บทที่ 11
อนุกรมเวลา-เส้นแนวโน้ม
(TIME SERIES-TREND LINE)

11.1 คำนำ

งานสำคัญที่เกี่ยวกับธุรกิจหรือองค์การหนึ่งองค์การใดจะต้องพบอยู่เสมอคือการวางแผนเพื่ออนาคต นี้เป็นหลักที่ต้องการความสามารถของคนหนึ่งคนใดที่ดีที่สุด อย่างไรก็ตาม ถ้ามีการดำเนินการวางแผนเกี่ยวกับธุรกิจหรือเศรษฐกิจ ผู้ที่รับผิดชอบจะต้องเป็นคนมองการไกลและทำนายระดับอนาคตของกิจกรรมภายใต้แผนการอันเฉียบแหลมเท่าที่เกี่ยวข้องและจำเป็น ไม่ว่านักธุรกิจหรือข้าราชการทุกระดับจะต้องวางแผนทั้งระยะสั้นและระยะยาวของค่าใช้จ่าย สิ่งเหล่านี้นักสถิติเศรษฐศาสตร์ ได้อธิบายออกมาเป็นตัวเลข และคำนวณข้อมูลชุดนี้ของคาบเวลาต่าง ๆ กัน เรียกว่า "อนุกรมเวลา"

ผลจากการวิเคราะห์อนุกรมเวลาพบว่ามีส่วนประกอบเป็น 4 หัวข้อดังนี้

(1) แนวโน้มระยะยาว *(Secular trend)*

(2) การเปลี่ยนแปลงตามฤดูกาล (Seasonal variations)

(3) การเปลี่ยนแปลงแบบวัฏจักร (Cyclical fluctuations)

(4) การเปลี่ยนแปลงแบบไม่สม่ำเสมอ (Irregular variations)

11.1.1 แนวโน้มระยะยาว

เมื่อเราพูดถึงแนวโน้มระยะยาวของอนุกรมเวลา โดยธรรมดาแล้วเราหมายถึงการเปลี่ยนแปลงที่เป็นไปโดยสม่ำเสมอและแน่นอนในกาลระยะยาว บางอนุกรมของข้อมูลแสดงแนวโน้มในทางสูงขึ้น บางอนุกรมแสดงแนวโน้มต่ำลง อย่างเช่น ผลิตภัณฑ์ทางอุตสาหกรรม แสดงแนวโน้มในทางสูงขึ้น ขณะเดียวกันอัตราการตายของคนด้วยโรคต่าง ๆ แสดงแนวโน้มต่ำลง

11.1.2 การเปลี่ยนแปลงตามฤดูกาล

การเปลี่ยนแปลงตามฤดูกาลประกอบด้วยแบบอย่างที่เกิดขึ้นซ้ำ ๆ กันอย่างสม่ำเสมอ ชนิดของการเปลี่ยนแปลงก็เป็นไปตามฤดูกาลของแต่ละปี หรือการเปลี่ยนแปลงไปตามเดือน

สัปดาห์ วันและชั่วโมง การเปลี่ยนแปลงแต่ละชนิดก็เป็นไปตามระยะเวลา ธรรมชาติ และระยะเวลานี้ก็ไม่ใช่ระยะยาวเกินกว่าหนึ่งปี อย่างเช่น ผู้โดยสารรถเมล์หรือการจราจรติดขัดจะมีคนหนาแน่นบนรถเมล์ หรือเวลาที่จราจรติดขัดมากในตอนเช้าหรือตอนเย็น ในระหว่างนั้นจะเบาบางลงไปมาก นี่แสดงถึงการเปลี่ยนแปลงระยะสั้น คือเป็นชั่วโมง ส่วนการเปลี่ยนแปลงตามฤดูกาลหรือเป็นเดือนนั้นได้แก่ราคาของผลไม้ หรือผลิตภัณฑ์ของเกษตรบางชนิด ราคาจะสูงขึ้นเมื่อหมดหน้าฤดูของผลไม้ชนิดนั้น ๆ หรือหน้าผลไม้หรือผลิตภัณฑ์ขาดแคลน ร้านจำหน่ายของขวัญก็ดี จะขายดีก็ช่วงระยะเวลาที่มีการแต่งงานหรือปีใหม่เท่านั้น ที่จะมีโอกาสขายได้ดีกว่าในช่วงระยะเวลาอื่น ๆ

11.1.3 การเปลี่ยนแปลงแบบวัฏจักร

การหมุนเวียนทางธุรกิจไม่มีอะไรมากไปกว่าการเปลี่ยนแปลงที่คงอยู่ในอนุกรมเวลาภายหลังได้ขจัดแนวโน้ม การเปลี่ยนแปลงตามฤดูกาลและการเปลี่ยนแปลงแบบไม่สม่ำเสมอ โดยทั่วไปเราอาจกล่าวได้ว่า การหมุนเวียนทางธุรกิจประกอบด้วยการเปลี่ยนแปลงขึ้น ๆ ลง ๆ ตามกิจกรรมธุรกิจบางชนิดของแนวโน้มเชิงสถิติหรือ “ปกติ” ในที่นี้ “ปกติ” เราหมายถึงส่วนเฉลี่ยในทางสถิติบางชนิด เราไม่ได้หมายความว่าทุกสิ่งทุกอย่างจะต้องถาวรหรือมีผลทั้งหมด

เป็นที่เข้าใจกันว่าการหมุนเวียนทางธุรกิจจากการเปลี่ยนแปลงตามฤดูกาลในระยะเวลานั้นจะปกคลุมไปด้วยวัฏจักรที่ขยายในระยะเวลาอันยาวนาน ยิ่งกว่านี้ การขึ้น ๆ ลง ๆ ในการหมุนเวียนทางธุรกิจยังถูกพิจารณาถึงผลลัพธ์จากกลุ่มต่าง ๆ ของสาเหตุ ระยะเวลาของความเจริญรุ่งเรือง ความตกต่ำ ความถอยหลังและการสู่สภาพเดิม ซึ่งบางครั้งก็เป็นทรรศนะทั้ง 4 ระยะของวัฏจักรที่สมบูรณ์ ปัจจัยที่ก่อให้เกิดสาเหตุอันนี้นอกเหนือจากลมฟ้าอากาศ ขนบธรรมเนียมทางสังคมแล้วยังทำให้เกิดฤดูอีก

จากแง่คิดของการวัดวัฏจักรทางธุรกิจ ซึ่งแสดงถึงความคล้ายคลึงกันพอที่จะพิสูจน์ได้ แต่ก็เป็นที่น่าเสียดายที่บางอย่างก็ไม่แสดงถึงความคล้ายคลึงกันเลยที่จะทำนายปรากฏการณ์ในอนาคต นี้ก็เป็นหลักความจริงที่ว่าวัฏจักรก็มีชนิดต่าง ๆ กันหลายชนิด บางครั้งก็เป็นระยะสั้น บางครั้งก็เป็นระยะยาว บางครั้งก็เร็ว บางครั้งก็ช้า แต่ในปัญหาทางปฏิบัติเพื่อแยกการเคลื่อนไหวทางวัฏจักรของอนุกรมเวลาได้โดยขจัดแนวโน้ม การเปลี่ยนแปลงตามฤดูกาลและการเคลื่อนไหวแบบไม่สม่ำเสมอ

11.1.4 การเปลี่ยนแปลงแบบไม่สม่ำเสมอ

ความไม่สม่ำเสมอหรือความไม่แน่นอน การขึ้น ๆ ลง ๆ ของอนุกรมเวลาเป็นการเปลี่ยน-แปลงซึ่งไม่สามารถทำนายได้ถูกต้องสมบูรณ์นัก สาเหตุเกิดจากปรากฏการณ์โดดเดี่ยวพิเศษ อย่างเช่น น้ำท่วม แผ่นดินไหว การนัดหยุดงาน สงคราม การล้มละลายของธนาคาร ข่าวดีและข่าวเลว อิทธิพลบางอย่างซึ่งเราแยกออกเป็นความไม่แน่นอนอย่างเห็นได้ชัด ได้แก่การทำงานด้วยตนเองก่อนที่จะสำนึกถึงสิ่งหนึ่งใด

การเปลี่ยนแปลงแบบไม่สม่ำเสมอ เป็นเหตุการณ์ส่วนใหญ่เกิดขึ้นโดยบังเอิญหรือแบบสุ่ม ซึ่งไม่สามารถทำนายได้ในชั่วระยะหนึ่งแต่ในระยะยาวก็พอจะทราบได้เล็กน้อยโดยเฉลี่ย

การตรวจสอบเกือบทุก ๆ อนุกรมเวลาพบว่ามีการเปลี่ยนแปลงคงที่และที่เป็นปัญหาที่ได้ยกขึ้นมาเกี่ยวกับลักษณะธรรมชาติของอำนาจโดยเฉพาะของการทำงาน อำนาจสุ่มเราสันนิษฐานว่าเป็นวิธีการที่แน่นอน อาจใช้ป้องกันอำนาจที่ปฏิบัติแบบไม่สุ่ม นั่นคือ ระบบสมัยนิยม ตัวอย่างเช่น การเปลี่ยนแปลงตามฤดูกาลสำหรับเดือนที่แน่นอนของปี วันของสัปดาห์ หรือชั่วโมงของวันที่สูงกว่าหรือต่ำกว่าของส่วนเฉลี่ย ถ้าหากว่าความเคลื่อนไหวหรือการเปลี่ยนแปลงแบบไม่สุ่มเหล่านี้ปรากฏอยู่ในอนุกรม เราต้องป้องกันและหาสาเหตุเพื่อสร้างแบบเหล่านี้ สิ่งหนึ่งที่เราต้องทำ คือ พยากรณ์เหตุการณ์ในอนาคต ดำเนินการตามเงื่อนไขที่ว่า อำนาจที่คล้ายคลึงกันจะยังคงทำไปเรื่อย ๆ ต่องานส่วนที่เหลืออยู่

11.2 การปรับปรุงข้อมูลเบื้องต้นของอนุกรมเวลา

ก่อนเริ่มคำนวณสมการแนวโน้มจริง ๆ หรือวัดการเปลี่ยนแปลงตามฤดูกาลเป็นสิ่งจำเป็นอย่างยิ่งที่ต้องทำการปรับปรุงข้อมูลดิบ โดยปกติการพิจารณาปรับปรุงเกี่ยวกับการเปลี่ยนแปลงตามปีปฏิทิน การเปลี่ยนแปลงของราคาและการเปลี่ยนแปลงของประชากร

วัตถุประสงค์ของการปรับปรุงสำหรับการเปลี่ยนแปลงตามปฏิทิน ก็เพื่อขจัดความแตกต่างที่ปลอมแปลงซึ่งเป็นเหตุลักษณะพิเศษของปฏิทินของเรา อย่างเช่น ตัวเลขผลิตภัณฑ์สำหรับเดือน ก.พ. อาจต่ำกว่าเดือนมกราคม ไม่ใช่เพราะว่าลดลงจริงในกิจกรรม แต่เพราะว่าเดือนกุมภาพันธ์มีน้อยกว่าเท่านั้น นี้อาจเอามาเป็นตัวเลขง่าย ๆ โดยการหารยอดรวมแต่ละเดือนด้วยจำนวนวันของแต่ละเดือน (บางครั้งหารด้วยจำนวนวันที่ทำงานของเดือน) ก็จะได้ค่าเฉลี่ยรายวันของแต่ละเดือน

การปรับปรุงสำหรับการเปลี่ยนแปลงของราคาก็เป็นสิ่งจำเป็น เมื่อไรก็ตามที่เรามีอนุกรมมูลค่าและมีความสนใจในปริมาณการเปลี่ยนแปลงอย่างเดียว จำนวนที่ขายทั้งหมดจะขึ้น ๆ ลง ๆ เปลี่ยนแปลงทั้งปริมาณและราคา ถ้าราคาสูงมากจำนวนหน่วยที่ขายจะลดต่ำลง การเพิ่มขึ้นของมูลค่าเป็นเหตุมาจากราคาสูงขึ้น เราสามารถขจัดผลของการเปลี่ยนแปลงราคาได้โดยการหารแต่ละรายการในอนุกรมมูลค่าด้วยดัชนีราคาที่เหมาะสมนี้ ก็ไม่มีอะไรนอกเหนือไปกว่ากระบวนการทำให้ราคาสูงขึ้น (deflating) ซึ่งได้อธิบายในบทที่ 10 หัวข้อ (10.9)

บางครั้งก็เป็นสิ่งจำเป็นเพื่อปรับอนุกรมสำหรับการเปลี่ยนแปลงของประชากรเนื่องจากการเปรียบเทียบตัวเลขของรายได้ ผลิตภัณฑ์และบริโภคสามารถทำให้ผิดรูปไปอย่างง่ายดายโดยการเปลี่ยนแปลงในขนาดของประชากร เมื่อเห็นสมควรที่จะปรับข้อมูลสำหรับการเปลี่ยนแปลงของประชากร ข้อมูลจะแสดงออกมาในรูปต่อประชากรคนหนึ่งได้โดยการหารตัวเลขเดิมด้วยประชากรทั้งหมดที่พอเหมาะ

ความสามารถของการเปรียบเทียบก็เป็นลักษณะสำคัญอย่างหนึ่งเกี่ยวกับการปรับปรุงเบื้องต้นของอนุกรมเวลา ความสามารถของการเปรียบเทียบของข้อมูลตลอดทั้งระยะเวลาภายใต้การตรวจสอบเป็นสิ่งจำเป็นต่อการวิเคราะห์ความหมายอย่างใดอย่างหนึ่งเป็นการยากหรือไม่อาจเป็นไปได้ที่จะได้ข้อมูลเปรียบเทียบที่แท้จริง บางครั้งเราต้องการตัวเลขที่ย้อนหลังกลับไป 20, 30, 50 ปี หรือมากกว่า ตัวเลขของผลิตภัณฑ์ ค่าจ้าง บริโภค การขายอาจรวบรวมได้โดยตัวแทนต่าง ๆ ถึงแม้ว่านิยามของรายการ (เช่น ตู้เสื้อผ้า) อาจไม่เปลี่ยนแปลง อาจมีการเปลี่ยนแปลงก็เฉพาะคุณภาพซึ่งราคาที่แจ้งก็ไม่สามารถที่จะเอามาเปรียบเทียบอย่างแท้จริงได้ การเปลี่ยนแปลงจากนิยามเวลาหนึ่งถึงนิยามอีกเวลาหนึ่งของห้างร้านหรือฟาร์มของปีนี้ ก็ไม่จำเป็นที่จะต้องเหมือนกับปีที่แล้ว ผลของการเปลี่ยนแปลงเช่นนั้นในนิยามเป็นตัวเลขที่ได้บันทึกได้มากมาย ก็ไม่สามารถเอามาเปรียบเทียบได้อย่างแท้จริง

11.3 แนวโน้ม (Trend)

แนวโน้มเป็นการเปลี่ยนแปลงระยะยาวของอนุกรมเวลา ตัวอย่างเช่นแนวโน้มของการขยายรายได้ประชาชาติ การเพิ่มขึ้นของประชากร ผลิตภัณฑ์ทางอุตสาหกรรมก็เป็นการเปลี่ยนแปลงระยะยาว เป็นต้น

ชนิดต่าง ๆ ของเส้นแนวโน้มแสดงการขยายตัวออกในลักษณะฐานนิยมต่าง ๆ กัน ตัวอย่างเช่น การขยายตัวของการกำเนิดของคนงานในสหรัฐอเมริกาประมาณอยู่ในลักษณะเป็นเส้นตรง

ที่มีความชันสูงขึ้น ดังนั้นจึงอาจใช้เส้นตรงได้ สำหรับการขยายพันธุ์ของแมลงหวี่อยู่ในลักษณะเรขาคณิต บางชนิดของเส้นแนวโน้มมีลักษณะเป็นเส้นโค้ง (parabolic) เป็นลอกการิทมมิค

เราสนใจในการคำนวณหาและแสดงออกของแนวโน้มในรูปของสมการและแสดงเป็นรูปกราฟจากข้อมูลที่กำหนดให้ก็เขียนกราฟได้ แต่ปัญหาที่จะปรับเส้นตรงให้เข้ากับข้อมูลเพื่อแสดงถึงการขยายตัวทีละน้อยในระยะยาวของอนุกรมเวลา มีวิธีต่าง ๆ ในการที่จะปรับเส้นตรงด้วยกัน เช่น วิธีกะเอา วิธีแบ่งตัวเลขที่มีอยู่เป็น 2 พวกเท่า ๆ กัน แล้วหาค่าเฉลี่ยของแต่ละพวก วิธีส่วนเฉลี่ยเคลื่อนที่ และวิธีกำลังสองน้อยที่สุด จุดประสงค์ที่สำคัญของเราก็จะเน้นหนักไปเกี่ยวกับวิธีกำลังสองน้อยที่สุด ส่วนวิธีกะเอา วิธีแบ่งตัวเลขที่มีอยู่เป็น 2 พวกเท่า ๆ กัน แล้วหาค่าเฉลี่ยของแต่ละพวกและวิธีส่วนเฉลี่ยเคลื่อนที่ก็จะใช้เป็นวิธีเบื้องต้นเท่านั้น

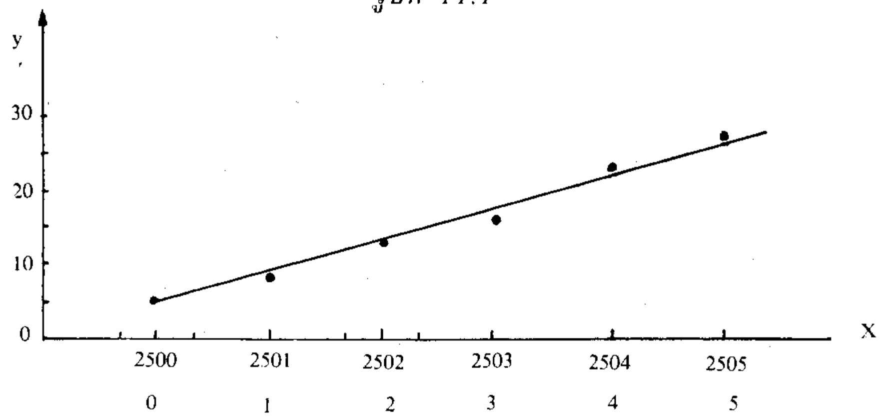
11.3.1 **วิธีกะเอา** (The Freehand Method)

วิธีง่ายที่สุดในการหาเส้นแนวโน้มเมื่อกำหนดกลุ่มของข้อมูลอนุกรมเวลาให้ คือ วิธีกะเอา กระบวนการก็คือต้องเขียนกราฟของอนุกรมเวลา โดยการสังเกต ลากเส้นตรงให้ผ่านจุดที่กะว่าแนวโน้มควรจะผ่านของอนุกรมเวลา วิธีนี้จัดว่าเป็นวิธีหาแนวโน้มที่รวดเร็วและง่ายที่สุด แต่มีข้อเสียอยู่ว่า เป็นวิธีหาแนวโน้มที่ขาดหลักเกณฑ์แน่นอน ไม่เหมาะสำหรับผู้ที่ยังไม่มีความชำนาญทางด้านนี้โดยเฉพาะ ตัวอย่างสมมุติข้อมูลดังในตาราง

ปี	x	y (ล้านกิโลกรัม)
2500	0	5
2501	1	8
2502	2	12
2503	3	15
2504	4	20
2505	5	23

ให้ y เป็นผลิตภัณฑ์รายปีของสินค้าบางชนิดมีหน่วยเป็นล้านกิโลกรัม พลอตจุดลงในรูป 11.1 เราลากเส้นโดยกะเอา แน่ละเส้นตรงที่ได้ไม่ถูกต้องแน่นอนตามข้อมูล แต่เราอาจได้แนวความคิดหยาบ ๆ ต่อเส้นตรงหรือเส้นโค้งว่าควรจะปรับหรือไม่ก่อนวิธีแก้ไข ในกรณีเช่นนั้นเส้นก็อาจช่วยให้เป็นจริงได้บ้าง

รูปที่ 11.1



11.3.2 วิธีแบ่งตัวเลขที่มีอยู่เป็น 2 พวกเท่า ๆ กัน แล้วหาค่าเฉลี่ยของแต่ละพวก (Method of Semiaverages)

วิธีการแบบนี้แบ่งอนุกรมเวลาเป็นสองส่วนแล้วหาค่าเฉลี่ยของแต่ละส่วน ต่อไปลากเส้นตรงผ่านค่าเฉลี่ยเหล่านี้ ตัวอย่าง ใช้ข้อมูลของตัวอย่างในหัวข้อ (11.3.1)

ปี	x	y	
2500	0	5	
2501	1	8	$\frac{25}{3} = 8.3$
2502	2	12	
2503	3	15	
2504	4	20	$\frac{58}{3} = 19.3$
2505	5	23	

ค่าเฉลี่ยของแต่ละส่วนคือ 8.3 (8,300,000 กก.) กับ 19.3 (19,300,000 กก.) เนื่องจาก 8.3 เป็นค่าเฉลี่ยของปี 2500, 2501 และ 2502 พลอตจุด 8.3 ที่ปี 2501 (ดูรูป 11.2) เช่นเดียวกันก็พลอตจุด 19.3 ที่ปี 2504 เส้นตรงก็จะผ่านจุด (1, 8.3) กับจุด (4, 19.3) ก็คือ เส้นแนวโน้มกึ่งเฉลี่ย (Semiaverages trendline) ที่ต้องการ เราคำนวณหาเส้นแนวโน้มได้ดังนี้

ให้สมการของเส้นตรงเป็น $\hat{y} = a + bx$

เราต้องการคำนวณหาค่า a และ b ได้โดยแทนค่าจุด (1, 8.3) กับจุด (4, 19.3) ลงในสมการของเส้นตรงได้

$$8.3 \quad a + b(1)$$

$$19.3 = a + b(4)$$

จากสมการข้างบนทั้งสองเราก็สามารถคำนวณหาค่า a ได้เท่ากับ 4.6 กับ b ได้เท่ากับ 3.7 ดังนั้นสมการของเส้นแนวโน้มได้

$$\hat{y} = 4.6 + 3.7x$$

x มีหน่วยเป็น 1 ปี

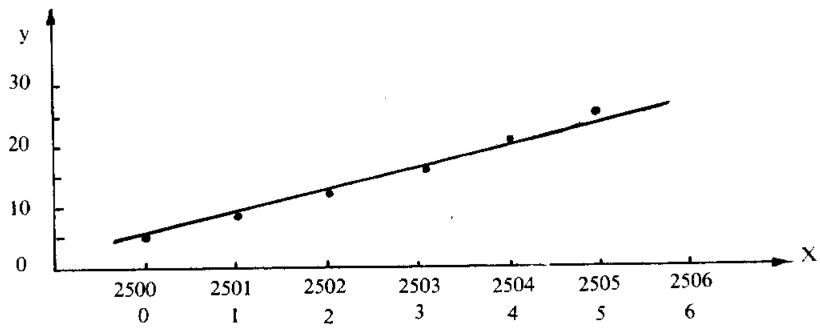
ค่าแนวโน้มของผลิตภัณฑ์สำหรับ x = 0 (นั่นคือ ปี 2500) เป็น

$$\hat{y} = 4.6 + 3.8(0) = 4.6$$

แต่ค่าผลิตภัณฑ์จริงคือ $y_{2500} = 5$ ค่าจริงกับค่าแนวโน้มของผลิตภัณฑ์ต่างกัน

$$\hat{y} - y_{2500} \quad 4.6 - 5 = -0.4$$

b = 3.7 แสดงว่าค่าแนวโน้มของผลิตภัณฑ์เพิ่มขึ้นรายปีเท่ากับ 3,700,000 กก.



รูปที่ 11.2

เมื่อมีจำนวนปีเป็นเลขคี่ อนุกรมก็ไม่สามารถแบ่งออกให้เท่ากันได้ ดังนั้นปีกลางก็ไม่มี หรืออนุกรมก็ถูกแบ่งออกเป็นสองส่วนไม่เท่ากัน

ถ้าเกิดมีค่าหนึ่งค่าใดของอนุกรมมีค่าสูงมากก็อาจมีอิทธิพลต่อกึ่งเฉลี่ยหรือเส้นแนวโน้มเสียรูปไป ในกรณีเช่นนั้น จึงต้องละทิ้งค่านั้นเสีย อย่างเช่น ต้องการเขียนเส้นแนวโน้มของผลิตภัณฑ์เหล็กกล้า เกิดมีปีหนึ่งคนงานนัดกันหยุดงานเป็นเหตุให้ผลิตภัณฑ์เหล็กกล้าได้น้อยมาก ในกรณีเช่นนั้น ผลิตภัณฑ์เหล็กกล้าของปีนั้นก็ต้องละทิ้งไป

นี่เป็นวิธีการเขียนเส้นแนวโน้มอย่างหยาบและง่ายซึ่งเป็นข้อดีของวิธีการนี้

11.3.3 วิธีส่วนเฉลี่ยเคลื่อนที่ (Method of Moving Averages)

วิธีการส่วนเฉลี่ยเคลื่อนที่เป็นวิธีการที่ใช้เวลาการขึ้น ๆ ลง ๆ ในอนุกรมเวลาและใช้ไม่แต่เส้นแนวโน้มเท่านั้น แต่ยังใช้กับการเปลี่ยนแปลงตามฤดูกาลและแบบวัฏจักร ตัวอย่างเช่น สมมติว่าข้อมูลที่กำหนดเกี่ยวกับการขายสินค้าในรูปตาราข้างล่าง

เริ่มแรกหาผลบวกเคลื่อนที่ 3 ปี เช่น สำหรับปี 2500, 2501 และ 2502 เราได้

$$3 + 4 + 8 = 15$$

ค่า 15 นี้ใส่ลงให้ตรงกับปีกลาง 2501 ผลบวกของ 3 ปี ต่อไปสำหรับ 2501, 2502 และ 2502 ก็จะได้

$$4 + 8 + 6 = 18$$

ปี	จำนวนเงินที่ขาย (ล้านบาท)	ผลบวกของ 3 ปีเคลื่อนที่	ค่าเฉลี่ยผลบวกของ 3 ปีเคลื่อนที่
2500	3		
2501	4	15	15/3 = 5
2502	8	18	18/3 = 6
2503	6	21	21/3 = 7
2504	7	24	24/3 = 8
2505	11	27	27/3 = 9
2506	9	30	30/3 = 10
2507	10	33	33/3 = 11
2508	14	36	36/3 = 12
2509	12		

ค่าที่ใส่ลงให้ตรงกับปีกลาง 2502 กระบวนการนี้ทำต่อไปจนหมด เราพบว่าคอลัมน์ของผลบวกเคลื่อนที่ 3 ปี จะไม่มีผลบวกสำหรับปี 2500 กับปี 2509

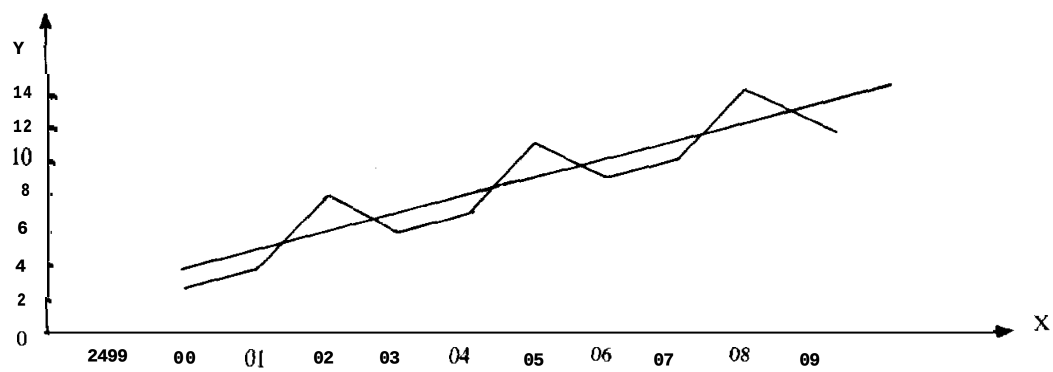
ต่อไปก็คำนวณหาส่วนเฉลี่ยเคลื่อนที่ 3 ปี โดยการหารผลบวกเคลื่อนที่ด้วย 3 ก็จะได้ส่วนเฉลี่ยเคลื่อนที่ 3 ปี ในคอลัมน์ที่ 4

กราฟแสดงถึงจำนวนเงินที่ขายของวัฏจักร 3 ปี สม่ำเสมอ ดังตัวอย่าง ยอดแหลมของกราฟอยู่ที่ปี 2502 ปี 2505 และปี 2508 ในช่วงระหว่าง 3 ปี เมื่อพลอตค่าเฉลี่ยเคลื่อนที่ 3 ปี ลงบนแผ่นกราฟดังรูป 11.3 จะได้เส้นตรง ส่วนการขึ้น ๆ ลง ๆ แบบวัฏจักรก็จะถูกเกลาออกไป เส้นตรงก็คือเส้นแนวโน้มที่เราหาได้

ภายหลังที่ได้พลอตจุดค่าเฉลี่ยเคลื่อนที่ 3 ปี เราจะมีปัญหาของการลากเส้นแนวโน้มของจุดเหล่านี้ วิธีกะเอา วิธีกึ่งเฉลี่ย วิธีกำลังสองน้อยที่สุด หรือวิธีอื่น ๆ อาจใช้ลากเส้นแนวโน้มของจุดค่าเฉลี่ยเคลื่อนที่เหล่านี้

ทำไมเราจึงได้เส้นโค้งที่เรียบ ผลลัพธ์ก็เพราะว่าข้อมูลที่อยู่ในวัฏจักรที่สม่ำเสมอและอยู่ช่วงระหว่างเวลา มี amplitude เหมือนกัน ข้อสังเกตว่าวัฏจักรก็คือ ช่วงระหว่าง 3 ปี เราก็เลือกค่าเฉลี่ยเคลื่อนที่ 3 ปี ถ้าช่วงระหว่างเวลาคือ 4 ปี เราก็ควรเลือกค่าเฉลี่ยเคลื่อนที่ 4 ปี ครึ่งหนึ่งของวัฏจักรจะอยู่เหนือจุดกลางและอีกครึ่งหนึ่งจะอยู่ต่ำกว่า เมื่อได้ค่าเฉลี่ย ผลก็จะหักลบกันหมด และถ้าหากว่าครึ่งของส่วนที่อยู่เหนือจุดกลางของวัฏจักรมีค่ามากกว่าครึ่งของส่วนล่างดังตัวอย่างของเรา ส่วนเฉลี่ยเคลื่อนที่ก็จะแสดงแนวโน้มสูงขึ้น

ถ้าต้องการเอาส่วนเฉลี่ยเคลื่อนที่ไปประยุกต์ให้ได้ผลอย่างเต็มที่ เริ่มแรกก็จำเป็นจะต้องกำหนดว่าเป็นวัฏจักรที่สม่ำเสมอมีระยะเท่ากันจริงหรือไม่ ในกรณีทางปฏิบัติในเมื่อวัฏจักรไม่เป็นจริง ช่วงระหว่างเวลาของวัฏจักรโดยทั่วไปไม่มีความสม่ำเสมอมาก แต่ก็มีจำนวนมากที่มีความสม่ำเสมอพอที่จะอนุโลมให้ใช้วิธีส่วนเฉลี่ยเคลื่อนที่ได้



รูปที่ 11.3

จะเห็นได้ว่าเส้นแนวโน้มเราคำนวณหาได้ออกมาเป็นเส้นตรงถ้าหากว่าลักษณะพื้นฐานของอนุกรมเวลาเป็นเส้นตรง แล้วเราก็ได้เส้นตรง ถ้าหากว่าลักษณะพื้นฐานเป็นเส้นโค้ง แนวโน้มก็จะได้ลักษณะเป็นเส้นโค้ง วิธีส่วนเฉลี่ยเคลื่อนที่ถูกนำไปใช้ได้ไม่แต่เส้นแนวโน้มเท่านั้น แต่ยังใช้กับข้อมูลทุกชนิดที่แสดงการขึ้น ๆ ลง ๆ ของระยะเวลาที่เท่ากันอย่างสม่ำเสมอ

11.3.4 วิธีกำลังสองน้อยที่สุด (Method of Least Squares)

วิธีกำลังสองน้อยที่สุดเป็นวิธีการที่ใช้กันกว้างขวางที่สุดของการกำหนดเส้นตรงของข้อมูลอนุกรม เพราะเส้นแนวโน้มที่ได้จะเป็นเส้นที่เรียกว่าเป็น best fit หรืออีกนัยหนึ่งมีคุณสมบัติดังนี้

ก. ส่วนเบี่ยงเบนของค่าข้อมูลซึ่งวัดตามแนวตั้งฉากเส้นแนวโน้มที่ได้รวมกันจะเท่ากับ 0

ข. ผลบวกกำลังสองของส่วนเบี่ยงเบนตามข้อ ก. ข้างต้นมีค่าน้อยที่สุด

ถ้าใช้สัญลักษณ์ x เป็นตัวแปรอิสระแทนปีหรือหน่วยเวลาอื่น ๆ y เป็นตัวแปรตามแทนค่าของข้อมูลในอนุกรม และ $\hat{y}$ แทนค่าแนวโน้ม เราก็สรุปได้ว่า

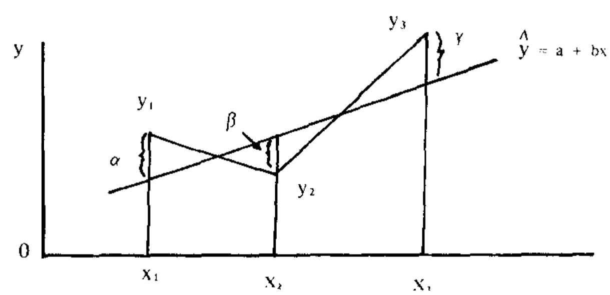
ก. $\Sigma(y - \hat{y}) = 0$

ข. $\Sigma(y - \hat{y})^2 =$ น้อยที่สุด

1. หลักการ

ดังรูป 11.4 $\hat{y}$ เป็นเส้นแนวโน้มที่คำนวณได้และ α, β และ γ เป็นส่วนเบี่ยงเบนของข้อมูลไปจากเส้นแนวโน้ม วิธีกำลังสองน้อยที่สุดเป็นวิธีซึ่งเราสามารถปรับเส้นแนวโน้ม $\hat{y}$ ให้เข้ากับข้อมูลจริงเพื่อว่าผลบวกกำลังสองของส่วนเบี่ยงเบนจะมีค่าน้อยที่สุด นั่นคือ

$$\alpha^2 + \beta^2 + \gamma^2 = \text{มีค่าน้อยที่สุด}$$



รูปที่ 11.4

แนวโน้มของอนุกรมเป็นเส้นตรง แนวโน้มนี้ก็จะเขียนได้ด้วยสมการ

$$\hat{y} = a + bx$$

ในเมื่อ a และ b เป็นค่าคงที่ เราสามารถคำนวณหาค่า a และ b ได้จากสมการปกติ

$$\Sigma y = na + b\Sigma x$$
$$\Sigma xy = a\Sigma x + b\Sigma x^2 \quad ....(11.3.1)$$

2. **จำนวนของปีเป็นเลขคี่** (Odd number of years)

สมมุติว่าเรามีข้อมูลของผลิตภัณฑ์น้ำมันปิโตรเลี่ยมรายปีมีหน่วยเป็นบาร์เรลเราประสงค์ที่จะกำหนดเส้นตรงของแนวโน้มโดยวิธีกำลังสองน้อยที่สุด ค่า a และ b คำนวณหาได้จากสูตร (11.3.1) x เป็นค่าของแกนนอน คือ ปีหนึ่ง แต่แทนที่จะใช้ปีจริง ๆ เราอาจเรียงใหม่เป็น ปี 1, 2, 3, ... หรือ 0, 1, 2, ... ได้ แต่วิธีง่ายกว่านี้ก็คือกำหนดให้ปีกลางเป็น 0 ปี ก่อนปีกลางเป็นลบและหลังปีกลางเป็นบวก เพื่อให้ผลบวกของปี (Σx) เป็นศูนย์ เพื่อความสะดวกในการคำนวณ

ปี	x	Y	xy	x^2
2500	- 2	5	- 10	4
2501	- 1	8	- 8	1
2502	0	12	0	0
2503	1	15	15	1
2504	2	20	40	4
	$\Sigma x = 0$	$\Sigma y = 60$	$\Sigma xy = 37$	$\Sigma x^2 = 10$

เราพบว่าเราต้องการ Σx, Σy, Σxy และ Σx^2 เพื่อนำไปคำนวณหาค่า a และ b ค่าเหล่านี้พบได้จากตารางที่ได้สร้างขึ้น ค่าของ x ในตารางเราออกแบบให้ $\Sigma x = 0$ เหตุผลก็เพื่อให้สะดวกในการคำนวณหาค่า a และ b จาก (11.3.1) จากตารางเราได้ $\Sigma x = 0$, $\Sigma y = 60$, $\Sigma xy = 37$, $\Sigma x^2 = 10$

n คือจำนวนของปี นั่นคือ $n = 5$ แทนค่าเหล่านี้ลงในสูตร (11.3.1) เราได้

$$60 = (5)(a) + b(0)$$
$$37 = a(0) + b(10)$$

a และ b คือ

$$a = \frac{60}{5} = 12$$

$$b = \frac{37}{10} = 3.7$$

ดังนั้น สมการสำหรับเส้นแนวโน้มคือ

$$\hat{y} = 12 + 3.7x \text{ (x มีหน่วยเป็น 1 ปี)} \qquad ....(11.3.2)$$

เพื่อที่จะให้คำนวณได้รวดเร็วขึ้นและง่ายเข้าในกรณี $\Sigma x = 0$ ได้ดังนี้

สมการปกติ

$$\Sigma y = na + b\Sigma x$$

$$\Sigma xy = a\Sigma x + b\Sigma x^2$$

เนื่องจาก $\Sigma x = 0$ สมการนี้ก็กลายเป็น

$$\Sigma y = na$$

$$\Sigma xy = b\Sigma x^2$$

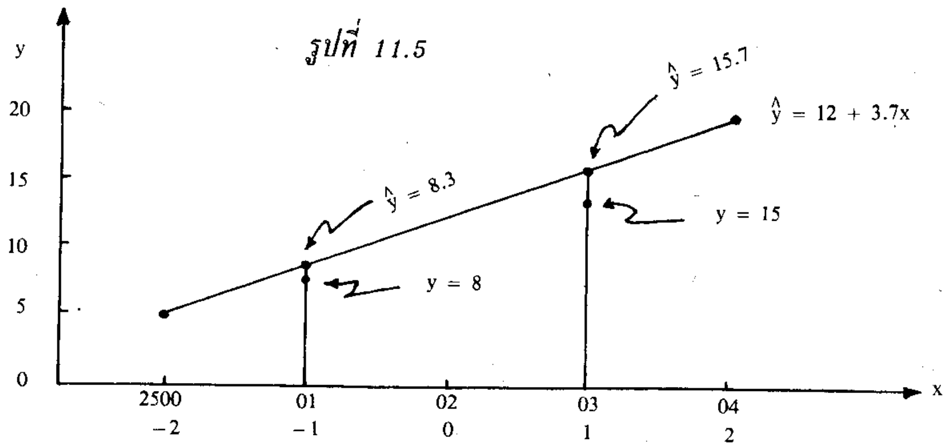
a และ b ก็คำนวณหาได้โดย

$$a = \frac{\Sigma y}{n} \text{ หรือ } \bar{y}$$

$$b = \frac{\Sigma xy}{\Sigma x^2} \qquad ....(11.3.3)$$

เราพลอตข้อมูลและเส้นแนวโน้มดังรูปที่ 11.5

รูปที่ 11.5



พลอตเส้นแนวโน้มจากสมการ (11.3.2) ค่าใด ๆ สองค่าของ $\hat{y}$ ก็คำนวณได้และเส้นตรงก็จะผ่านสองจุดนั้น ตัวอย่าง

$$x = -1 : \hat{y} = 12 + (3.7)(-1) = 8.3$$

$$x = +1 : \hat{y} = 12 + (3.7)(1) = 15.7$$

แล้วพลอตสองจุด (−1, 8.3) กับ (1, 15.7) และลากเส้นตรงผ่านก็จะได้เส้นแนวโน้มตามความต้องการ ค่า y เป็นผลรวมรายปี การแปลความของสมการได้ดังนี้ b = 3.7 คือการเปลี่ยน-แปลงรายปีโดยประมาณในผลิตภัณฑ์น้ำมันปิโตรเลียม มีหน่วยเป็นล้านบาร์เรล ค่าแนวโน้มหรือค่าผลิตภัณฑ์โดยประมาณของปี 2504 (แทนค่า x = 2) เป็น

$$\hat{y} = 12 + 3.7(2) = 19.4$$

3. จำนวนของปีเป็นเลขคู่ *(Even number of years)*

ความแตกต่างระหว่างการประยุกต์ของวิธีการกำลังสองน้อยที่สุดต่ออนุกรมเวลาของจำนวนปีเป็นเลขคี่กับเลขคู่เป็นระบบการนับที่ได้ประยุกต์กับค่า x เพื่อว่า $\Sigma x = 0$ สมมุติว่าเรามีข้อมูลต่อไปนี้ของผลิตภัณฑ์น้ำมันปิโตรเลียม ในที่นี้มีจำนวน 6 ปี ด้วยกัน

ปี	x	Y	xy	x^2
2500	-5	5	-25	25
2501	-3	8	-24	9
2502	-1	12	-12	1
2503	1	1.5	1.5	1
2504	3	20	60	9
2505	5	25	125	25
	$\Sigma x = 0$	$\Sigma y = 85$	$\Sigma xy = 139$	$\Sigma x^2 = 70$

มีหลายวิธีของการนับค่า x เพื่อให้ $\Sigma x = 0$ วิธีการหนึ่งที่ต้องออกแบบคือ 1, 3, 5 ดังในตาราง การที่เราไม่สามารถใช้ −3, −2, −1, 1, 2, 3, เพราะว่าระหว่าง −1 กับ 1 มีสองหน่วย (นั่นคือ −1, 0, 1,) ขณะที่ 1, 2, 3, แต่ละปีต่างกัน 1 หน่วย $\Sigma x = 0$ ที่สามารถทำให้เราใช้สูตร (11.3.3)

จากตารางเราคำนวณหา

$$\Sigma y = 85,\ \Sigma xy = 139,\ \Sigma x^2 = 70,\ n = 6$$

แทนค่าเหล่านี้ลงในสูตร (11.3.3) เราได้

$$a = \frac{\Sigma y}{n} = \frac{85}{6} = 14.2$$

$$b = \frac{\Sigma xy}{\Sigma x^2} = \frac{139}{70} = 1.99$$

ดังนั้น เส้นแนวโน้มก็คือ

$$\hat{y} = 14.2 + 1.99x$$

ในเมื่อ x มีหน่วยเป็นครึ่งปี b = 1.99 แสดงการเพิ่มขึ้นของค่าแนวโน้มต่อครึ่งปี (ล้านบาร์เรล)

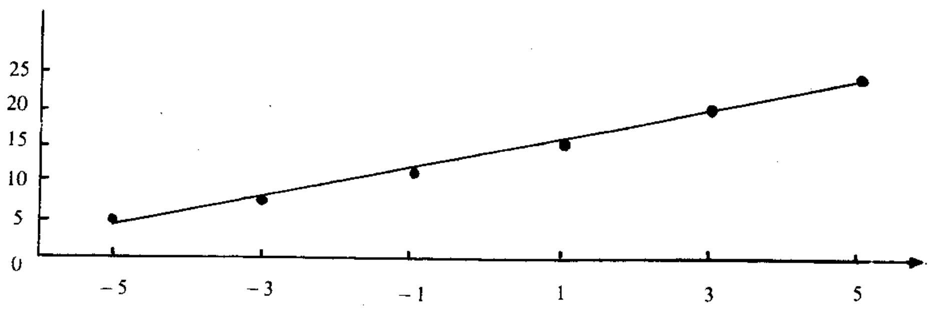
ค่าแนวโน้มผลิตภัณฑ์น้ำมันปิโตรเลียมสำหรับปี 2504 ได้

$$\hat{y} = 14.2 + (1.99)(3) = 20.17$$

สำหรับปี 2505 ได้

$$\hat{y} = 14.2 + (1.99)(5) = 24.15$$

เส้นแนวโน้มกราฟได้ดังรูป 11.6



รูปที่ 11.6

การนับค่า x วิธีอื่นได้โดยการคูณระบบการนับ 1, 3, 5, ฯลฯ ด้วยจำนวนเลขใดเลขหนึ่งก็ได้ ในกรณีนี้เราลองคูณด้วย $\frac{1}{2}$ ดังนั้นระบบการนับทั้งสองเป็น

-5	-3	-1	1	3	5
-2.5	-1.5	-0.5	0.5	1.5	2.5

เมื่อใช้ระบบที่สอง หน่วยของ x ก็กลับมาเป็น 1 ปี แทนที่จะเป็นครึ่งปี บางคนชอบใช้ระบบนี้เพราะว่าต้องการหลีกเลี่ยงหน่วยครึ่งปี

11.4 การเปลี่ยนหน่วยมูลค่าและการเปลี่ยนออริจิน
(Changing the Unit Value and Shifting the Origin)

11.4.1 การเปลี่ยนหน่วยมูลค่า

สมการและข้อมูลของอนุกรมเวลามักจะเป็นยอดรวมรายปี แต่มีหลายกรณี ข้อมูลถูกกำหนดให้เป็นค่าเฉลี่ยรายเดือน ตัวอย่างพอจะแสดงความแตกต่างระหว่างยอดรวมรายปี ค่าเฉลี่ยรายเดือนและสมการของเส้นแนวโน้มที่สมนัยต่อกัน (Corresponding trend line equations)

สมมุติว่า มีนักสถิติคนหนึ่งได้รับค่าจ้างรายปี 2505 เท่ากับ 12,000 บาท และค่าจ้างของเขาได้เพิ่มขึ้นเป็น 12,720 บาท ในปี 2506 และ 13,440 บาท ในปี 2507 แล้วค่าเฉลี่ยรายเดือนสำหรับค่าจ้างปี 2505 เท่ากับ 1,000 บาท สำหรับปี 2506 เท่ากับ 1,060 และสำหรับปี 2507 เท่ากับ 1,120 บาท ดังนั้น ค่าจ้างเพิ่มขึ้นเฉลี่ยรายเดือน คือ 60 บาท สมการยอดรวมรายปี

$$\hat{y} = 12{,}000 \text{ บาท} + 720x \qquad \text{.........(11.4.1)}$$

$x = 0$ คือ วันที่ 1 กรกฎาคม 2505 และ x มีหน่วยเป็น 1 ปี

สมการเฉลี่ยรายเดือน

$$\hat{y} = \frac{12{,}000}{12} + \frac{720}{12}x$$

$$\hat{y} = 1000 + 60x \qquad \text{..........(11.4.2)}$$

$x = 0$ คือ วันที่ 1 กรกฎาคม 2505 และ x ยังมีหน่วยเป็น 1 ปี เราต้องการเปลี่ยนให้เป็นสมการของเดือนจึงจำเป็นจะต้องให้ค่าของ x มีหน่วยเป็น 1 เดือน ด้วย ค่าของ x จึงต้องหารด้วย 12 สมการของเดือนก็ได้เป็น

$$\hat{y} = 1{,}000 + 60\frac{x}{12}$$

$$= 1{,}000 + 5x \qquad ....(11.4.3)$$

$x = 0$ ก็ยังเป็นวันที่ 1 กรกฎาคม 2505 และ x มีหน่วยเป็น 1 เดือน ด้วย

ดังนั้น เมื่อกำหนดสมการให้เป็นแนวโน้ม เราจำเป็นต้องพิจารณาว่าเป็นสมการชนิดไหนของสามสมการนั้น ควรจะสังเกตว่าในสมการ (11.4.3) เราจะมีจุด origin วันที่ 1 ก.ค. เพื่อที่จะให้ถูกกับกฎของเราในการใช้เป็นข้อมูลกลางหรือค่ากลางปีของปีหรือของเดือน จึงควรเคลื่อนไปอยู่กลางเดือนกรกฎาคม นั่นคือ วันที่ 15 กรกฎาคม เนื่องจากเป็นสมการของเดือน

มีอนุกรมเวลาจำนวนมากที่ถูกนำเสนอเป็นข้อมูลเฉลี่ยรายเดือนดังตัวอย่างข้อมูลสำหรับผลิตภัณฑ์กระดาษในตารางที่ 11.1

ตารางที่ 11.1

ผลิตภัณฑ์กระดาษ

ปี	ผลิตภัณฑ์ (พันตัน)
1950	887
1952	908
1954	971
1956	1,166
1958	1,127

ที่มา : สถิติธุรกิจ U.S. Dept. of Commerce 1959

887 หมายความว่าผลิตภัณฑ์เฉลี่ยของกระดาษสำหรับปี 1950 เท่ากับ 887 พันตัน

ตัวอย่างที่ 11.4.1 สมการยอดรวมรายปีสำหรับผลิตภัณฑ์วิทยุของห้างหุ้นส่วนหนึ่งดังนี้ (หน่วยเป็นเครื่อง)

$$\hat{y} = 144 + 72x$$

x มีหน่วยเป็น 1 ปี และ $x = 0$ คือวันที่ 1 กรกฎาคม 1958 นี้หมายความว่าในปี 1958 ค่าแนวโน้มคือ 144 เครื่อง ในปี 1959 ค่าแนวโน้มคือ $144 + 72 \times 1 = 216$ เครื่อง ถ้าต้องการเปลี่ยนสมการนี้ให้เป็นสมการเฉลี่ยรายเดือนได้ดังนี้

$$\hat{y} = \frac{144}{12} + \frac{72X}{12} = 12 + 6x$$

x มีหน่วยเป็น 1 ปี จุด origin หรือ x = 0 คือวันที่ 1 กรกฎาคม 1958 หมายความว่าในปี 1958 ค่าแนวโน้มเฉลี่ยรายเดือนคือ 12 เครื่อง นั่นคือ ทุก ๆ เดือนระหว่างปี 1958 ผลิตวิทยุได้เฉลี่ยเท่ากับ 12 เครื่อง สำหรับปี 1959 เราคำนวณได้

$$\hat{y} = 12 + (6)(1) = 18$$

นั่นคือ ทุก ๆ เดือนระหว่างปี 1959 ผลิตวิทยุได้เฉลี่ยเท่ากับ 18 เครื่อง สำหรับสมการรายเดือนคือ

$$\hat{y} = 12 + 6\frac{x}{12} = 12 + 0.5x$$

x มีหน่วยเป็น 1 เดือน จุด origin หรือ x = 0 คือวันที่ 1 กรกฎาคม 1958 ผลิตภัณฑ์รายเดือนสำหรับวันที่ 1 กรกฎาคม 1958 คือ

$$\hat{y} = 12 + (0.5)(0) = 12$$

เราเปลี่ยนจุด origin เป็นวันที่ 15 กรกฎาคม นั่นคือ ล่วงหน้าไปครึ่งเดือน ผลิตภัณฑ์สำหรับวันที่ 15 กรกฎาคม 1958 ได้

$$\hat{y} = 12 + (0.5)\frac{1}{2} = 12.25$$

ดังนั้น สมการรายเดือนสำหรับวันที่ 15 กรกฎาคม 1958 เป็น

$$\hat{y} = 12.25 + 0.5x$$

x มีหน่วยเป็น 1 เดือน จุด origin คือวันที่ 15 กรกฎาคม 1958 นี้หมายความว่า ค่าแนวโน้มของผลิตภัณฑ์สำหรับเดือนกรกฎาคม 1958 ได้

$$\hat{y} = 12.25 + (0.5)(0) = 12.25$$

นั่นคือ 12.25 เครื่องสำหรับเดือนกรกฎาคม ส่วนสำหรับเดือนสิงหาคม ผลิตภัณฑ์ได้เท่ากับ

$$\hat{y} = 12.25 + (0.5)(1) = 12.75$$

และสำหรับเดือนมิถุนายน

$$\hat{y} = 12.25 + (0.5)(-1) = 11.75$$

ตัวอย่าง 11.4.2 กำหนดให้ตารางข้อมูลสำหรับผลิตภัณฑ์กระเป๋าของห้างหุ้นส่วนหนึ่ง จงคำนวณหาสมการของเส้นแนวโน้มเฉลี่ยรายเดือนโดยวิธีกำลังสองน้อยที่สุด

ปี	เฉลี่ยรายเดือน (1000) x	y	xy	x^2
2496	-2	4	-8	4
249-l	-1	7	-7	1
2498	0	8	0	0
2499	1	10	10	1
2500	2	15	30	4
	$\Sigma x = 0$	$\Sigma y = 44$	$\Sigma xy = 25$	$\Sigma x^2 = 10$

จากข้อมูลเหล่านี้ เราคำนวณหา

$$a = \frac{\Sigma x}{n} = \frac{44}{5} = 8.8$$

$$b = \frac{\Sigma xy}{\Sigma x^2} = \frac{25}{10} = 2.5 \; ; \hat{y} = 8.8 + 2.5x$$

จุด origin คือวันที่ 1 กรกฎาคม 2498 x มีหน่วยเป็น 1 ปี ค่าแนวโน้มของผลิตภัณฑ์เฉลี่ยรายเดือนสำหรับปี 2499 คือ

$$\hat{y} = 8.8 + (2.5)(1) = 11.3$$

นั่นคือ 11.3 พันใบ (หรือกระเป๋า 11.3 × 100 = 11300 ใบ) ต่อเดือน ระหว่างปี 2499 สำหรับยอดรวมรายปีของผลิตภัณฑ์ปี 2499 ได้เท่ากับ

$$(11.3 \times 12) \times 1000 = 135,600$$

นั่นคือ กระเป๋าสำหรับปี 2499 เท่ากับ 135.6 พันใบ หรือ 135,600 ใบ

11.4.2 การเปลี่ยนออริจิน (ORIGIN)

สมการเฉลี่ยรายเดือนเกี่ยวกับค่าจ้างใช้ในหัวข้อ (11.4.1) คือ

$$\hat{y} = 1,000 + 60x$$

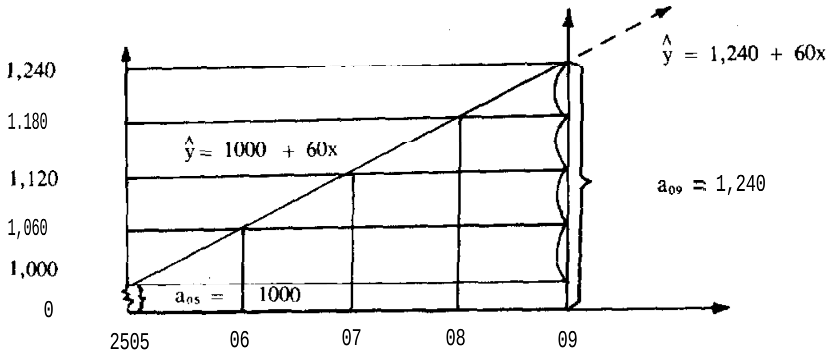
x มีหน่วยเป็น 1 ปี จุด origin คือวันที่ 1 กรกฎาคม 2505 สมมุติว่าเราต้องการเปลี่ยน origin เป็นวันที่ 1 กรกฎาคม 2509 ปัญหาของการเปลี่ยน origin เป็นปี 2509 ในเทอมของรูปที่ 11.7 คือต้องหาจุดตัดที่แกน y ใหม่ a_{09} นี้คือ

$$a_{09} = 1,000 + (60 \times 4) = 1,240$$

ดังนั้น สมการก็กลายเป็น

$$\hat{y} = 1,240 + 60x$$

x มีหน่วยเป็น 1 ปี จุด origin คือ วันที่ 1 กรกฎาคม 2509



รูปที่ 11.7

ถ้าเราต้องการดำเนินการเพียงขั้นเดียวก็ทำได้ดังนี้

$$\hat{y} = 1000 + 60(x + 4)$$
$$= 1000 + 60x + 240$$
$$\hat{y} = 1240 + 60x$$

ในกรณีที่เราต้องการเปลี่ยนจุด origin เป็นวันที่ 1 กรกฎาคม 2500 ก็ทำได้ โดยเอา 5 ไปแทน 4 ในวงเล็บ (5 คือผลต่างระหว่าง ปี 2505 กับ ปี 2500) แต่ปี 2500 เป็นปีก่อนหน้าของปี 2505 เครื่องหมายบวกในวงเล็บก็เปลี่ยนเป็นลบ เราก็ได้สมการ

$$\hat{y} = 1000 + 60(x - 5)$$
$$= 1000 + 60x - 300$$
$$= 700 + 60x$$

นั่นคือ สมการเฉลี่ยรายเดือนที่ x มีหน่วยเป็น 1 ปี จุด origin วันที่ 1 ก.ค. 2500

แบบฝึกหัดที่ 11.1

1. มูลค่าดัชนีของผลิตภัณฑ์นมปี 1941-1954 ในเวอร์จิเนียร์ โดยให้ 1935 – 39 = 100 มีดังนี้ 114, 116, 117, 120, 124, 125, 128, 135, 139, 142, 138, 136, 145 และ 147 (ที่มา : สถานีการทดลองการเกษตรเวอร์จิเนียร์)
   ก) ใช้วิธีกึ่งเฉลี่ย (Semi average) หามูลค่าแนวโน้มสำหรับปี 1944 กับ 1951 พลอตข้อมูลและลากเส้นตรงผ่านมูลค่าแนวโน้มปี 1944 กับ 1951 (120.6 กับ 140.3)
   ข) คำนวณหาแนวโน้มรายปีเพิ่มขึ้นและคำนวณมูลค่าแนวโน้มปี 1941, 1948 และ 1953 (2.8, 112.2, 131.9 และ 145.9)
2. จากแบบฝึกหัดข้อ 1 ใช้ข้อมูลอันเดียวกัน จงคำนวณหาส่วนเฉลี่ยเคลื่อนที่ 3 ปี พร้อมทั้งกราฟเส้นตรงส่วนเฉลี่ยเคลื่อนที่ 3 ปี
3. จำนวนลูกจ้างทั้งหมดของปี 1935 – 1949 ในสหรัฐ ดังต่อไปนี้ 28.9, 31.9, 33.1, 31.5, 33.0, 34.8, 38.2, 40.7, 42.2, 41.5, 39.8, 41.4, 43.2, 44.3 และ 42.8 (ล้านคน)
   ก) จงหาสมการกำลังสองน้อยที่สุด ($\hat{y} = 37.82 + 1.07x$)
   ข) จงคำนวณหาค่าแนวโน้มของปี 1935 กับ 1949 (30.3 กับ 45.3)
4. จากแบบฝึกหัดข้อ 1 คำนวณหาสมการเส้นกำลังสองน้อยที่สุดและใช้สมการที่คำนวณหาค่าแนวโน้มปี 1941 และ 1954 ($\hat{y} = 130.43 + 1.3x$; 113.53 กับ 147.3)
5. ดัดแปลง สมการแนวโน้มในแบบฝึกหัดข้อ 4 โดยการเปลี่ยนจุด origin มาเป็นปี 1941 และ x มีหน่วยเป็น 1 ปี ($\hat{y} = 113.53 + 2.6x$)
6. การขนส่งเหล็กกล้าของบริษัทหนึ่งสำหรับปี 1930 – 1955 ดังนี้ 2.1, 1.4, 0.8, 1.4, 1.5, 2.2, 3.4, 3.4, 1.9, 3.3, 4.2, 5.9, 5.8, 5.9, 5.9, 5.4, 4.7, 6.1, 6.4, 5.1, 6.4, 7.0, 6.0, 7.1, 5.0 และ 7.0 ล้านตัน จงหาเส้นแนวโน้มกำลังสองน้อยที่สุด ($\hat{y} = 4.435 + 0.118x$)
7. ดัดแปลงสมการแนวโน้มในแบบฝึกหัดข้อ 6 โดยการเปลี่ยนจุด origin ของ x ไปเป็นปี 1950 แล้วดัดแปลงสมการนี้ให้เป็น สมการเฉลี่ยรายเดือนของการขนส่งเหล็กกล้า ($\hat{y} = .517 + 0197x$)
8. กำหนดให้สมการ $\hat{y} = 8.32 + 0.22x$ (ใช้ปี 1950 เป็นจุด origin x มีหน่วยเป็น 1 ปี y กำไรรายปี

หน่วยพันล้านดอลลาร์) ซึ่งเราต้องการหาค่าแนวโน้มของกำไร เปลี่ยนจุด origin เป็นปี 1946 แล้วคำนวณหามูลค่าแนวโน้มรายปีของปี 1946 ถึงปี 1954 (7.44, 7.66, 7.88, 8.1, 8.32, 8.54, 8.76, 8.98, 9.2)

11.5 แนวโน้มของเส้นโค้งแบบพาราโบล่า (Parabola Trends)

เราไม่สามารถที่จะอธิบายค่าแนวโน้มด้วยเส้นตรงเสมอไปในกรณีที่อนุกรมไม่อยู่ในลักษณะเป็นเส้นตรง เราจึงต้องพิจารณาปรับให้เส้นเป็นเส้นโค้งชนิดใดชนิดหนึ่ง มีเส้นโค้งอยู่หลายชนิดที่ใช้ในการปรับแนวโน้ม มีอยู่ชนิดหนึ่งที่มีประโยชน์มากคือ เส้นโค้งแบบพาราโบล่า ซึ่งสมการก็เขียนได้เป็น

$$\hat{y} = a + bx + cx^2 \qquad \text{........ (il.5.1)}$$

เรียกว่า "สมการโปลีโนเมียลกำลังที่สอง" (second degree polynomial equation) ซึ่งกำลังสองของ x เป็นกำลังสูงสุดในสมการนี้

ปัญหาของการกำหนดแนวโน้มแบบพาราโบล่าเป็นสิ่งจำเป็นที่จะต้องคำนวณหาค่าคงที่ a, b และ c ดังปรากฏในสมการ (11.5.1) เพื่อที่จะให้ได้เส้นโค้งอยู่ในลักษณะเป็น best fit นิยามของ goodness of fit ในลักษณะของกำลังสองน้อยที่สุด (least squares) เราจะต้องกำหนด a, b และ c เพื่อให้ $\Sigma(y-\hat{y})^2$ มีค่าน้อยที่สุด แน่ละก็เพราะว่าค่าแนวโน้ม $\hat{y}$ คำนวณได้จากสมการ (11.5.1)

การคำนวณหาค่า a, b และ c ก็หาได้จากสมการปกติ 3 สมการ ดังนี้

$$\Sigma y = n \cdot a + b \cdot \Sigma x + c \cdot \Sigma x^2 \qquad \text{..........(11.5.2)}$$

$$\Sigma xy = a \cdot \Sigma x + b \cdot \Sigma x^2 + c \cdot \Sigma x^3 \qquad \text{..........(11.5.3)}$$

$$\Sigma x^2y = a \cdot \Sigma x^2 + b \cdot \Sigma x^3 + c \cdot \Sigma x^4 \qquad \text{.........(11.5.4)}$$

เพื่อทำให้ง่ายเข้าในการใช้สมการเหล่านี้ เราใช้วิธีการอย่างเดียวกันกับหัวข้อ (11.3.4) คือเปลี่ยนสเกลของ x เป็น ....., −3, −2, −1, 0, 1, 2, 3, ..., หรือ ....., −5, −3, −1, 1, 3, 5, ..., ขึ้นอยู่กับว่ามีจำนวนปีเป็นเลขคู่หรือเลขคี่ เพื่อให้ Σx กับ Σx^3 ทั้งสองมีค่าเท่ากับ 0 สมการปกติก็ลดรูปลงได้เป็น

$$\Sigma y = n \cdot a + c \cdot \Sigma x^2 \qquad \text{.......... (11.5.5)}$$

$$\Sigma xy = b \cdot \Sigma x^2 \qquad \text{.......... (11.5.6)}$$

$$\Sigma x^2 y = a \cdot \Sigma x^2 + c \cdot \Sigma x^4 \qquad \text{..........(11.5.7)}$$

เราสามารถหาค่า b ได้โดยตรงจากสมการ (11.5.6) ส่วนค่า a และ c คำนวณหาได้จากการแก้สมการ (11.5.5) กับ (11.5.7)

ตัวอย่างแสดงวิธีกำลังสองน้อยที่สุดของการกำหนดแนวโน้มแบบพาราโบล่าใช้ข้อมูลเกี่ยวกับการประกันชีวิตแบบสามัญในมลรัฐเทนเนสซี่ช่วงระหว่างปี 1938 ถึง 1954 จำนวนขายประจำปีเหล่านี้ได้มาจากสมาคมประกันชีวิตดังแสดงในคอลัมน์ y ของตาราง

คำนวณค่า a, b และ c ได้จากสมการ (11.5.5), (11.5.6) และ (11.5.7) ซึ่งเราต้องหาค่า n, Σy, Σxy, $\Sigma x^2 y$, Σx^2 และ Σx^4 จากตารางต่อไปนี้

ปี	จำนวนเงินที่ขายรายปี (ล้านดอลลาร์) Y	X	xy	x^2y	x^2	x^4
1938	88.1	- 8	- 704.8	5,638.4	64	4,096
1939	89.1	- 7	- 623.7	4,365.9	49	2,401
1940	88.6	- 6	-531.6	3,189.6	36	1,296
1941	101.9	- 5	- 509.5	2,547.5	25	625
1942	86.7	- 4	- 346.8	1,387.2	16	256
1943	96.8	- 3	- 290.4	871.2	9	81
1944	112.7	- 2	- 225.4	450.8	4	16
1945	129.2	- 1	- 129.2	129.2	1	1
1946	202.0	0	0	0	0	0
1947	195.4	1	195.4	195.4	1	1
1948	192.8	2	385.6	771.2	4	16
1949	191.9	3	575.7	1,727.1	9	81
1950	237.4	4	949.6	3,798.4	16	256
1951	234.6	5	1,173.0	5,865,0	25	625
1952	270.9	6	1,625.0	9,752.4	36	1,296
1953	320.0	7	2,240.0	15,680.0	49	2,401
1954	338.0	8	2,704.0	21,632.0	64	4,096
	$\Sigma y = 2{,}976.1$	$\Sigma x = 0$	$\Sigma xy = 6{,}487.3$	$\Sigma x^2y = 78{,}001.3$	$\Sigma x^2 = 408$	$\Sigma x^4 = 17{,}544$

แทนค่ายอดรวมของคอลัมน์ xyและคอลัมน์ x^2 ลงในสมการ (11.5.6) เราได้

$$6,487.3 = b \times 408$$

$$b = \frac{6,487.3}{408} = 15.90$$

และแทนค่า $n = 17$ กับยอดรวมของคอลัมน์ y, x^2y, x^2 และ x^4 ลงในสมการ (11.5.5) กับ (11.5.7) เราได้

$$2,976.1 = 17a + 408c \qquad \text{..........(1)}$$

$$78,001.3 = 408a + 17,544c \qquad \text{..........(2)}$$

แก้สมการทั้งสองนี้ได้ด้วยสมการที่หนึ่งหารตลอดด้วย 17 และสมการที่สองหารด้วย 408 แล้ว เอาสองสมการมาลบกันหาค่า c ต่อไปแทนค่า c ลงไปในสมการหนึ่งสมการใดก็สามารถคำนวณ หาค่า a ได้ ในที่นี้เราคำนวณหาค่า a และ c ได้

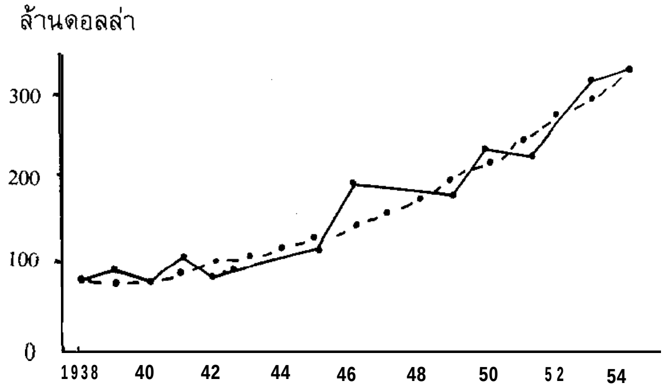
$a = 154.71$ และ $c = 0.8482$

และสมการของแนวโน้มแบบพาราโบล่าก็เป็น

$$\hat{y} = 154.71 + 15.90x + 0.8482x^2$$

x มีหน่วยเป็น 1 ปี จุด origin คือปี 1946 y คือยอดรวมจำนวนเงินที่ขายมีหน่วยเป็นล้านดอลลาร์ สมการนี้ $a = 154.71$ เป็นค่าแนวโน้มสำหรับปี 1946 $b = 15.90$ เป็นความชันของเส้นโค้งที่ $x = 0$ และ $2c = 1.6964$ เป็นอันตราการเปลี่ยนแปลงของความชันที่จุดนี้เฉพาะ

ต้องการหาค่าแนวโน้มสำหรับปีใดปีหนึ่งที่กำหนดให้ เราก็แทนค่า x ลงไปในสมการแนวโน้ม ดังตัวอย่าง เราต้องการค่าแนวโน้มสำหรับปี 1948 ก็แทนค่า $x = 2$ เราได้ $\hat{y} = 154.71 + 15.9(2) + 0.8482(2)^2 = 189.90$ สำหรับปี 1940 ก็แทนค่า $x = -6$ เราก็ได้ $\hat{y} = 154.71 + 15.9(-6) + 0.8482(-6)^2 = 89.85$ ค่าแนวโน้มสำหรับปี 1938 ถึง 1954 ดังแสดงในรูป 11.8 พร้อมกับข้อมูลจริง ลากเส้นโค้งผ่านจุดที่ใช้แทนค่าแนวโน้มก็จะได้เส้นโค้งแบบพาราโบล่าซึ่งกำหนดด้วยวิธีกำลังสองน้อยที่สุด



รูปที่ 11.8

จำนวนขายของการประกันชีวิตแบบสามัญในเทนเนสซี่

ปัญหาเกี่ยวกับเส้นโค้งแนวโน้มแบบพาราโบล่าที่พอจะยกมากล่าวเสมอก็คือ เมื่อไรหรือภายใต้โอกาสอะไรก็ควรจะใช้แนวโน้มเช่นนั้น แม้ว่าจะยากลำบากก็ยังเป็นไปตามเกณฑ์ที่โดยทั่วไปจะช่วยกันได้บ้าง ตัวอย่างแสดงการพิจารณาเส้นโค้งแบบพาราโบล่า

$$y = x^2 + 3x + 2$$

เราคำนวณค่า y สำหรับ x = 0, 1, 2, 3, 4, 5 และ 6 ได้โดยแทนค่า x เหล่านี้ลงในสมการก็จะได้ค่า y ผลต่างครั้งแรกและครั้งที่สองของ y ดังแสดงได้จากข้างล่าง ผลต่างครั้งแรกคำนวณได้จากการลบกันของค่า y ตัวก่อนเอาไปลบค่า y ตัวถัดไปทำอย่างนี้ไปเรื่อย ๆ ส่วนผลต่างครั้งที่สองก็คำนวณได้โดยวิธีเดียวกันกับครั้งแรก แต่แทนที่จะเป็นค่า y ก็ได้ค่าของผลต่าง

x	y	ผลต่างครั้งแรก	ผลต่างครั้งที่สอง
0	2		
		4	
1	6		2
		6	
2	12		2
		8	
3	20		2
		10	
4'	30		2
		12	
5	42		2
		14	
6	56		

นี้แสดงว่า ถ้าอนุกรมของจุดวางอยู่บนเส้นโค้งแบบพาราโบล่าจริงแล้ว ผลต่างครั้งที่สองของค่าก็จะอยู่บนเส้นโค้งแบบพาราโบล่าเช่นเดียวกัน ถึงแม้ว่าผลต่างครั้งที่สองของอนุกรมจะมีความคล้ายคลึงกันมากหรือน้อยก็อาจใช้เป็นเครื่องแสดงว่า เส้นโค้งแบบพาราโบล่าให้ good fit อย่างไรก็ตาม ถ้าหากว่าผลต่างครั้งที่สองมีค่าห่างไกลกันมาก เส้นโค้งแบบพาราโบล่าก็ยังให้ลักษณะที่ดีของแนวโน้ม

แนวโน้มมีอีกหลายชนิด เช่น สมการแนวโน้มของเส้นโค้งที่มีกำลังสูงกว่าสอง $\hat{y} = a + bx + cx^2 + dx^3$, $\hat{y} = a + bx + cx^2 + dx^3 + ex^4$ ฯลฯ ซึ่งไม่ได้ใช้บ่อยนัก ส่วนแนวโน้มแบบเอกซ์โปเนนเชียลและแนวโน้มอื่น ๆ จะไม่กล่าวในที่นี้

แบบฝึกหัดที่ 11.2

1. จากตัวอย่างในหัวข้อที่ (11.5) เราคำนวณหา $\hat{y} = 154.71 + 15.90x + .8482x^2$ (origin คือ ปี 1946 x มีหน่วยเป็น 1 ปี $\hat{y}$ คือยอดรวมจำนวนขายมีหน่วยเป็นล้านดอลลาร์) เป็นเส้นโค้งแบบพาราโบล่า

(ก) ใช้สมการข้างต้นคำนวณหาค่าแนวโน้มของแต่ละปีที่กำหนดให้ (81.8, 85.0, 89.8, 96.4, 104.7, 114.6, 126.3, 139.6, 154.7, 171.4, 189.9, 210.0, 231.9, 255.4, 280.6, 307.6, 336.2)

(ข) พลอตค่าแนวโน้มเหล่านี้พร้อมกับข้อมูลจริง เวลาเส้นโค้งผ่านจุดที่ใช้แทนค่าแนวโน้มเพื่อแสดงแนวโน้มของเส้นโค้งแบบพาราโบล่า

2. ตัวเลขเหล่านี้แสดงถึงการสูบใบยาเส้นต่อหัวในสหรัฐ ของปี 1930 ถึง 1955

ปี	จำนวน (ปอนด์)	ปี	จำนวน (ปอนด์)
1930	8.8	1943	11.5
1931	8.4	1944	11.2
1932	7.6	1945	12.5
1933	7.8	1946	12.2
1934	8.3	1947	12.0
1935	8.2	1948	12.1
1936	8.8	1949	11.9

1937	9.0	1950	12.0
1938	8.8	1951	12.5
1939	8.8	1952	12.9
1940	9.1	1953	12.9
1941	9.8	1954	12.2
1942	10.7	1955	12.3

(ก) ปรับแนวโน้มของเส้นโค้งแบบพาราโบล่า โดยวิธีกำลังสองน้อยที่สุด ($\hat{y} = 10.6013 + 0.1129x - 0.0057x^2$; origin คือ ปี 1942-43; x มีหน่วยครึ่งปี)

(ข) ใช้สมการที่คำนวณได้จาก (ก) คำนวณหาค่าแนวโน้มสำหรับปี 1930 ตลอดจนถึงปี 1955 (7.4, 7.7, 8.0, 8.3, 8.5, 8.8, 9.0, 9.3, 9.5, 9.8, 10.0, 10.3, 10.5, 10.7, 10.9, 11.2, 11.4, 11.6, 11.8, 12.0, 12.2, 12.4, 12.5, 12.7, 12.9, 13.1)

3. ประชากรของประเทศสหรัฐระหว่างปี 1850–1950 ในช่วง 10 ปี

ปี	1850	1860	1870	1880	1890	1900	1910	1920	1930	1940	1950
ประชากร (ล้านคน)	23.2	31.4	39.8	50.2	62.9	76.0	92.0	105.7	122.8	131.7	151.1

(ก) จงคำนวณหาสมการแนวโน้มของเส้นโค้งแบบพาราโบล่า ($\hat{y} = 76.64 + 13.00x + .3974x^2$; จุด origin คือปี 1900; x มีหน่วยเป็น 10 ปี)

(ข) จงหาค่าแนวโน้มของปีที่กำหนดให้เปรียบเทียบกับค่าจริง (21.6, 31.0, 41.2, 52.2, 64.0, 76.6, 90.0, 104.2, 119.2, 135.0, 151.6)

(ค) จงหาค่าแนวโน้มของปี 1945, 1960 และปี 1840 (143.2, 168.9, 12.9)

11.6 การพยากรณ์ (Forecasting)

ในบทก่อน ๆ ของหนังสือเล่มนี้ เราก็ได้ทราบถึงความแตกต่างระหว่างสถิติภาคพรรณนา (descriptive statistics) กับสถิติภาคอนุมาณ (inductive statistics) จะสังเกตเห็นได้ว่าสถิติภาคพรรณนาเกี่ยวข้องอยู่กับเบื้องต้นส่วนสถิติภาคอนุมาณเกี่ยวข้องในลักษณะทั่ว ๆ ไป เช่น การทำนาย การประมาณค่าและการตัดสินใจ ความแตกต่างนี้ก็พอที่จะทำให้เข้าใจแจ่มแจ้งขึ้น การกำหนดค่าแนวโน้มที่ได้อธิบายการเปลี่ยนแปลงของอนุกรมในบทนี้ก็ยังอยู่ในขอบเขตสถิติภาค

พรรณนา การปรับค่าแนวโน้มเพื่ออธิบายข้อมูลในอดีตยังไม่เพียงพอสำหรับนักธุรกิจผู้ซึ่งจะต้องรอบรู้เกินขอบเขตของประสบการณ์ในอดีต ก็เพื่อทำการอนุมาณอะไรที่อาจเกิดขึ้นในอนาคต นักธุรกิจจะต้องทำนายหรือพยากรณ์ระดับอนาคตของกิจกรรมธุรกิจและเขาจะต้องใช้อย่างน้อยที่สุดก็ส่วนหนึ่งของเหตุการณ์ที่เกิดขึ้นในอดีตเป็นพื้นฐานในการทำนาย

การประมาณค่าซึ่งอยู่นอกขอบเขตของค่าพื้นฐานซึ่งใช้คำนวณหาสมการ อย่างเช่น เราต้องการประมาณค่าหรือทำนายปี 2510 ก็แทนค่า $x = 8$ ลงในสมการ (11.3.2) $\hat{y} = 12 + 3.7x$ เราได้ $\hat{y} = 12 + 3.7(8) = 41.6$ นั่นคือ ในปี 2510 เราทำนายการผลิตน้ำมันปิโตรเลียมรายปีได้ 41.6 บาร์เรล

การประมาณค่าแนวโน้มเป็นสิ่งจำเป็น ถึงแม้ว่าจะเป็นการเดาแต่กระบวนการจะสำเร็จได้ก็ขึ้นอยู่กับปัจจัยหลายอย่าง ปัญหาพื้นฐานที่เราพบเสมอก็คือสิ่งที่ได้กระทำไปแล้วในอดีตจะนำไปทำต่อไปในอนาคต จะต้องเป็นวิธีการเดียวกัน ความพยายามที่จะทำนายก็ยิ่งลำบากมากขึ้น ถ้าหากว่าระยะเวลาของอนาคตยิ่งห่างไกลมากขึ้นเท่านั้น สมการแนวโน้มที่ให้ค่าประมาณอาจถูกดัดแปลงไปตามความประสงค์ของคนใดคนหนึ่งเพื่อให้เป็นไปตามผลของปัจจัยต่าง ๆ

ปัญหาทั้งหมดของการพยากรณ์เกี่ยวกับจำนวนข้อบังคับนอกเหนือสถิติอื่น ๆ จะไม่ขอกล่าวในที่นี้ ตามที่ได้กล่าวมาแล้วจะเห็นได้ว่า อย่างน้อยที่สุดสถิติก็ช่วยพัฒนาให้คำตอบที่สมบูรณ์ในปัญหาของการพยากรณ์ทางเศรษฐกิจ ถ้าหากว่านักธุรกิจใช้สถิติอย่างถูกต้องด้วยความเหมาะสมเท่าที่จะเป็นไปได้ และอยู่ในขอบเขตของสถิติแล้ว สถิติก็จะให้คุณค่าอันหามิได้ที่จะนำไปสู่จุดหมายและการตัดสินใจอย่างถูกต้องและฉลาด

แบบฝึกหัดที่ 11.3

1. ใช้สมการกำลังสองน้อยที่สุดในแบบฝึกหัดที่ 11.1 ข้อ 3 เพื่อพยากรณ์จำนวนลูกจ้างปี1965 ในสหรัฐอเมริกา (62.43)

2. ใช้สมการกำลังสองน้อยที่สุดในแบบฝึกหัดที่ 11.1 ข้อ 6 เพื่อพยากรณ์การขนส่งเหล็กกล้าของบริษัทปี 1965 (9.745)

3. ใช้สมการแนวโน้มของเส้นโค้งแบบพาราโบล่าในแบบฝึกหัดที่ 11.2 ข้อ 2 เพื่อทำนายการสูบยาสูบต่อหัวของปี 1925 (5.95155)