บทที่ 2

สถิติที่ใช้ในการวิจัย

ในกรณีที่ผู้ทำการพยากรณ์ต้องการนำเทคนิคการทำให้เรียบแบบเอกซ์โพเนนเชียลไปใช้ในการวิเคราะห์อนุกรมเวลาที่มีข้อมูลสูญหายนั้น วิธีพยากรณ์ที่ใช้สำหรับวิเคราะห์ข้อมูลที่เหลืออยู่มีหลายวิธี แต่วิธีพยากรณ์ที่ใช้ในการศึกษาครั้งนี้มี 3 วิธี คือ วิธีพยากรณ์ที่มีการปรับแก้ด้วยวิธีของอัลด์รินและแดมส์เลท วิธีพยากรณ์ที่มีการปรับแก้ด้วยวิธีของไรท์ และวิธีพยากรณ์ที่มีการประมาณค่าสูญหาย ดังนั้นในบทนี้จะขอกล่าวถึงรายละเอียดของวิธีพยากรณ์แต่ละวิธี โดยมีรายละเอียดต่างๆ ดังนี้

2.1 วิธีพยากรณ์ที่ใช้ในการศึกษา

เนื่องจากเทคนิคการทำให้เรียบแบบเอกซ์โพเนนเชียลที่ใช้ในการวิจัยนี้มี 2 เทคนิค คือ เทคนิคการทำให้เรียบแบบเอกซ์โพเนนเชียลครั้งเดียวและซ้ำสองครั้ง ดังนั้นจะกล่าวถึงรายละเอียดของวิธีพยากรณ์แต่ละวิธี แยกเป็นแต่ละเทคนิคดังนี้

2.1.1 วิธีพยากรณ์เมื่อใช้กับเทคนิคการทำให้เรียบแบบเอกซ์โพเนนเชียลครั้งเดียว

การทำให้เรียบแบบเอกซ์โพเนนเชียลครั้งเดียว เป็นเทคนิคการพยากรณ์ที่เหมาะสมกับข้อมูลอนุกรมเวลาที่มีรูปแบบคงที่ (constant model):

$$y_t = \mu + \varepsilon_t \qquad , \; t=1,2,\ldots,n$$

โดยที่ μ คือระดับ หรือค่าเฉลี่ยของข้อมูล

การพยากรณ์โดยใช้เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียว จะกำหนดน้ำหนัก (weights) ให้กับข้อมูลหรือค่าสังเกตในอดีตไม่เท่ากัน โดยจะให้น้ำหนักกับค่าสังเกตปัจจุบันมากกว่าค่าสังเกตที่ผ่านมา และวิธีการที่ใช้กำหนดน้ำหนักให้กับค่าสังเกต คือ จากน้ำหนักทั้งหมด 1 จะกระจายให้กับค่าสังเกตในลักษณะที่ลดลงแบบเรขาคณิต หรือแบบเอกซโพเนนเชียล โดยเริ่มจากค่าสังเกตปัจจุบันจะมีน้ำหนักมากที่สุด และน้ำหนักจะลดลงเรื่อยๆ แบบเรขาคณิตสำหรับค่าสังเกตในอดีต ฉะนั้น ค่าพยากรณ์สำหรับค่าสังเกตในอนาคตที่เวลา $n+\tau$ (τ = 1, 2,...) จากเวลาปัจจุบัน n คำนวณได้ดังนี้

$$\hat{y}_n(\tau) = c \sum_{t=0}^{n-1} a^t y_{n-t} \qquad \ldots(2.1)$$

ค่าคงที่ a ($|a| < 1$) คือ สัมประสิทธิ์ส่วนลด (discount coefficient) แต่โดยทั่วไปมักจะใช้ a ที่มีค่าอยู่ระหว่าง 0 ถึง 1 ($0<a<1$) ตัวคูณ c ในสมการที่ (2.1) เป็นค่าคงที่ซึ่งใช้เพื่อปรับผลรวมของน้ำหนักให้เป็น 1

เนื่องจาก $$c \sum_{t=0}^{n-1} a^t = 1 \qquad \ldots(2.2)$$

เพราะฉะนั้น $$c = (1-a)/(1-a^n) \qquad \ldots(2.3)$$

เพราะว่า $$\sum_{t=0}^{n-1} a^t = (1-a^n)/(1-a)$$

จากสมการที่ (2.3) ถ้า n มีขนาดใหญ่ ($n \to \infty$) พบว่า a^n จะมีค่าเข้าใกล้ศูนย์ ($a^n \to 0$) เพราะฉะนั้น

$$\hat{y}_n(\tau) = (1-a)\sum_{t\geq 0} a^t y_{n-t}$$

$$= (1-a)[y_n + ay_{n-1} + a^2 y_{n-2} + \ldots] \qquad \ldots(2.4)$$

สำหรับที่เวลา t ใดๆ ($t=1,2,\ldots,n$) จะกำหนดให้

$$\hat{y}_t(\tau) = S_t = S_t^{(1)} = (1-a)[y_t + ay_{t-1} + a^2 y_{t-2} + \ldots] \qquad \ldots(2.5)$$

จากการแทนค่าอย่างง่าย จะแสดงได้ว่า

$$\hat{y}_t(\tau) = S_t = (1-a)y_t + aS_{t-1} \qquad ...(2.6)$$

และจะเรียก S_t ว่า ค่าของการทำให้เรียบ (smoothed value) แบบเอกซโพเนนเชียลอันดับที่หนึ่ง ณ เวลา t สำหรับการพยากรณ์โดยใช้เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียวนี้ จะใช้สมการที่ (2.6) ปรับค่าของการทำให้เรียบ S_t ไปทุกๆ คาบเวลา (t=1, 2,...,n) และที่เวลา t=n , ค่า S_n จะเป็นค่าพยากรณ์ของข้อมูลในอนาคตทุกค่าดังนี้

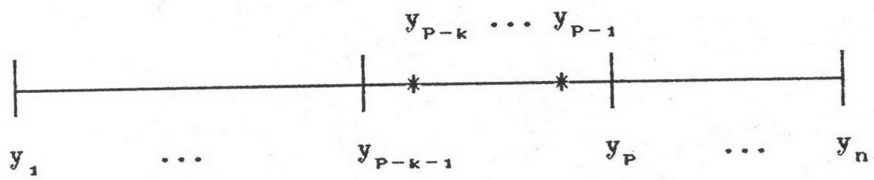
$$\hat{y}_n(\tau) = S_n = (1-a)y_n + aS_{n-1} \quad ; \tau=1,2,... \qquad ...(2.7)$$

เมื่อ $\hat{y}_n(\tau)$ แทน ค่าพยากรณ์ล่วงหน้า τ คาบเวลา จากเวลาที่ t = n

จากสมการที่ (2.7) จะพบว่า การคำนวณหาค่า S_t ที่เวลา t=1 จะต้องทราบค่า S_o หรือค่าเริ่มต้นของการทำให้เรียบ (initial smoothed value) เพราะว่า S_1 = $(1-a)y_1 + aS_o$ ดังนั้นจะต้องทำการกำหนดค่าเริ่มต้น S_o ก่อน จึงจะทำการพยากรณ์ต่อไปได้ และการกำหนดค่าเริ่มต้น S_o ว่าควรมีค่าเป็นเท่าไร สามารถทำได้หลายวิธี เช่น ใช้ค่าเฉลี่ยเลขคณิตของข้อมูลที่มีอยู่ : y_1, y_2 ,..., y_n (Brown, 1962 and Montgomery and Johnson, 1976, quoted in Abraham and Ledolter, 1983) ซึ่งการกำหนดค่าเริ่มต้นด้วยวิธีนี้ จะใช้ได้ดีถ้าระดับค่าเฉลี่ย (mean level) ของข้อมูลมีการเปลี่ยนแปลงอย่างช้าๆ (a มีค่าเข้าใกล้ 1) หรือใช้ข้อมูลตัวแรกเป็นค่าเริ่มต้น (Makridakis and Wheelwright, 1978, quoted in Abraham and Ledolter, 1983) ถ้าระดับค่าเฉลี่ยของข้อมูลมีการเปลี่ยนแปลงอย่างรวดเร็ว (a มีค่าเข้าใกล้ 0) นอกจากจะมีปัญหาในการกำหนดค่าเริ่มต้นแล้ว การพยากรณ์โดยใช้เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียว ยังมีปัญหาในการกำหนดค่าสัมประสิทธิ์ส่วนลด a ซึ่งการกำหนดค่า a ว่าควรเป็นเท่าไรนั้น อาจพิจารณาจากการเปลี่ยนแปลงระดับค่าเฉลี่ยของข้อมูลว่ามีการเปลี่ยนแปลงช้าหรือเร็วอย่างไร ถ้าระดับของข้อมูลมีการเปลี่ยนแปลงอย่างช้าๆ ค่า a ควรมีค่าเข้าใกล้หนึ่ง แต่ถ้าระดับค่าเฉลี่ยของข้อมูลเปลี่ยนแปลงอย่างรวดเร็ว ค่า a ควรมีค่าเข้าใกล้ศูนย์ ในทางปฏิบัติโดยทั่วไปจะเลือก a ที่มีค่า

อยู่ระหว่าง 0.7 และ 0.95 (Brown, 1962, quoted in Abraham and Ledolter, 1983) หรืออาจจะเลือกค่า a จากการสร้างค่าความคลาดเคลื่อนในการพยากรณ์ล่วงหน้า 1 คาบเวลา (one-step-ahead forecast error) โดยใช้ a ที่ค่าต่างๆ แล้วเลือกค่า a ที่ทำให้ค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง (Mean Square Error : MSE) ต่ำที่สุด ไปใช้ในการพยากรณ์ต่อไป และในการวิจัยครั้งนี้ จะเลือกใช้ค่า a ที่ให้ค่า MSE ต่ำสุด

เมื่อข้อมูลเกิดการสูญหายไป 1 ช่วง รายละเอียดการสูญหายของข้อมูลจะแสดงดังรูป



กำหนดให้ $y_1,...,y_{p-k-1}$ และ $y_p,...,y_n$ เป็นค่าสังเกตหรือข้อมูลที่ทราบค่า ส่วน $y_{p-k},...,y_{p-1}$ เป็นข้อมูลที่สูญหาย ดังนั้น ข้อมูลชุดนี้จะมีจำนวนข้อมูลสูญหาย k ตัว ซึ่งข้อมูลที่มีการสูญหายในลักษณะดังกล่าว สามารถนำมาพยากรณ์โดยใช้เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียว ด้วยวิธีพยากรณ์ดังต่อไปนี้

1. <u>วิธีพยากรณ์โดยใช้เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียวที่มีการปรับแก้ด้วยวิธีของอัลด์รินและแดมส์เลท</u>

เป็นวิธีการปรับแก้เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียวที่เสนอโดย Magne Aldrin และ Eivind Damsleth (1987) วิธีนี้จะไม่ประมาณข้อมูลที่สูญหาย แต่จะปรับค่าคงที่ (a) ที่ใช้กำหนดน้ำหนักให้กับข้อมูลหลังช่วงข้อมูลสูญหายตัวแรก, y_p (first observation after the gap) เท่านั้น ส่วนข้อมูลตัวต่อไปจะกำหนดน้ำหนักด้วยค่าคงที่ (a) ค่าเดิม การคำนวณหาค่าพยากรณ์จะมีขั้นตอนดังนี้

1.1 คำนวณหาค่าของการทำให้เรียบแบบเอกซโพเนนเชียลอันดับที่หนึ่ง ($S_t^{(1)} = S_t$) ของข้อมูลช่วงแรก (ช่วงก่อนข้อมูลสูญหาย) จากสูตร

$$S_t = (1-a)y_t + aS_{t-1} \text{ , } t = 1,2,...,p-k-1 \qquad ...(2.8)$$

1.2 คำนวณหาค่า S_t, $t = p,p+1,p+2,\ldots,n$ ของข้อมูลหลังช่วงที่มีข้อมูลสูญหาย ดังนี้

1.2.1 ที่เวลา $t = p$ จะคำนวณหาค่า S_p จากสูตร

$$S_p = (1-a^*)y_p + a^*S_{p-k-1} \qquad \ldots(2.9)$$

โดยที่

$$a^* = a/[1+k(1-a)^2] \qquad \ldots(2.10)$$

และ k = จำนวนข้อมูลสูญหาย

1.2.2 ที่เวลา $t = p+1,p+2,\ldots,n$ จะคำนวณหาค่า S_t, จากสมการที่ (2.8) เช่นเดียวกับข้อมูลในช่วงแรก

1.3 คำนวณหาค่าพยากรณ์จากสูตร

$$\hat{y}_n(\tau) = S_n = (1-a)y_n + aS_{n-1}\ , \ \tau = 1,2,\ldots \qquad \ldots(2.11)$$

2. <u>วิธีพยากรณ์โดยใช้เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียวที่มีการปรับแก้ด้วยวิธีของไรท์</u>

เป็นวิธีการปรับแก้เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียวที่เสนอโดย D.H. Wright (1986) วิธีของไรท์นี้ จะไม่ประมาณข้อมูลที่สูญหายเช่นเดียวกับวิธีของอัลด์รินและแดมส์เลท แต่ค่าถ่วงน้ำหนัก (a) ที่ใช้กำหนดน้ำหนักให้กับข้อมูลหลังช่วงข้อมูลสูญหาย $(y_p,y_{p+1},\ldots,y_n)$ จะถูกปรับไปทุกๆ คาบเวลา ซึ่งขั้นตอนของการคำนวณหาค่าพยากรณ์มีดังนี้

2.1 คำนวณหาค่า S_t ของข้อมูลช่วงแรก $(y_1,y_2,\ldots,y_{p-k-1})$ จากสูตร

$$S_t = (1-a)y_t + aS_{t-1}\ ,\ t = 1,2,\ldots,p-k-1$$

2.2 คำนวณหาค่า S_t ของข้อมูลหลังช่วงข้อมูลสูญหาย $(y_p,y_{p+1},\ldots,y_n)$ ดังนี้

2.2.1 ที่เวลา $t = p$ จะคำนวณหาค่า S_p จากสูตร

$$S_p = (1-a_p)y_p + a_p S_{p-k-1} \qquad ...(2.12)$$

เมื่อ $a_p = w_p/(w_p+1-a_{p-k-1})$ และ $w_p = a^{k+2}$

ค่าเริ่มต้นสำหรับ a_o คือ $a_o = a^{s+1}$, k = จำนวนข้อมูลสูญหาย
โดยที่ s คือ ค่าเฉลี่ยของช่วงเวลาระหว่างข้อมูลที่มีความต่อเนื่อง (ไม่สูญหาย) ทั้งก่อนและหลังข้อมูลสูญหาย (the average time interval between successive data points) ซึ่งคำนวณได้จากสูตร

$$s = (n-k-I)/I$$

เมื่อ I คือ จำนวนช่วงที่ข้อมูลมีความต่อเนื่อง ตัวอย่างเช่น กรณีที่ข้อมูลสูญหายไป 1 ช่วง จะมีจำนวนช่วงที่ข้อมูลต่อเนื่องกัน 2 ช่วง คือ ข้อมูลในช่วงก่อนและหลังข้อมูลสูญหาย ดังนั้น I จะมีค่าเท่ากับ 2 ในกรณีดังกล่าว

2.2.2 ที่เวลา $t = p+1, p+2, ..., n$ จะคำนวณหาค่า S_t จากสูตร

$$S_t = (1-a_t)y_t + a_t S_{t-1} \qquad ...(2.13)$$

โดยที่ $a_t = w_t/(w_t+1-a_{t-1})$, $w_t = a^2$ ($\because k=0$)

2.3 คำนวณหาค่าพยากรณ์จากสูตร

$$\hat{y}_n(\tau) = S_n = (1-a_n)y_n + a_n S_{n-1} \quad , \tau = 1,2,... \qquad ...(2.14)$$

3. <u>วิธีพยากรณ์โดยใช้เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียวที่มีการประมาณค่าสูญหาย</u>

วิธีที่มีการประมาณค่าสูญหายนี้ จะประมาณข้อมูลส่วนที่สูญหายโดยใช้สมการพยากรณ์ ที่ได้จากการวิเคราะห์ข้อมูลในช่วงแรก $(y_1, y_2, \ldots, y_{p-k-1})$ พยากรณ์ไปทุกช่วงเวลาที่มีข้อมูลสูญหาย โดยมีขั้นตอนการคำนวณหาค่าพยากรณ์ดังนี้

3.1 คำนวณหาค่า S_t, $t = 1,2,\ldots,p-k-1$ ของข้อมูลช่วงแรกจากสูตร

$$S_t = (1-a)y_t + aS_{t-1}$$

3.2 ใช้ค่าของการทำให้เรียบแบบเอกซโพเนนเชียลที่เวลา $t=p-k-1$ (ค่า S_{p-k-1}) เป็นค่าประมาณของข้อมูลสูญหายทุกช่วงเวลา

$$\begin{aligned}\hat{y}_{p-k-1}(\tau) &= S_{p-k-1} \quad , \tau = 1,2,\ldots,k \\ &= (1-a)y_{p-k-1} + aS_{p-k-2} \qquad \ldots(2.15)\end{aligned}$$

3.3 คำนวณหาค่า $S_t = (1-a)y_t + aS_{t-1}$, $t = p-k, p-k+1, \ldots, p-1, p, p+1, \ldots, n$ ของข้อมูลในช่วงที่สูญหายและข้อมูลหลังช่วงสูญหาย

3.4 หาค่าพยากรณ์จากสมการที่ (2.11)

2.1.2 <u>วิธีพยากรณ์เมื่อใช้กับเทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้ง</u>

การทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้ง เป็นเทคนิคการพยากรณ์ที่เหมาะสมกับข้อมูลอนุกรมเวลาที่มีรูปแบบเป็นแนวโน้มเชิงเส้นตรง (Linear trend model):

$$y_t = \mu + \beta t + \varepsilon_t \quad , \quad t=1,2,\ldots,n$$

โดยที่ μ , β แทนระดับและความชัน (slope) ของข้อมูลตามลำดับ

เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้ง จะคำนวณหาค่าของการทำให้เรียบแบบเอกซโพเนนเชียล 2 อันดับ คือ ค่าของการทำให้เรียบแบบเอกซโพเนนเชียลอันดับที่หนึ่งและอันดับที่สอง ($S_t^{(1)}$ และ $S_t^{(2)}$) โดยที่ $S_t^{(1)} = S_t$ ซึ่งถูกกำหนดไว้แล้วในสมการที่ (2.6) และ $S_t^{(2)}$ ถูกกำหนดเป็น

$$S_t^{(2)} = (1-a)S_t^{(1)} + aS_{t-1}^{(2)} \quad , t = 1,2,\dots,n \qquad \dots(2.16)$$

การประมาณค่าพารามิเตอร์ μ และ β ในรูปแบบของอนุกรมเวลา สามารถเขียนอยู่ในรูปฟังก์ชันของค่าของการทำให้เรียบแบบเอกซโพเนนเชียลทั้งสองอันดับดังนี้

$$\hat{\mu}_t = 2S_t^{(1)} - S_t^{(2)} \qquad \dots(2.17)$$

$$\hat{\beta}_t = [(1-a)/a](S_t^{(1)} - S_t^{(2)}) \qquad \dots(2.18)$$

สมการพยากรณ์สำหรับหาค่าพยากรณ์ล่วงหน้า τ คาบเวลาจากข้อมูล n ค่า โดยใช้เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้งนี้ คำนวณได้จากสูตร

$$\hat{y}_n(\tau) = \hat{\mu}_n + \tau\, \hat{\beta}_n \qquad \dots(2.19)$$

$$= [2 + \tau(1-a)/a]\, S_n^{(1)} - [1 + \tau(1-a)/a]\, S_n^{(2)}$$

การหาค่าพยากรณ์จากสมการข้างต้น จะต้องทราบ (1) ค่าสัมประสิทธิ์ส่วนลด, a และ (2) ค่าของการทำให้เรียบแบบเอกซโพเนนเชียลที่เวลาปัจจุบัน (t=n) ; $S_n^{(1)}$ และ $S_n^{(2)}$ ซึ่งค่าดังกล่าว ได้จากการคำนวณในลักษณะซ้ำๆ กัน โดยใช้สมการที่ (2.6) และ (2.16) ซึ่งจะเริ่มทำการคำนวณจากเวลา t=1 ไปจนกระทั่ง t=n จึงจะได้ค่า $S_n^{(1)}$ และ $S_n^{(2)}$ เนื่องจากที่เวลา t=1 จะต้องกำหนดค่าเริ่มต้น $S_o^{(1)}$ และ $S_o^{(2)}$ จึงจะเริ่มต้นทำการคำนวณได้ ดังนั้นผู้วิเคราะห์ควรกำหนดค่าเริ่มต้นขึ้นมาก่อน สำหรับการวิจัยครั้งนี้ จะใช้ข้อมูลตัวแรกเป็นค่าเริ่มต้น ; $S_o^{(1)} = S_o^{(2)} = y_1$

การพยากรณ์โดยใช้เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้งนี้ นอกจากจะมีปัญหาในการกำหนดค่าเริ่มต้นแล้ว ยังมีปัญหาในการกำหนดค่าสัมประสิทธิ์ส่วนลด a ว่าควรเป็นเท่าใด Brown (1962) แนะนำว่า ควรจะเลือก a ที่มีค่าอยู่ระหว่าง $\sqrt{0.70} = 0.84$ และ $\sqrt{0.95} = 0.97$ หรืออาจจะเลือกค่า a จากการสร้างค่าความคลาดเคลื่อนในการพยากรณ์ล่วงหน้า 1 คาบเวลา ณ ค่า a ต่างๆ แล้วเลือกค่า a ที่ให้ค่า MSE ต่ำสุด ไปใช้ในการพยากรณ์ต่อไป สำหรับการวิจัยครั้งนี้ จะเลือกใช้ค่า a ที่ให้ค่า MSE ต่ำที่สุด

นอกจากจะใช้ค่าของการทำให้เรียบแบบเอกซโพเนนเชียลอันดับที่หนึ่งและอันดับที่สอง ($S_t^{(1)}$ และ $S_t^{(2)}$) ประมาณค่า μ และ β ด้วยสมการที่ (2.17) และ (2.18) แล้ว ยังสามารถประมาณค่าดังกล่าวด้วยวิธีของโฮลท์ (Holt's method) ซึ่งเป็นการทำให้เรียบแบบเอกซโพเนนเชียลที่ใช้สำหรับพยากรณ์ข้อมูลอนุกรมเวลาที่มีแนวโน้มเป็นเส้นตรงอีกวิธีหนึ่ง วิธีของโฮลท์นี้ จะใช้ค่าสัมประสิทธิ์ส่วนลด 2 ค่า (a_1 และ a_2) สำหรับประมาณพารามิเตอร์ μ และ β แยกกันคนละค่า ดังสูตรต่อไปนี้

$$\hat{\mu}_t = (1-a_1)y_t + a_1(\hat{\mu}_{t-1} + \hat{\beta}_{t-1})$$

$$= (1-a_1)y_t + a_1\hat{y}_{t-1}(1) \qquad \ldots(2.20)$$

$$\hat{\beta}_t = (1-a_2)(\hat{\mu}_t - \hat{\mu}_{t-1}) + a_2\hat{\beta}_{t-1} \qquad \ldots(2.21)$$

จากสมการที่ (2.20) และ (2.21) ถ้า $a_1 = a^2$ และ $a_2 = 2a/(1+a)$ วิธีของโฮลท์ จะเหมือนกับการทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้ง เมื่อค่าสัมประสิทธิ์ส่วนลดเท่ากับ a[1] ด้วยเหตุนี้ถ้ากำหนดให้ a_1 มีค่าเท่ากับ a^2 และ a_2 มีค่าเท่ากับ $2a/(1+a)$ แล้ว ค่าของ μ และ β ที่ประมาณจากวิธีทั้งสองจะเป็นค่าเดียวกัน

---

[1]Magne Aldrin and Eivind Damsleth, "Forecasting non-seasonal time series with missing observations," Journal of Forecasting 8 (1989): 97-116.

สำหรับวิธีพยากรณ์ที่ใช้ในการวิจัยครั้งนี้ วิธีพยากรณ์ที่มีการปรับแก้ด้วยวิธีของอัลด์รินและแดมส์เลท จะประมาณค่าพารามิเตอร์ในรูปแบบของอนุกรมเวลาที่เวลา t ; μ_t และ β_t ด้วยวิธีของโฮลท์ โดยการแทนค่า $a_1 = a^2$ และ $a_2 = 2a/(1+a)$ ซึ่งเป็นการประมาณค่าพารามิเตอร์ทั้งสอง (μ_t, β_t) โดยตรง ไม่ต้องประมาณจากค่าของการทำให้เรียบแบบเอกซโพเนนเชียลอันดับที่หนึ่งและอันดับที่สอง ($S_t^{(1)}$ และ $S_t^{(2)}$) ส่วนวิธีพยากรณ์ที่มีการปรับแก้ด้วยวิธีของไรท์ และวิธีพยากรณ์ที่มีการประมาณค่าสูญหายจะประมาณค่า μ_t และ β_t จากค่า $S_t^{(1)}$ และ $S_t^{(2)}$

เมื่อข้อมูลเกิดการสูญหาย จะทำการพยากรณ์โดยใช้เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้ง ด้วยวิธีพยากรณ์ดังต่อไปนี้

1. <u>วิธีพยากรณ์โดยใช้เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้งที่มีการปรับแก้ด้วยวิธีของอัลด์รินและแดมส์เลท</u>

วิธีของอัลด์รินและแดมส์เลทนี้ จะไม่ประมาณข้อมูลที่สูญหาย แต่จะปรับค่าคงที่ (a) ที่ใช้กำหนดน้ำหนักในการประมาณค่าพารามิเตอร์ μ และ β ที่เวลา t = p เท่านั้น (p คือ เวลาที่ต่อจากเวลาของข้อมูลที่สูญหายตัวสุดท้าย) สำหรับการประมาณค่าพารามิเตอร์ดังกล่าว จะใช้วิธีของโฮลท์ (ใช้สมการที่ 2.20 และ 2.21 โดยแทน $a_1 = a^2$ และ $a_2 = 2a/(1+a)$) ส่วนการประมาณค่าพารามิเตอร์ทั้งสอง (μ, β) ที่เวลาถัดไป (p+1,p+2,...,n) จะกลับไปใช้ค่าคงที่ a ค่าเดิมในการกำหนดน้ำหนักให้กับข้อมูล โดยมีขั้นตอนการคำนวณหาค่าพยากรณ์ดังนี้

1.1 ประมาณค่า μ_t และ β_t ของข้อมูลในช่วงแรก (t = 1,2,..., p-k-1) จากสูตร

$$\hat{\mu}_t = (1-a^2)y_t + a^2(\hat{\mu}_{t-1} + \hat{\beta}_{t-1}) \qquad ...(2.22)$$

$$\hat{\beta}_t = [1-2a/(1+a)](\hat{\mu}_t - \hat{\mu}_{t-1}) + [2a/(1+a)]\hat{\beta}_{t-1} \qquad ...(2.23)$$

โดยที่เวลา t=1 จะกำหนดค่าเริ่มต้น $\mu_o = y_1$ และ $\beta_o = 0$*

1.2 ประมาณค่า μ_t และ β_t ของข้อมูลหลังช่วงที่มีข้อมูลสูญหาย (t = p,p+1,p+2,...,n) ดังนี้

1.2.1 ที่เวลา t = p จะประมาณค่าดังกล่าวจากสูตร

$$\hat{\mu}_p = (1-a^*)y_p + a^*(\hat{\mu}_{p-k-1} + (k+1)\hat{\beta}_{p-k-1}) \qquad ...(2.24)$$

$$\hat{\beta}_p = (1-b^*)(\hat{\mu}_p - \hat{\mu}_{p-k-1})/(k+1) + b^*\hat{\beta}_{p-k-1} \qquad ...(2.25)$$

โดยที่

$$a^* = \frac{a^2}{1+(1-a^2)^2k[1+(1-\frac{2a}{1+a})(k+1)(1+(1-\frac{2a}{1+a})(\frac{2k+1}{6}))]} \qquad ...(2.26)$$

$$b^* = \frac{1-(1-\frac{2a}{1+a})(k+1)\{1+(1-a^2)k[1+(1-\frac{2a}{1+a})(\frac{k+1}{2})]\}}{1+(1-a^2)k[1+(1-\frac{2a}{1+a})(k+1)(1+(1-\frac{2a}{1+a})(\frac{2k+1}{6}))]} \qquad ...(2.27)$$

1.2.2 ที่เวลา t = p+1,p+2,...,n จะประมาณค่า μ_t และ β_t จากสมการที่ (2.22) และ (2.23) เช่นเดียวกับข้อมูลในช่วงแรก

1.3 หาค่าพยากรณ์จากสูตร

$$\hat{y}_n(\tau) = \hat{\mu}_n + \tau\,\hat{\beta}_n \quad , \tau = 1,2,... \qquad ...(2.28)$$

---

*การกำหนดค่าเริ่มต้นดังกล่าว เนื่องจากกำหนด $S_o^{(1)} = S_o^{(2)} = y_1$ และเนื่องจาก $\mu_t = 2\,S_t^{(1)} - S_t^{(2)}$ และ $\beta_t = [a/(1-a)](S_t^{(1)} - S_t^{(2)})$

2. วิธีพยากรณ์โดยใช้เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้งที่มีการปรับแก้ด้วยวิธีของไรท์

การปรับแก้ด้วยวิธีของไรท์นี้ จะไม่ประมาณค่าข้อมูลที่สูญหาย แต่จะปรับค่าคงที่ a ที่ใช้กำหนดน้ำหนักให้กับข้อมูลหลังช่วงที่มีข้อมูลสูญหายทุกคาบเวลา ($t=p+1, p+2, \ldots, n$) และจะประมาณค่าพารามิเตอร์ μ และ β ที่เวลาต่างๆ จากค่าของการทำให้เรียบแบบเอกซโพเนนเชียลอันดับที่หนึ่งและอันดับที่สอง ($S_t^{(1)}$ และ $S_t^{(2)}$) โดยมีขั้นตอนการคำนวณหาค่าพยากรณ์ดังนี้

2.1 คำนวณหาค่า $S_t^{(1)}$ และ $S_t^{(2)}$ พร้อมทั้งประมาณค่า μ_t และ β_t ของข้อมูลช่วงแรก ($t = 1,2,\ldots,p-k-1$) จากสูตร

$$S_t^{(1)} = (1-a)y_t + aS_{t-1}^{(1)} \qquad \ldots(2.29)$$

$$S_t^{(2)} = (1-a)S_t^{(1)} + aS_{t-1}^{(2)} \qquad \ldots(2.30)$$

$$\hat{\mu}_t = 2S_t^{(1)} - S_t^{(2)} \qquad \ldots(2.31)$$

$$\hat{\beta}_t = [(1-a)/a](S_t^{(1)} - S_t^{(2)}) \qquad \ldots(2.32)$$

2.2 คำนวณหาค่า $S_t^{(1)}$ และ $S_t^{(2)}$ ของข้อมูลหลังช่วงข้อมูลสูญหาย ($t = p,p+1,p+2,\ldots,n$) ดังนี้

2.2.1 ที่เวลา $t = p$ จะคำนวณหาค่า $S_p^{(1)}$ และ $S_p^{(2)}$ จากสูตร

$$S_p^{(1)} = (1-a_p)y_p + a_p\, S_{p-k-1}^{(1)} \qquad \ldots(2.33)$$

$$S_p^{(2)} = (1-a_p)S_p^{(1)} + a_p\, S_{p-k-1}^{(2)} \qquad \ldots(2.34)$$

โดยที่ $a_p = w_p/(w_p+1-a_{p-k-1})$, $w_p = a^{k+2}$ และ k = จำนวนข้อมูลที่สูญหาย

สำหรับค่า μ_p และ β_p จะประมาณจากสูตร

$$\hat{\mu}_p = 2S_p^{(1)} - S_p^{(2)} \qquad ...(2.35)$$

$$\hat{\beta}_p = [1/Q_p(1-a_p)](S_p^{(1)} - S_p^{(2)}) \qquad ...(2.36)$$

เมื่อ $Q_p = a_p Q_{p-k-1} + (k+1)[a_p/(1-a_p)]$

และค่าเริ่มต้นสำหรับ Q_o หาได้จาก $Q_o = s.a^{s+1}/(1-a^{s+1})^2$, $s = (n-k-2)/2$

2.2.2 ที่เวลา $t = p+1, p+2, ..., n$ จะคำนวณหาค่า $S_t^{(1)}$ และ $S_t^{(2)}$ จากสูตร

$$S_t^{(1)} = (1-a_t)y_t + a_t S_{t-1}^{(1)} \qquad ...(2.37)$$

$$S_t^{(2)} = (1-a_t)S_t^{(1)} + a_t S_{t-1}^{(2)} \qquad ...(2.38)$$

โดยที่ $a_t = w_t/(w_t+1-a_{t-1})$, $w_t = a^2$

และจะประมาณค่า μ_t และ β_t จากสูตร

$$\hat{\mu}_t = 2S_t^{(1)} - S_t^{(2)} \qquad ...(2.39)$$

$$\hat{\beta}_t = [1/Q_t(1-a_t)](S_t^{(1)} - S_t^{(2)}) \qquad ...(2.40)$$

เมื่อ $Q_t = a_t Q_{t-1} + a_t/(1-a_t)$

2.3 หาค่าพยากรณ์จากสมการที่ (2.28)

3. วิธีพยากรณ์โดยใช้เทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลซ้ำสองครั้งที่มีการประมาณข้อมูลสูญหาย

วิธีนี้จะคล้ายกับวิธีที่ใช้กับเทคนิคการทำให้เรียบแบบเอกซโพเนนเชียลครั้งเดียว กล่าวคือ จะประมาณข้อมูลในช่วงที่สูญหายด้วยค่าพยากรณ์ จากสมการพยากรณ์ที่เวลา $t=p-k-1$ ซึ่งเป็นเวลาล่าสุดก่อนที่จะมีข้อมูลสูญหาย โดยมีขั้นตอนการคำนวณหาค่าพยากรณ์ดังนี้

3.1 คำนวณหาค่า $S_t^{(1)}$ และ $S_t^{(2)}$ พร้อมทั้งประมาณค่า μ_t และ β_t ของข้อมูลในช่วงแรก ($t=1,2,\ldots,p-k-1$) ด้วยสมการที่ (2.29), (2.30), (2.31) และ (2.32) เช่นเดียวกับวิธีพยากรณ์ที่มีการปรับแก้ด้วยวิธีของไรท์

3.2 คำนวณหาสมการพยากรณ์ที่เวลา $t = p-k-1$ จากสูตร

$$\hat{y}_{p-k-1}(\tau) = \hat{\mu}_{p-k-1} + \tau\,\hat{\beta}_{p-k-1} \qquad , \tau = 1,2,\ldots,k$$

จากนั้นจะนำสมการดังกล่าว ไปใช้สำหรับประมาณค่าของข้อมูลที่สูญหายทุกๆ ช่วงเวลา จนได้ข้อมูลครบถ้วน

3.3 คำนวณหาค่า $S_t^{(1)}$ และ $S_t^{(2)}$ พร้อมทั้งประมาณค่า μ_t และ β_t ของข้อมูลในช่วงที่สูญหายและหลังช่วงที่สูญหาย ($t = p-k,p-k+1,\ldots,p-1,p,p+1,\ldots,n$) จากสมการที่ (2.29), (2.30), (2.31) และ (2.32) เช่นเดียวกับข้อมูลในช่วงแรก

3.4 หาค่าพยากรณ์จากสมการที่ (2.28)

2.2 เกณฑ์ในการเปรียบเทียบความสามารถของวิธีพยากรณ์

การเปรียบเทียบว่าวิธีพยากรณ์ใด มีความคลาดเคลื่อนของการพยากรณ์ต่ำกว่ากันนั้น จะพิจารณาจากค่าความคลาดเคลื่อนระหว่างค่าพยากรณ์กับค่าจริงในรูปของ ค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง (The squares root of mean square error : RMSE) วิธีที่ให้ค่า RMSE ต่ำกว่า จะเป็นวิธีพยากรณ์ที่มีความคลาดเคลื่อนโดยเฉลี่ยต่ำกว่า สำหรับค่า RMSE ที่เกิดจากการพยากรณ์ ณ คาบเวลา t ($RMSE_t$) คำนวณได้จากสูตรต่อไปนี้

$$RMSE_t = \sqrt{\frac{\sum_{i=1}^{1000} (y_{ti} - \hat{y}_{ti})^2}{1000}} \quad ; t = n+1, n+2, ..., n+12$$

โดยที่ y_{ti} คือ ค่าสังเกต ณ เวลา t ในการทำซ้ำรอบที่ i

$\hat{y}_{ti}$ คือ ค่าพยากรณ์ ณ เวลา t ในการทำซ้ำรอบที่ i

i คือ จำนวนรอบของการทำซ้ำ ; $i = 1, 2, ..., 1000$

t คือ คาบเวลาของการพยากรณ์