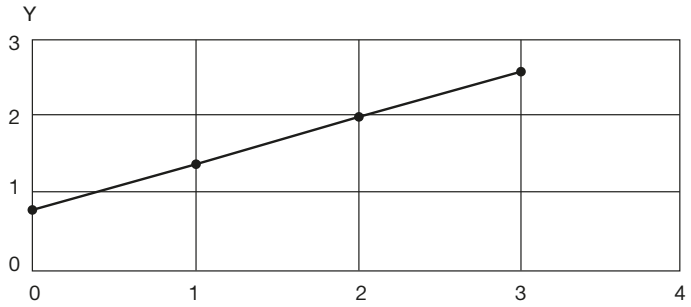
ตอบ ถ้า x = 1, 2, 3 จะได้ค่า y ดังนี้

x	y
1	1.38
2	1.98
3	2.58

จากตารางความสัมพันธ์จะเห็นได้ว่าทุกหนึ่งหน่วยที่เพิ่มขึ้นของ GDP หรือ x จะส่งผลต่อการเพิ่มขึ้นเป็นปริมาณ 0.6 หน่วยของรายรับหรือ y

1.2 จงอธิบายค่าจุดตัดแกน y

ตอบ จุดตัดแกน y คือค่า y เมื่อ x = 0 นั่นก็คือ 0.78



2. วิศวกรคุณภาพของโรงงานแห่งหนึ่งได้ทำการตรวจสอบคุณภาพของมอเตอร์ไฟฟ้าที่ผลิตออกมาจากสายการผลิตและได้ทำการสังเกตความสัมพันธ์ระหว่างปริมาณของมอเตอร์ที่ผลิตออกมาในแต่ละล็อตกับจำนวนมอเตอร์ที่ไม่ผ่านการตรวจสอบคุณภาพและได้ผลแสดงดังในตารางต่อไปนี้

ล็อตที่	จำนวนของเสีย	ขนาดของล็อต
1	4	25
2	8	50
3	6	75
4	16	100
5	22	125
6	27	150
7	36	175
8	49	200
9	53	225
10	70	250
11	82	275
12	95	300
13	109	325

จากข้อมูลดังกล่าว

(ก) จงสร้างสมการการถดถอยโดยมีจำนวนของเสียเป็นตัวแปรตามและขนาดของล็อตเป็นตัวแปรอิสระ

```
> x<-c(25,50,75,100,125,150,175,200,225,250,275,300,325)
> y<-c(4,8,6,16,22,27,36,49,53,70,82,95,109)
> fit<-lm(y~x)
> line(x,fit$f)

Call:
line(x, fit$f)

Coefficients:
[1]  -17.7308     0.3549
```

ตอบ $y = -17.7308 + 0.3549\,x + \varepsilon$

(ข) จงคำนวณค่า Var $(\hat{\beta}_0)$ และ Var $(\hat{\beta}_1)$

ตอบ

```
> summary(fit)

Call:
lm(formula = y ~ x)

Residuals:
   Min     1Q Median     3Q    Max
-9.132 -4.637 -1.764  6.247 12.857

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -17.73077    4.62644  -3.832  0.00278 **
x             0.35495    0.02332  15.224 9.75e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 7.863 on 11 degrees of freedom
Multiple R-squared:  0.9547,    Adjusted R-squared:  0.9506
F-statistic: 231.8 on 1 and 11 DF,  p-value: 9.75e-09
```

$\text{Var}(\hat{\beta}_0) = 4.62644^2 = 21.40395$

$\text{Var}(\hat{\beta}_1) = 0.02332^2 = 0.000544$

(ค) ทดสอบค่าสมมติฐานเกี่ยวกับค่า β_1 ที่ระดับนัยสำคัญ 0.05

ตอบ จากการทดสอบ พบว่าค่า p-value ของ β_1 มีค่าน้อยมาก $(9.75 * 10^{-9})$ ซึ่งมีค่าน้อยกว่า 0.05 ดังนั้นค่าสัมประสิทธิ์ β_1 จึงมีความเหมาะสมที่จะรวมอยู่ในตัวแบบ

(ง) คำนวณค่าสัมประสิทธ์สหสัมพันธ์ของกลุ่มตัวอย่าง

ตอบ $r = \hat{\beta}_0 = \sqrt{\dfrac{SS_{xx}}{SS_{yy}}} = 0.3549\sqrt{\dfrac{113{,}750}{15{,}011.08}} = 0.976958$

(จ) ทดสอบค่าสมมติฐานเกี่ยวกับค่า ρ ที่ระดับนัยสำคัญ 0.05

ตอบ $t = r\sqrt{\dfrac{n-2}{1-r^2}} = 0.976958\sqrt{\dfrac{13-2}{1-0.976958^2}} = 15.18137 > t_{0.025,\,13-2} = 2.2$

ดังนั้นจึงสามารถปฏิเสธ H_0 และสรุปได้ว่า $\rho \neq 0$

(ฉ) จงคำนวณค่าสัมประสิทธิ์การตัดสินใจ

ตอบ $CD = r^2 = 0.976958^2 = 0.954446$

(ช) คำนวณค่าช่วงของจำนวนของเสียที่ระดับความเชื่อมั่น 95% เมื่อขนาดของล็อตเท่ากับ 220

ตอบ $y_p \in \hat{y} \pm t_{\alpha/2} S \sqrt{1 + \frac{1}{n} + \frac{(x_p - \bar{x})^2}{SS_{xx}}}$

$y = -17.7308 + 0.3549 * 220 = 60.3472$

$t_{0.025,13-2} = 2.2$

$S = 7.863$

$X_p = 220$

$SS_{xx} = 113,750$

X bar = 175

$y_p = 60.3472 \pm 18.09937 = (42.24783, 78.44657) = (43, 79)$

ฉ) คำนวณค่าช่วงของค่าเฉลี่ยจำนวนของเสียที่ระดับความเชื่อมั่น 95% เมื่อขนาดของล็อตเท่ากับ 220

ตอบ $\mu_p \in \hat{y} \pm t_{\alpha/2,n-2} * S \sqrt{\frac{1}{n} + \frac{(x_p - \bar{x})^2}{SS_{xx}}}$

$\mu_p = 60.3472 \pm 5.324072 = (55.02313, 65.67127) = (56, 66)$

เฉลยบทที่ 4

1. จงอธิบายลักษณะของข้อมูลอนุกรมเวลาที่มีลักษณะเป็นไปตามกระบวนการคงที่และให้ความแตกต่างระหว่างกระบวนการคงที่เชิงกว้าง (Wide Sense) และเชิงแคบ (Strict Sense)

ตอบ ข้อมูลอนุกรมเวลาที่มีลักษณะเป็นไปตามกระบวนการคงที่เชิงแคบก็คือ ข้อมูลอนุกรมเวลาจะมีคุณสมบัติทางสถิติซึ่งไม่แปรผันไปตามเวลา หรือไม่ขึ้นอยู่กับจุดเริ่มต้นของเวลา

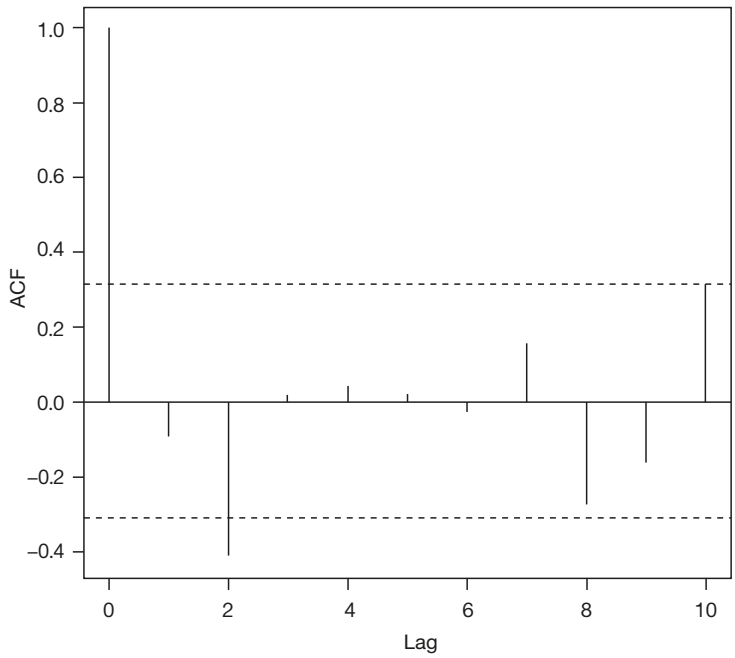
2. จากข้อมูลอนุกรมเวลาซึ่งแสดงถึงยอดขายของสินค้าชนิดหนึ่ง (หน่วย : พันชิ้น) ตั้งแต่เดือนที่ 1 ถึงเดือนที่ 40

เดือน	ยอดขาย	เดือน	ยอดขาย	เดือน	ยอดขาย	เดือน	ยอดขาย
1	29	11	23	21	22	31	33
2	24	12	22	22	29	32	23
3	30	13	29	23	31	33	23
4	32	14	39	24	40	34	38
5	27	15	31	25	28	35	31
6	29	16	22	26	21	36	27
7	35	17	34	27	36	37	30
8	28	18	29	28	36	38	24
9	28	19	25	29	36	39	33
10	39	20	28	30	24	40	33

จงดำเนินการดังต่อไปนี้

2.1 สร้างกราฟ ACF และวิเคราะห์ว่าอนุกรมเวลาดังกล่าวมีลักษณะเป็นไปตามรูปแบบของกระบวนการคงที่หรือไม่

ตอบ



2.2 ทำการทดสอบทางสถิติด้วยวิธีดังต่อไปนี้

2.2.1 Augmented Dickey–Fuller

ตอบ Dickey–Fuller = −4.2422, Lag order = 3, p–value = 0.01035

Alternative Hypothesis : Stationary

จากการทดสอบสมมติฐานจะเห็นได้ว่าสามารถปฏิเสธ H_0 ได้ (p–value = 0.01035 < 0.05) และสรุปว่าข้อมูลดังกล่าวมีลักษณะเป็นแบบ Stationary

2.2.2 KPSS

ตอบ สามารถแบ่งการทดสอบออกเป็น Level และ Trend

KPSS Test for Level Stationarity

data: ex

KPSS Level = 0.0411, Truncation lag parameter = 1, p–value = 0.1

จากการทดสอบสมมติฐานจะเห็นได้ว่าไม่สามารถปฏิเสธ H_0 ได้ (p–value = 0.1 > 0.05) และ

สรุปว่าข้อมูลดังกล่าวมีลักษณะเป็นแบบ Stationary ในรูปแบบ Level

KPSS Test for Trend Stationarity

data: ex

KPSS Trend = 0.0249, Truncation Lag parameter = 1, p–value = 0.1

จากการทดสอบสมมติฐานจะเห็นได้ว่าไม่สามารถปฏิเสธ H_0 ได้ (p–value = 0.1 > 0.05) และสรุปว่าข้อมูลดังกล่าวมีลักษณะเป็นแบบ Stationary ในรูปแบบ Trend

2.2.3 ทำการเปรียบเทียบผลลัพธ์ที่ได้ในข้อที่ 2.2.1 และ 2.2.2

ตอบ จากผลการทดสอบ ADF และ KPSS จะเห็นได้ว่าข้อมูลดังกล่าวนั้นมีลักษณะเป็นแบบ Stationary

เฉลยบทที่ 5

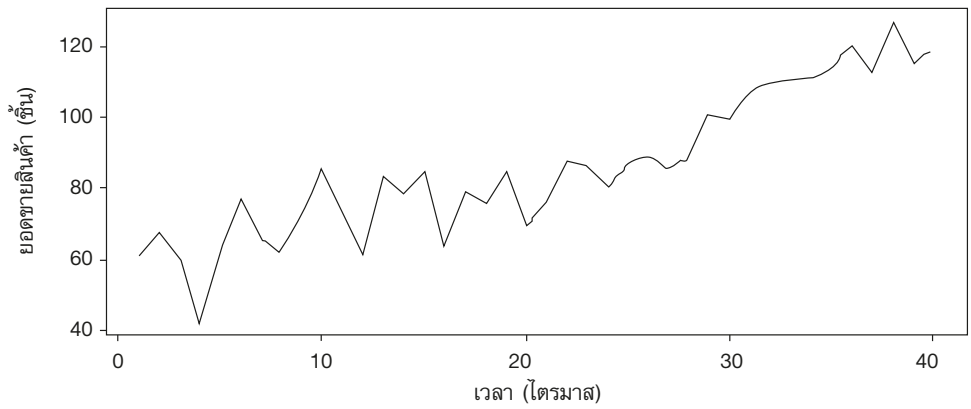
1. จากข้อมูลยอดขายสินค้าประเภทหนึ่งเป็นรายไตรมาส

ไตรมาส	y_i	ไตรมาส	y_i	ไตรมาส	y_i	ไตรมาส	y_i	ไตรมาส	y_i
1	60	9	72	17	79	25	87	33	111
2	68	10	86	18	75	26	89	34	111
3	59	11	75	19	85	27	85	35	114
4	42	12	61	20	69	28	89	36	120
5	63	13	84	21	76	29	101	37	112
6	77	14	78	22	88	30	99	38	127
7	66	15	84	23	86	31	107	39	115
8	62	16	63	24	80	32	110	40	119

จงตอบคำถามดังต่อไปนี้

1.1 พล็อตกราฟอนุกรมเวลา

ตอบ

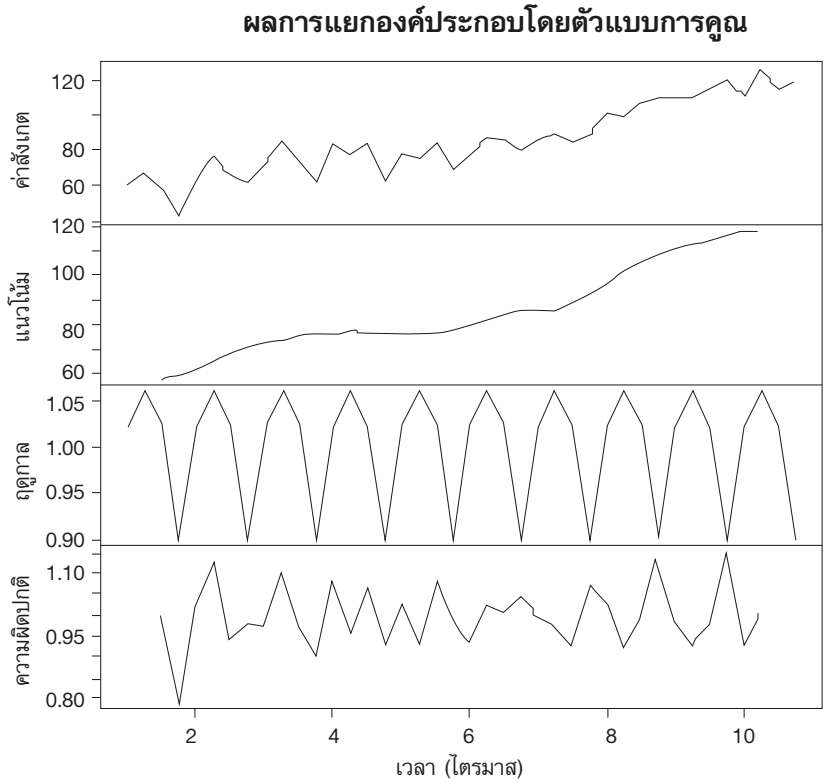


1.2 จากกราฟอนุกรมเวลา จงระบุตัวแบบการกระทำการทางคณิตศาสตร์ที่เหมาะสมในการแยกองค์ประกอบ

ตอบ จากกราฟจะเห็นได้ว่าข้อมูลอนุกรมเวลามีการกระจายตัวอย่างมากในช่วงแรก (0–10) และกระจายตัวอย่างคงที่ในช่วงถัดไป ดังนั้นจึงเห็นได้ว่าควรใช้การแยกองค์ประกอบแบบตัวแบบการคูณมาใช้

1.3 จงใช้โปรแกรม R ในการแยกองค์ประกอบต่างๆ ของอนุกรมเวลา

ตอบ



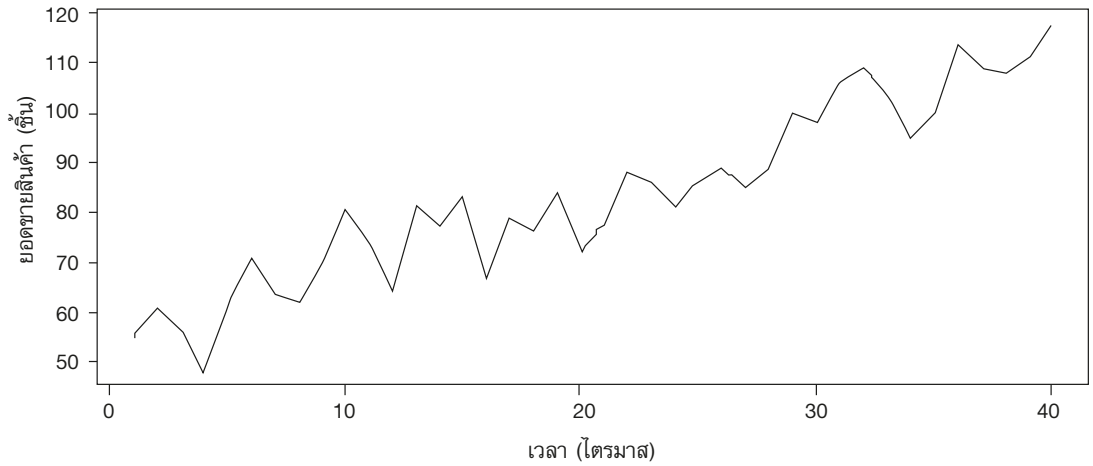
2. จากข้อมูลยอดขายสินค้าประเภทหนึ่งเป็นรายไตรมาส

ไตรมาส	y_i	ไตรมาส	y_i	ไตรมาส	y_i	ไตรมาส	y_i	ไตรมาส	y_i
1	55	9	70	17	79	25	87	33	104
2	61	10	81	18	76	26	89	34	95
3	57	11	74	19	84	27	85	35	100
4	48	12	64	20	72	28	89	36	114
5	61	13	81	21	78	29	100	37	109
6	71	14	77	22	88	30	98	38	108
7	64	15	83	23	86	31	106	39	111
8	62	16	67	24	81	32	109	40	118

จงตอบคำถามดังต่อไปนี้

2.1 พล็อตกราฟอนุกรมเวลา

ตอบ

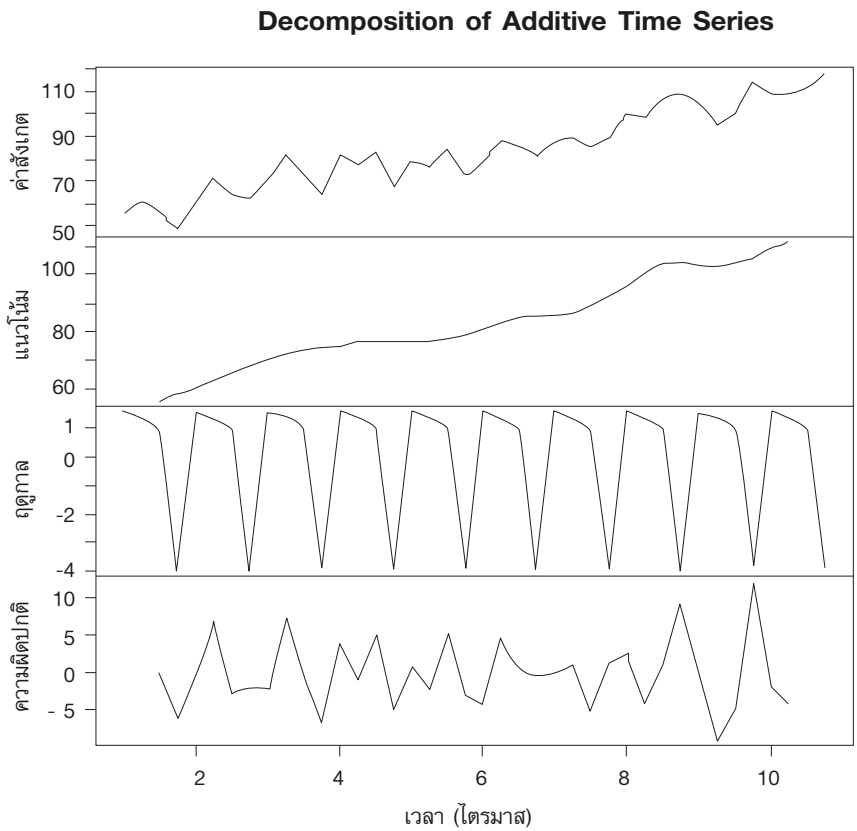


2.2 จากกราฟอนุกรมเวลา จงระบุตัวแบบการกระทำการทางคณิตศาสตร์ที่เหมาะสมในการแยกองค์ประกอบ

ตอบ จากการกระจายตัวของข้อมูล จะเห็นได้ว่าข้อมูลมีลักษณะการกระจายตัวแบบคงที่ ดังนั้นการแยกองค์ประกอบการบวกจึงอาจมีความเหมาะสมในการใช้แยกองค์ประกอบ

2.3 จงใช้โปรแกรม R ในการแยกองค์ประกอบต่างๆ ของอนุกรมเวลา

ตอบ



เฉลยบทที่ 6

1. จากข้อมูลอนุกรมเวลาดังต่อไปนี้

เดือนที่	ยอดขาย	เดือนที่	ยอดขาย
1	94	9	139
2	108	10	155
3	109	11	151
4	101	12	147
5	121	13	158
6	128	14	179
7	123	15	175
8	119	16	173

จงใช้ตัวแบบต่อไปนี้ในการพยากรณ์ค่าของข้อมูลข้างต้น

1.1 วิธีพยากรณ์อย่างง่าย

ตอบ

เดือนที่	y	yhat
1	94	
2	108	94
3	109	108
4	101	109
5	121	101
6	128	121
7	123	128
8	119	123
9	139	119
10	155	139
11	151	155
12	147	151
13	158	147
14	179	158
15	175	179
16	173	175

1.2 วิธีค่าเฉลี่ยเคลื่อนที่

ตอบ

Time Series:

Start = 1

End = 16

Frequency = 1

[1] NA NA 103.6667 106.0000 110.3333 116.6667 124.0000 123.3333

[9] 127.0000 137.6667 148.3333 151.0000 152.0000 161.3333 170.6667 175.6667

1.3 วิธีค่าเฉลี่ยเคลื่อนที่สองชั้น

ตอบ

t	y	SMA	DMA	Level	Trend	yhat
1	94					
2	108					
3	109	103.6667				
4	101	106.0000				
5	121	110.3333	106.6667	114	3.6667	
6	128	116.6667	111.0000	122.3333	5.6667	117.6667
7	123	124.0000	117.0000	131	7.0000	128.0000
8	119	123.3333	121.3333	125.3333	2.0000	138.0000
9	139	127.0000	124.7778	129.2222	2.2222	127.3333
10	155	137.6667	129.3333	146	8.3333	131.4444
11	151	148.3333	137.6667	159	10.6667	154.3333
12	147	151.0000	145.6667	156.3333	5.3333	169.6667
13	158	152.0000	150.4444	153.5556	1.5556	161.6667
14	179	161.3333	154.7778	167.8889	6.5556	155.1111
15	175	170.6667	161.3333	180	9.3333	174.4444
16	173	175.6667	169.2222	182.1111	6.4444	189.3333

1.4 วิธีทำให้เรียบด้วยเลขชี้กำลังชั้นเดียว $\alpha = 0.9$

ตอบ

t	y	level	yhat
1	94	94	
2	108	106.6	94
3	109	108.76	106.6
4	101	101.776	108.76
5	121	119.0776	101.776
6	128	127.1078	119.0776
7	123	123.4108	127.1078
8	119	119.4411	123.4108
9	139	137.0441	119.4411
10	155	153.2044	137.0441
11	151	151.2204	153.2044
12	147	147.4220	151.2204
13	158	156.9422	147.422
14	179	176.7942	156.9422
15	175	175.1794	176.7942
16	173		175.1794

1.5 วิธีทำให้เรียบด้วยเลขชี้กำลังสองชั้น $\alpha = 0.5$ และ $\beta = 0.4$

ตอบ

t	y	level	Trend	yhat
1	94			
2	108	108	14	
3	109	115.5	11.4	122
4	101	113.95	6.22	126.9
5	121	120.585	6.386	120.17
6	128	127.4855	6.5918	126.971
7	123	128.53865	4.37634	134.0773
8	119	125.957495	1.593342	132.915
9	139	133.2754185	3.883175	127.5508
10	155	146.0792966	7.451456	137.1586
11	151	152.2653763	6.945305	153.5308
12	147	153.1053409	4.503169	159.2107
13	158	157.804255	4.581467	157.6085
14	179	170.6928611	7.904323	162.3857
15	175	176.7985919	7.184886	178.5972
16	173	178.4917389	4.98819	183.9835

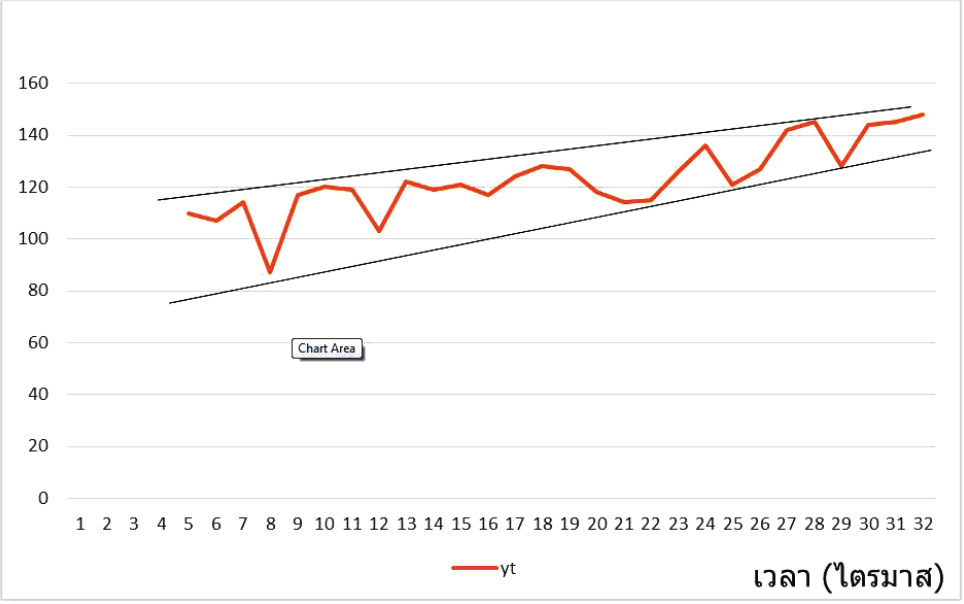
2. จากข้อมูลอนุกรมเวลาของยอดขายสินค้าชนิดหนึ่งซึ่งมีการเก็บรวบรวมไว้เป็นเวลาเจ็ดปี (28 ไตรมาส)

ปี	ไตรมาส	ค่าสังเกต
ปีที่ 1	Q1	110
	Q2	107
	Q3	114
	Q4	87
ปีที่ 2	Q1	117
	Q2	120
	Q3	119
	Q4	103
ปีที่ 3	Q1	122
	Q2	119
	Q3	121
	Q4	117
ปีที่ 4	Q1	124
	Q2	128
	Q3	127
	Q4	118
ปีที่ 5	Q1	114
	Q2	115
	Q3	126
	Q4	136
ปีที่ 6	Q1	121
	Q2	127
	Q3	142
	Q4	145
ปีที่ 7	Q1	128
	Q2	144
	Q3	145
	Q4	148

ถ้าค่าพารามิเตอร์ที่เหมาะสมในการคำนวณคือ $\alpha = 0.3$ $\beta = 0.1$ $\gamma = 0.1$

2.1 จงวิเคราะห์ว่าตัวแบบการบวกหรือการคูณเป็นตัวแบบที่เหมาะสมที่จะใช้ในการอธิบายข้อมูลดังกล่าวโดยใช้การพิจารณาจากกราฟ

ตอบ จากรูป จะเห็นได้ว่าเมื่อเวลาเพิ่มขึ้น ช่วงกว้างของการกระจายตัวของข้อมูลจะลดลง ซึ่งเป็นการบ่งชี้ว่าตัวแบบการคูณน่าที่จะมีความเหมาะสมในการอธิบายข้อมูลชุดนี้



2.2 ใช้วิธีทำให้เรียบด้วยเลขชี้กำลังสามชั้นในการคำนวณค่าส่วนประกอบระดับ แนวโน้ม ฤดูกาล พร้อมทั้งค่าพยากรณ์

ตอบ

ไตรมาส	yt	Lt	Tt	St	yhat
-3				1.052631579	
-2				1.023923445	
-1				1.090909091	
0		104.5	2.5625	0.832535885	
1	110	106.29375	2.485625	1.050112146	
2	107	107.4955625	2.35724375	1.021070093	111.3817524
3	114	108.2469644	2.196659563	1.087132913	119.839425
4	87	108.6605368	2.018350844	0.829348163	91.94828022
5	117	110.9002228	2.040484363	1.05060117	116.2252441
6	120	114.315624	2.177976049	1.023935611	115.3203784
7	119	114.384191	1.96703514	1.082454978	126.6440268
8	103	118.7040345	2.202315979	0.833183776	96.49567562
9	122	119.4716424	2.058845172	1.047657336	127.0243533
10	119	119.9368149	1.899477902	1.02076096	124.4393941
11	121	118.820287	1.597877323	1.076043942	131.8823016
12	117	126.4202756	2.198088454	0.842413844	100.3304608
13	124	125.540648	1.89031685	1.041664391	134.7479726
14	128	126.820669	1.829287262	1.019614784	130.076554
15	127	125.4624453	1.510536169	1.069665058	138.4330061
16	118	130.9031885	1.903556869	0.848315413	106.9637974
17	114	125.7967935	1.202561679	1.028120294	138.3400576
18	115	122.7358567	0.776211836	1.011350456	129.4904201
19	126	121.7966127	0.604666251	1.066149705	132.1165439
20	136	133.7762138	1.742159737	0.865146189	103.8348915
21	121	130.170014	1.20732378	1.018263621	139.3291901
22	127	129.6365375	1.033243754	1.008181618	132.8685304
23	142	131.4257121	1.108836833	1.067580563	139.3135487
24	145	143.0547039	2.160852334	0.879991397	114.6617599
25	128	139.3621453	1.57551124	1.008284295	147.8677181
26	144	141.505782	1.632323782	1.009126083	142.0907546
27	145	140.9430134	1.412814553	1.063700966	152.8114595
28	148	150.1041183	2.187643584	0.890590485	125.271904

เฉลยบทที่ 7

1. จากตัวแบบ AR(1) ดังต่อไปนี้ $y_t = 0.7 + 0.3\, y_{t-1} + \varepsilon_t$ ถ้า $\varepsilon_t \sim WN(0, 1)$ จงคำนวณหาค่า

1.1 โมเมนต์อันดับที่ 1

ตอบ เนื่องจาก $\xi = 0.7,\ \phi_1 = 0.3$

$$\mu = \xi/(1 - \phi_1) = 0.7/(1 - 0.3) = 1$$

1.2 โมเมนต์อันดับที่ 2

ตอบ $\rho_1 = \phi_1 = 0.3$

$$\gamma_0 = \sigma_y^2 = \frac{\sigma_\varepsilon^2}{1 - \phi_1\rho_1} = 1/(1 - 0.3 * 0.3) = 1.098901099$$

1.3 ความแปรปรวนร่วมในตัว ณ Lag ที่ 1

ตอบ $\gamma_1 = \phi_1\gamma_0 = \dfrac{\phi_1\sigma_\varepsilon^2}{1 - \phi_1\rho_1} = 0.3 * 1.098901099 = 0.32967$

1.4 สหสัมพันธ์ในตัว ณ Lag ที่ 1

ตอบ $\rho_1 = \gamma_1/\gamma_0 = \phi_1 = 0.3$

2. จากตัวแบบ MA(1) ดังต่อไปนี้ $y_t = 0.5 + 0.4\varepsilon_{t-1} + \varepsilon_t$ ถ้า $\varepsilon_t \sim WN(0, 1)$ และ $\mu = 0.5$ จงคำนวณหาค่า

2.1 โมเมนต์อันดับที่ 1

ตอบ เนื่องจากเป็นตัวแบบ MA(1) ดังนั้นโมเมนต์อันดับที่ 1 เท่ากับ 0

2.2 โมเมนต์อันดับที่ 2

ตอบ $Var[\tilde{y}_t] = \sigma_\varepsilon^2(1 + \theta_1^2) = 1(1 + 0.4^2) = 1.16$

2.3 ความแปรปรวนร่วมในตัว ณ Lag ที่ 1

ตอบ $\gamma_t = \sigma_\varepsilon^2\theta_1 = 1 * 0.4 = 0.4$

2.4 สหสัมพันธ์ในตัว ณ Lag ที่ 1

ตอบ $\rho_k = \dfrac{\theta_1}{1 + \theta_1^2} = 0.4/(1 + 0.4^2) = 0.344828$

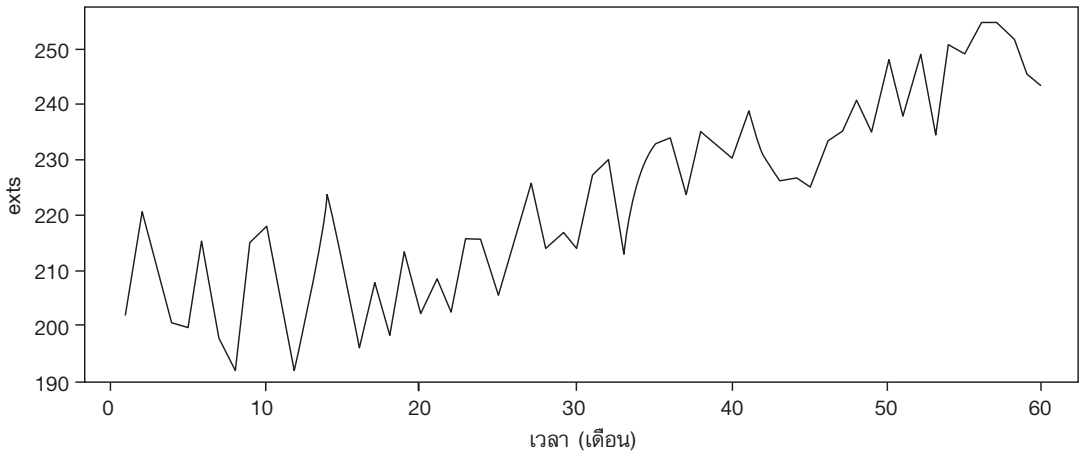
3. จากข้อมูลการใช้พลังงานไฟฟ้ารายเดือน (หน่วย MegaWatt, MW) ของพื้นที่แห่งหนึ่งดังต่อไปนี้

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2010	202	221	212	201	200	216	198	192	215	218	203	192
2011	205	225	212	196	209	198	214	202	209	202	216	216
2012	205	217	226	214	217	214	227	230	213	227	233	234
2013	223	235	233	230	239	231	226	227	225	233	235	241
2014	235	248	238	250	234	251	249	255	255	252	246	243

จงใช้โปรแกรม R ในการ

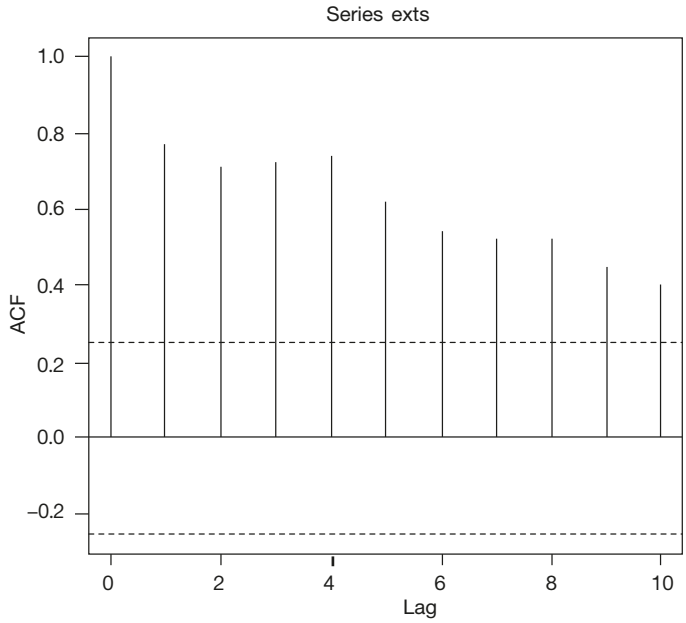
3.1 พล็อตข้อมูลอนุกรมเวลาดังกล่าว

ตอบ



3.2 สร้างกราฟ ACF ที่ Lag 1–10

ตอบ



3.3 จากกราฟในข้อที่ 3.2 จงวิเคราะห์เลือกประเภทของตัวแบบ ARIMA ที่มีความเหมาะสมที่สุดในการพยากรณ์ข้อมูลดังกล่าว

ตอบ จากกราฟ ACF จะเห็นได้ว่า ค่าของ ACF จะค่อยๆ ลดลง ดังนั้นจะมีลักษณะเป็นแบบ Non–stationary ดังนั้นตัวแบบที่เหมาะสมก็อาจจะเป็น IMA(1,1)

เฉลยบทที่ 8

1. จากข้อมูลอนุกรมเวลาชุดหนึ่ง

```
 [1] 48.54444 46.90893 49.16669 50.27591 50.45985 51.39441 50.60098 49.30251
 [9] 48.98619 48.55207 50.25295 49.17123 50.52526 51.08968 49.14077 51.02810
[17] 50.04657 49.02418 49.49457 51.30304 52.64640 54.11204 54.94484 53.08955
[25] 53.97284 53.95765 53.50182 53.39024 54.46148 52.76819 53.41712 54.16316
[33] 53.11404 50.56188 52.25581 52.18514 51.01109 52.41957 50.18504 50.50104
[41] 51.94850 52.42830 52.93569 50.80080 49.79196 50.94880 51.50110 49.22291
[49] 50.87704 50.63306 48.86644 49.37134 48.50727 47.96207 49.49451 50.64768
[57] 51.80521 52.17960 51.27763 52.48296
```

ถ้าต้องการใช้ตัวแบบ AR(1) ในการพยากรณ์ชุดข้อมูลดังกล่าว

1.1 จงคำนวณค่าสัมประสิทธิ์ AR และ ค่าจุดตัดแกน y โดยใช้วิธีกำลังสองน้อยสุด

ตอบ จากการคำนวณโดยใช้หลักกำลังสองน้อยสุด จะได้ว่า

Intercept 12.85255511

β_1 0.74964142 = ϕ

1.2 จงใช้โปรแกรม R ในการคำนวณค่าสัมประสิทธิ์ในข้อที่ 1.1 โดยกำหนดตัวแบบที่ใช้คือ AR(1)

ตอบ

```
> arima(arts,order=c(1,0,0))
Coefficients:
        ar1     intercept
      0.7614    51.0378
s.e.  0.0833    0.6113
```

1.3 จงใช้โปรแกรม R ในการเลือกตัวแบบที่เหมาะสมที่สุดในการพยากรณ์

ตอบ

```
> auto.arima(arts)
Series: arts
ARIMA(0,1,1) with drift
Coefficients:
         ma1     drift
      -0.2927    0.0670
s.e.   0.1409    0.1142
```

2. จากข้อมูลอนุกรมเวลาชุดหนึ่ง

```
 [1] 20.53172 18.43413 19.21637 21.69705 22.62497 22.31641 21.95769 18.90365
 [9] 17.73094 19.05077 17.90767 17.90082 20.31899 21.20725 18.52357 17.59815
[17] 20.58481 22.23272 22.10585 22.32999 22.12615 22.19074 20.52519 21.88054
[25] 19.59788 20.71425 19.57213 18.42645 17.92505 20.04024 21.79462 21.02761
[33] 21.87919 19.42313 17.77162 18.82022 18.50957 18.64250 18.34020 19.28987
[41] 19.66354 21.63841 21.96898 21.57616 20.89908 18.48772 18.02525 18.15466
[49] 20.22590 18.44170 18.24805 20.24619 22.03597 19.82746 21.17978 20.39568
[57] 20.52220 21.86053 21.21747 19.37643
```

ถ้าต้องการใช้ตัวแบบ MA(1) ในการพยากรณ์ชุดข้อมูลดังกล่าว

2.1 จงคำนวณค่าสัมประสิทธิ์ MA และ ค่าจุดตัดแกน y โดยใช้วิธีกำลังสองน้อยสุด

ตอบ

Autocorrelations of series ‘mats’, by lag

0	1	2	3	4	5	6	7	8	9	10
1.000	0.557	0.145	–0.049	–0.168	–0.278	–0.212	–0.195	–0.215	–0.089	–0.038

ดังนั้นจะได้ว่า

$$\frac{\theta_1}{1 + \theta_1^2} = 0.557$$

$$0 = 0.557\,\theta_1^2 - \theta_1 + 0.557$$

จากสมการดังกล่าวจะพบว่าการใช้สมการ Quadratic ไม่สามารถจะหาผลเฉลยได้ ดังนั้น ในกรณีการใช้หลักการโมเมนต์อาจจะไม่เหมาะสมในการคำนวณ

2.2 จงใช้โปรแกรม R ในการคำนวณค่าสัมประสิทธิ์ในข้อที่ 2.1 โดยกำหนดตัวแบบที่ใช้คือ MA(1)

ตอบ

```
arima(x = mats, order = c(0, 0, 1))

Coefficients:
         ma1     intercept
      0.5627     20.0949
s.e.  0.0882     0.2542

sigma^2 estimated as 1.607:  log likelihood = –99.56,  aic = 205.13
```

2.3 จงใช้โปรแกรม R ในการเลือกตัวแบบที่เหมาะสมที่สุดในการพยากรณ์

ตอบ

```
ARIMA(1,0,1) with non–zero mean

Coefficients:
         ar1      ma1      intercept
      0.3281   0.3608    20.0938
s.e.  0.1777   0.1658    0.3195
sigma^2 estimated as 1.53:  log likelihood=–98.14
AIC=204.27   AICc=205   BIC=212.65
```

เฉลยบทที่ 9

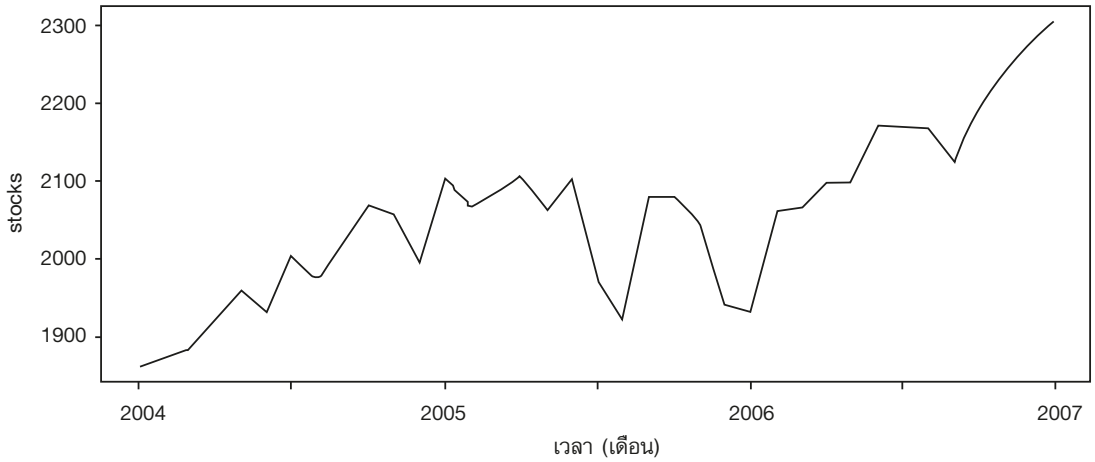
จากข้อมูลราคาปิดเฉลี่ยรายเดือนของดัชนีราคาหุ้น S&P 500 ระหว่างเดือน กุมภาพันธ์ 2014 ถึง กุมภาพันธ์ 2017

เดือน	ราคาปิด	เดือน	ราคาปิด	เดือน	ราคาปิด	เดือน	ราคาปิด
2014/1	1859.45	2015/1	1994.99	2016/1	1940.24	2017/1	2278.87
2014/2	1872.34	2015/2	2104.5	2016/2	1932.23	2017/2	2307.87
2014/3	1883.95	2015/3	2067.89	2016/3	2059.74		
2014/4	1923.57	2015/4	2085.51	2016/4	2065.3		
2014/5	1960.23	2015/5	2107.39	2016/5	2096.95		
2014/6	1930.67	2015/6	2063.11	2016/6	2098.86		
2014/7	2003.37	2015/7	2103.84	2016/7	2173.6		
2014/8	1972.29	2015/8	1972.18	2016/8	2170.95		
2014/9	2018.05	2015/9	1920.03	2016/9	2168.27		
2014/10	2067.56	2015/10	2079.36	2016/10	2126.15		
2014/11	2058.9	2015/11	2080.41	2016/11	2198.81		
2014/12	1994.99	2015/12	2043.94	2016/12	2238.83		

1. จงทำการตรวจสอบสภาวะความเป็นวิวิธพันธุ์ของข้อมูลดังกล่าวโดยใช้การวิเคราะห์ดังต่อไปนี้

1.1 จากกราฟอนุกรมเวลา

ตอบ

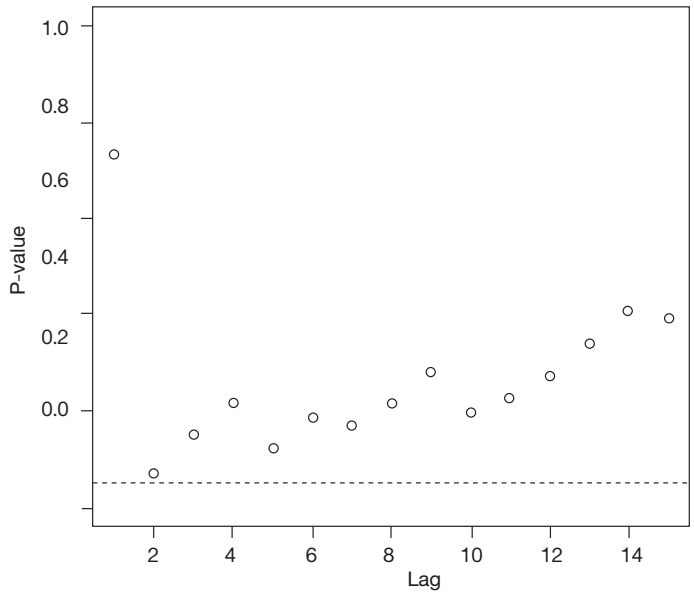


1.2 จากการทดสอบทางสถิติแบบ McLeod–Li

ตอบ

```
> fit<–auto.arima(stockts)
> fit
Series: stockts
ARIMA(0,1,0) with drift
Coefficients:
        drift
      12.4561
s.e.   9.8716

sigma^2 estimated as 3508:  log likelihood=–192.02
AIC=388.04   AICc=388.4   BIC=391.21
```

แล้วจึงทำการทดสอบ McLeod–Li กับค่าส่วนเหลือ ได้ผลดังต่อไปนี้



ซึ่งจะเห็นได้ว่าไม่สามารถปฏิเสธ H_0 ได้จึงสรุปว่าข้อมูลดังกล่าวไม่ได้รับผลจากสภาวะวิวิธพันธุ์

1.3 จากการทดสอบทางสถิติแบบ Lagrange Multiplier

เมื่อค่าระดับความมีนัยสำคัญเท่ากับ 0.05

ตอบ

```
ArchTest(stockts,lags=12,demean = FALSE)

    ARCH LM-test; Null hypothesis: no ARCH effects
data:  stockts
Chi-squared = 18.6003, df = 12, p-value = 0.09864
```

จากการทดสอบจะเห็นได้ว่า ค่า p–value > 0.05 ดังนั้นจึงไม่สามารถปฏิเสธ H_0 และสรุปได้ว่าข้อมูลดังกล่าวไม่ได้รับผลจากสภาวะวิวิธพันธุ์

2. จงใช้โปรแกรม R ในการสร้างตัวแบบ ARCH หรือ GARCH ที่มีความเหมาะสมในการอธิบายชุดข้อมูลดังกล่าว

ตอบ เนื่องจากผลการทดสอบให้ผลที่ตรงกันคือ ข้อมูลไม่ได้รับผลจากสภาวะวิวิธพันธุ์ ดังแสดงในข้อที่ 1 ดังนั้นจะเห็นได้ว่าตัวแบบ ARIMA ก็มีความพอเพียงในการอธิบายข้อมูลดังกล่าว จึงไม่มีความจำเป็นที่จะต้องใช้ตัวแบบ ARCH หรือ GARCH ในการสร้างตัวแบบเพื่ออธิบายข้อมูลดังกล่าว

เฉลยบทที่ 10

อ้างถึงฐานข้อมูล Open Source http://homes.esat.kuleuven.be/~smc/daisy/daisydata.html

1. จากค่าของข้อมูลสัญญาณนำเข้าและข้อมูลสัญญาณส่งออกชุดที่ 99–001 ซึ่งเป็นข้อมูลของระบบทำความร้อนแบบ SISO (Single Input Single Output) โดยอาศัยหลอดฮาโลเจนขนาด 300 วัตต์ สำหรับอินพุตก็คือ ค่าความต่างศักย์ที่ป้อนให้กับหลอดฮาโลเจน ในขณะที่สัญญาณส่งออกก็คืออุณหภูมิที่วัดได้ในหน่วยองศาเซลเซียส จากข้อมูลดังกล่าว จงใช้โปรแกรม Scilab ในการสร้างตัวแบบตามหลักการระบุเอกลักษณ์ดังต่อไปนี้ โดยใช้ข้อมูลชุดที่ 1–400 ในการสร้างตัวแบบและข้อมูลชุดที่ 401–800 ในการ Validation

1.1 ตัวแบบ AR

ตอบ

```
–>[m] = armax(y,u,3,0,0);
 m.a
 ans  =

    1.  – 1.7304246     0.5310742     0.1994270
–>m.b
 ans  =

    []

––>m.c
 ans  =
```

1.

1.2 ตัวแบบ ARX

ตอบ

```
–>[m]=armax(y, u, 3, 3, 0);

––>m.c
 ans  =

    1.

––>m.b
 ans  =

    0.0043987     0.0975953     0.1790260

––>m.a
 ans  =

    1.  – 1.4637867     0.3930474     0.0837967
```

1.3 ตัวแบบ ARMAX

ตอบ

```
––>[m] = armax(y, u, 3, 3, 3);

––>m.c
 ans  =

    1.  – 0.9212276     0.4815650     0.0894576
```

```
-->m.b
 ans  =

    0.0024147     0.1186100     0.0628516

-->m.a
 ans  =

    1.  - 2.1521823     1.6416265  - 0.4810792
```

1.4 ตัวแบบ BJ

ตอบ

```
->[m] = armax(y, u, 0, 3, 3);

-->m.c
 ans  =

    1.     1.0132305     1.0069656     0.9657013

-->m.b
 ans  =

  - 0.6358336  - 3.2086502     21.891782

-->m.a
 ans  =

    1.
```

1.5 ตัวแบบ OE

ตอบ

```
–>[m] = armax(y, u, 0, 3, 0);

––>m.c
 ans  =

    1.

––>m.b
 ans  =

  – 0.1480887     0.1059745     21.296231

––>m.a
 ans  =

    1.
```