

## บทที่ 5

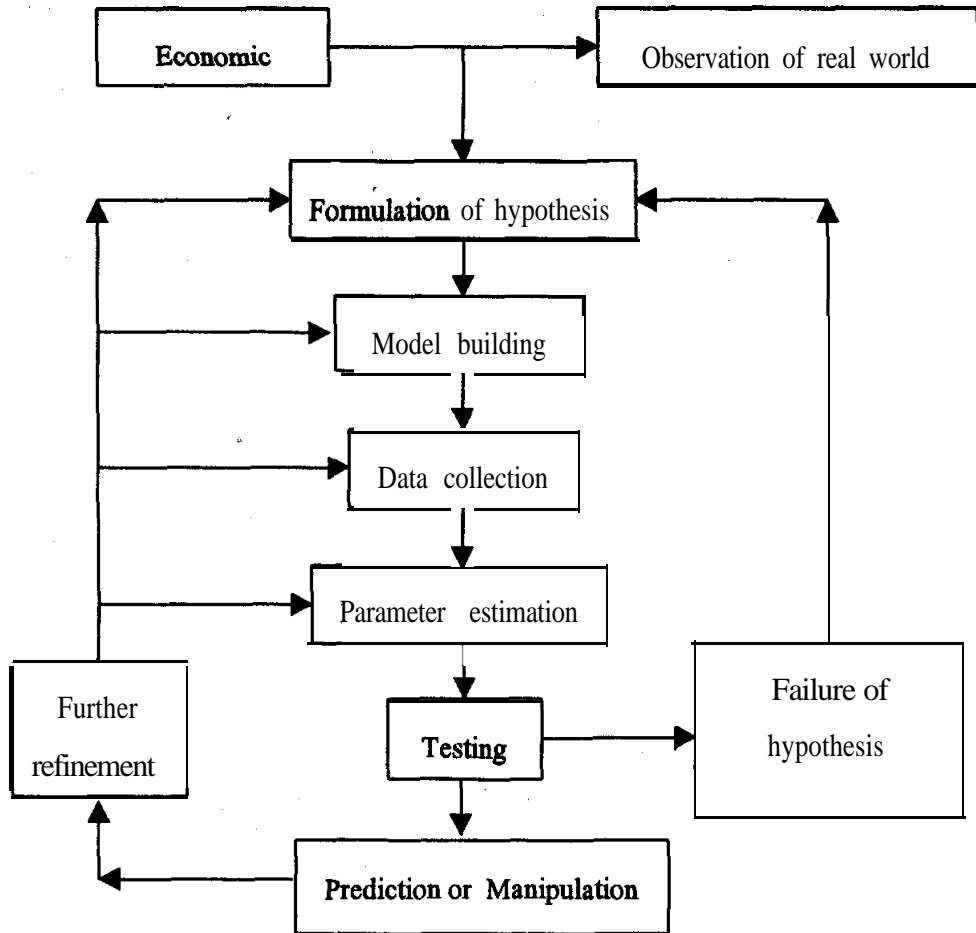
### วิธีการทางเศรษฐมิติและสมการถดถอยอย่างง่าย (Econometric Method and Simple Regression)

#### 5.1 บทนำ

ในหนังสือเล่มนี้ทั้งเล่ม เราใช้ตัวอย่างการคำนวณโดยใช้สมการที่เคยเรียนมาแล้ว เช่น สมการอุปสงค์และอุปทาน สมการบริโภค สมการสาธารณูปโภค สมการการผลิตและอื่นๆ เราได้ชี้แจงไว้แล้วว่าสมการเชิงปฏิบัติเหล่านี้ จะไม่ปรากฏขึ้นในชีวิตจริงโดยปาฏิหาริย์ แต่สมการเหล่านี้จะอธิบายพฤติกรรมของหน่วยต่างๆทางเศรษฐศาสตร์ เช่น บริษัท ผู้บริโภค หรืออุตสาหกรรมต่างๆ สมการเหล่านี้จะรวบรวมมาจากการสังเกตสิ่งที่เกิดขึ้นจริงในโลก การเก็บรวบรวมข้อมูลและการประมาณการของสมการเชิงปฏิบัตินำไปสู่การประยุกต์ใช้ วิธีการทางสถิติเหล่านี้มีเนื้อหาหากไม่สามารถอธิบายให้ละเอียดได้ด้วยพื้นที่จำกัดเพียงไม่กี่หน้า อย่างไรก็ตามนักศึกษาผู้สนใจวิชาเศรษฐศาสตร์อย่างจริงจังควรศึกษาความรู้เบื้องต้นของวิชาสถิติหรือวิชาเศรษฐมิติเพิ่มเติม ซึ่งสิ่งที่สำคัญมากคือความเข้าใจเกี่ยวกับแนวคิดพื้นฐานและวิธีการของเศรษฐมิติจะทำให้สามารถเรียนรู้เศรษฐศาสตร์เชิงปริมาณซึ่งได้อธิบายเนื้อหาเบื้องต้นของวิชานี้ไว้ในหัวข้อต่างๆ โดยมีวัตถุประสงค์คือ เพื่อให้นักศึกษารู้จักเนื้อหาเบื้องต้นของวิชานี้ นักศึกษาทุกคนควรอ่านบทนี้ก่อนที่จะอ่านบทต่อไป

#### 5.2 ขั้นตอนวิธีการดำเนินงานของเศรษฐมิติ

ก่อนที่จะลงไปรายละเอียดทางคณิตศาสตร์หรือสถิติศาสตร์ นักศึกษาจำเป็นต้องเข้าใจว่านักเศรษฐมิติมีวิธีการทำงานอย่างไร วิธีการของเขาก็คือ ตั้งอยู่บนพื้นฐานทางวิธีการทางวิทยาศาสตร์อย่างเคร่งครัด ไม่เช่นนั้นแล้ว การนำตัวเลขต่างๆ มาใช้ก็จะไร้ทิศทางทำให้ได้ผลลัพธ์ที่ไม่มีประโยชน์ เพื่อให้เห็นภาพในวิชาเศรษฐมิติ อันที่จริงวิธีทางวิทยาศาสตร์มิได้จำกัดเฉพาะในวิชาเศรษฐมิติเท่านั้น แต่ยังสามารถนำไปประยุกต์ใช้ได้กับการวิเคราะห์ทางเศรษฐศาสตร์ด้วยการทดลองและการสังเกตใดๆก็ได้ ที่จริงแล้ว นักเศรษฐศาสตร์ที่แท้จริงอาจจะคำนึงถึง เรื่องของลำดับขั้นในการประมาณตัวพารามิเตอร์ และอาจจะพอใจที่จะทิ้งส่วนอื่นให้นักทฤษฎีเศรษฐศาสตร์ทั่วไป



รูปที่ 5.1 แสดงขั้นตอนการดำเนินงานของเศรษฐมิติ

พิจารณาแต่ละขั้นตอนในลำดับต่อไปนี้

### 5.2.1 ทฤษฎีและการสังเกต

การกำเนิดแนวคิดต่างๆ เกี่ยวกับว่าระบบเศรษฐศาสตร์ทำงานอย่างไร อาจจะมาจากข้อพิจารณาทางทฤษฎีจากการสังเกตพฤติกรรมทางเศรษฐศาสตร์ในโลกแห่งความจริง หรือมาจากทั้ง 2 แหล่งผสมผสานกันสำหรับความเป็นไปได้ที่ข้อมูลทั้ง 2 แหล่งมาผสมผสานกันให้พิจารณาตัวอย่างต่อไปนี้

จากการสังเกต สถิติสำหรับเศรษฐกิจโดยรวมอาจจะแสดงให้เห็นว่าบางช่วงเวลาในอดีต การนำเข้าสินค้าอุปโภคและการใช้จ่ายเพื่อการบริโภคเพิ่มขึ้นทั้งสองอย่าง การสังเกตนี้ทำให้เราตั้งสมมติฐานว่าการเพิ่มขึ้นของตัวเลขการนำเข้าสินค้าขึ้นอยู่กับค่าใช้จ่าย แต่ก็อาจเลือกอีกวิธีว่าตัวเลขการใช้จ่ายขึ้นอยู่กับตัวเลขการนำเข้า เราอาจจะแยกสมมติฐาน 2 ข้อนี้ออกจากกันด้วยการใช้

ทฤษฎี ซึ่งจะบอกเราว่าผู้บริโภครับสนองความต้องการของตนเอง โดยเลือกสรรสินค้าตอบสนองจากสินค้าภายในประเทศ และความต้องการส่วนที่เหลือได้รับการตอบสนองจากสินค้านำเข้าจากต่างประเทศ การสร้างทฤษฎีด้วยวิธีที่ต่างๆไปแบบนี้บอกว่าการนำเข้านั้นจะขึ้นอยู่กับค่าใช้จ่ายกับการอุปโภคบริโภคมากกว่าสิ่งอื่นๆ ดังนั้นเราจึงสร้างสมมติฐานในรูปแบบนี้ออกมาไว้สำหรับการสำรวจด้วยการทดลองและการสังเกตภายหลัง

### 5.2.2 การสร้างสมมติฐาน

สมมติฐานประกอบด้วยข้อความที่เกี่ยวกับความเป็นจริงซึ่งเอามาจากทฤษฎี และหรือเอามาจากการสังเกตจากเหตุการณ์จริง ในการสร้างทฤษฎีขึ้นมาทฤษฎีหนึ่งจำเป็นมากที่จะต้องแน่ใจว่าข้อสมมติฐานที่ถูกสร้างขึ้นมานั้นสามารถพิสูจน์ว่าผิดได้ (Refutation) การให้ความสำคัญของการพิสูจน์ว่าผิด แทนที่จะเป็นการทดสอบสมมติฐานมักจะทำให้นักศึกษาสับสนซึ่งควรจะได้รับการอธิบายดังต่อไปนี้

ทฤษฎีที่พิสูจน์ไม่ได้ว่าผิด ไม่สามารถนำไปทดสอบอย่างมีความหมายอะไรได้ และบอกอะไรเราไม่ได้เกี่ยวกับโลกนี้ ทฤษฎีนี้เราเรียกว่า ความว่างเปล่า (Empty) พิจารณาตัวอย่างของสมมติฐานว่างเปล่า 2 ข้อ ข้อแรกพิจารณาทฤษฎีที่ว่าอุปสงค์ของกล้วยลดลงเมื่อราคากล้วยสูงขึ้น กำหนดให้สิ่งอื่นๆยังคงเหมือนเดิม ถ้ามิได้กำหนดอนุประโยคสุดท้ายเอาไว้ สมมติฐานนี้ก็จะเป็นสมมติฐานที่ไม่สามารถนำมาทดสอบได้ สมมติว่าเราสังเกตว่าราคากล้วยเพิ่มขึ้น และอุปสงค์ลดลง เราก็สรุปได้ว่าสมมติฐานนี้ถูกต้อง เพราะสิ่งอื่นๆไม่ต้องเหมือนเดิม หรือกล่าวอีกอย่างหนึ่งก็คือ ถ้าอนุประโยคมิได้ระบุไว้ ไม่ว่าจะการสังเกตจะออกมาเป็นอย่างไร ก็ใช้ได้กับสมมติฐานที่เราตั้งไว้ เพราะฉะนั้นสมมติฐานนี้ก็พิสูจน์ไม่ได้ว่าผิด

ตัวอย่างที่ 2 ของสมมติฐานว่าง คือสมมติฐานที่มีตัวแปรที่ไม่ได้อธิบายเอาไว้อย่างอิสระ ยกตัวอย่างสมมติฐานที่ว่าผู้บริโภครจัดสรรรายได้เพื่อจะได้ใช้สาธารณูปโภคมากที่สุด ถ้าเราสามารถให้คำอธิบายอย่างอิสระของคำว่า สาธารณูปโภค ทฤษฎีที่พิสูจน์ไม่ได้ว่าผิดเพราะเราสามารถจะให้ความหมายของคำว่า สาธารณูปโภค อย่างไรก็ได้เพื่อจะให้เหมาะสมกับการสังเกตพฤติกรรมผู้บริโภค

อะไรทำให้เกิดการพิสูจน์ว่าผิดของทฤษฎี คำทฤษฎีนั้นเป็นความเชื่อว่าจะเกิดขึ้นจากสาเหตุที่เกิดขึ้นมาก่อนและเปลี่ยนแปลงอะไรไม่ได้ การพิสูจน์ว่าผิดก็ง่ายมาก ยกตัวอย่างเช่น สมมติฐานที่ว่าหงส์ทุกตัวมีสีขาว พิสูจน์ได้ว่าผิดเมื่อพบกับหงส์ที่มีสีดำ(แม้ว่าทฤษฎีจะไม่ได้ถูกพิสูจน์มาว่ากลุ่มตัวอย่างในการสังเกตมีแต่หงส์สีขาว) อย่างไรก็ตามทฤษฎีทางเศรษฐศาสตร์ไม่ได้

ถูกสร้างขึ้นมาเป็นแบบขาวกับดำอย่างนั้น แต่มักจะกล่าวว่ามีสภาพบางอย่างเกิดขึ้น ก็จะมีอิทธิพลที่จะมีแนวโน้มที่จะเกิดผลลัพธ์เช่นนั้นเช่นนี้

ความเชื่อมโยงของผลกระทบและสาเหตุก็จะถูกสร้างขึ้นมาในเชิงทฤษฎี แล้วอนุกรมของกรรສັงເກດที่เป็นข้อเท็จจริงก็ถูกสร้างขึ้น ซึ่งก็จะตรงหรือไม่ตรงกันกับทฤษฎี วิธีการทางสถิติกำหนดเงื่อนไขในการพิจารณาว่าการสังเกตที่ไม่ตรงกันเป็นจำนวนเท่าไร จึงมีความเกี่ยวข้องกับ ความบังเอิญทางสถิติอย่างแท้จริง และจำนวนเท่าไรที่จะพิสูจน์ทฤษฎีได้ว่าผิด แต่ก็เพียงแสดงได้ แต่มั่นตรงกันหรือไม่ หรือตรงกันกับกลุ่มข้อมูลจริงที่ได้จากการสังเกตเท่านั้น

ถ้าห้บกระบวนการสร้างสมมติฐาน ต้องระมัดระวังมากเพื่อที่จะได้แน่ใจว่า ความสัมพันธ์ที่เป็นเหตุเป็นผลระหว่างปรากฏการณ์นั้นได้ถูกสร้างไว้อย่างเหมาะสมตั้งแต่เริ่มแรก ถ้าไม่ได้ทำเช่นนั้น ผลลัพธ์ของการทดลองนั้นก็คงจะแปลความหมายผิดได้ง่าย มีนักเศรษฐศาสตร์หลายคน ที่ติดกับดัก ศาสตราจารย์ทางจิตวิทยาท่านหนึ่งที่ใช้เวลาหลายปีสอนแมลงสาบให้เดินด้วยคำสั่งว่า “เดินเร็ว” ขึ้นตอนถัดไปศาสตราจารย์ทางจิตวิทยาท่านนั้นดึงขาแมลงสาบออกทั้งหมด แล้วก็สั่งให้มันเดิน สิ่งมีชีวิตที่น่าสงสารตัวนั้นมันก็ขยับเขยื้อนไม่ได้ จากสิ่งที่มองเห็นได้ทั้งหมด ท่านศาสตราจารย์ก็สรุปว่าการดึงขาแมลงสาบออกมีผลกระทบต่อการได้ชินของมัน

### 5.2.3 การสร้างแบบจำลอง

สมมติฐานสร้างข้อความที่บอกว่าโลกมีพฤติกรรมอย่างไร แบบจำลองที่เราเห็นในบทก่อนหน้านี้สร้างตัวแปรหลักเกี่ยวกับการทดสอบสมมติฐานและเชื่อมโยงตัวแปรทั้งหลายนี้โดยใช้คำพูด และหรือใช้ข้อความทางคณิตศาสตร์ แบบจำลองจึงให้กลวิธีหรือวิธีการที่มีรูปแบบที่ใช้ในการทดสอบสมมติฐานได้

แบบจำลองเศรษฐมิติสร้างมาจากสมการโครงสร้างหนึ่งสมการหรือมากกว่า (Structural equations) สมการที่มีความสำคัญมากที่สุดคือ สมการพฤติกรรม (behavioral equations) อันได้แก่ สมการการบริโภค สมการการลงทุน สมการสาธารณูปโภคและอื่นๆ ซึ่งสมการเหล่านี้จะอธิบายพฤติกรรมของหน่วยต่างๆ ในเศรษฐศาสตร์

เราได้อธิบายถึงความแตกต่างระหว่าง ตัวแปรภายในและตัวแปรภายนอกแล้ว ในแบบจำลองทางเศรษฐศาสตร์ นักศึกษาคควรกลับไปอ่านบทที่ 3 เพื่อทบทวนแนวคิดที่สำคัญนี้

### 5.2.4 การเก็บรวบรวมข้อมูล

ตอนนี้เราอยู่ในขั้นที่เราได้สร้างรูปแบบจำลองทางทฤษฎีเพื่อทดสอบสมมติฐานบางข้อแล้ว ขณะนี้เราจึงเข้าใกล้การเอารูปแบบจำลองของเราไปใช้ในการทดสอบข้อเท็จจริงจากการ

ทดลองและการสังเกต อย่างไรก็ตามจำเป็นต้องเก็บรวบรวมข้อมูลจากการทดลองและการสังเกตก่อน

โดยทั่วไปมีวิธีการรวบรวมข้อมูลในวิชาเศรษฐมิติอยู่ 2 วิธี วิธีแรกเราอาจจะสังเกตจากพฤติกรรมของบุคคล องค์กร อุตสาหกรรม หรือเศรษฐกิจโดยรวม ภายในระยะเวลาที่ต่อเนื่องกัน ช่วงหนึ่ง ยกตัวอย่างเช่น เก็บมาอย่างต่อเนื่องรายเดือนหรือรายปี ผลของการสังเกตจุดนี้เรียกว่าการวิเคราะห์แบบอนุกรมเวลา (Time series analysis) หรือเราอาจจะเก็บข้อมูล ณ จุดๆ หนึ่งของเวลา โดยใช้ข้อมูลจากบุคคลหลายๆคน องค์กรต่างๆ หรือหลายอุตสาหกรรม โดยปกติแล้วจะต้องใช้กลุ่มตัวอย่างจากประชากร บุคคล องค์กรหรืออุตสาหกรรมทั้งหมด จากการศึกษา ด้วยวิธีการนี้เราเรียกว่า การวิเคราะห์แบบข้ามเขต (Cross section analysis) บางครั้งเราอาจจะใช้วิธีการอนุกรมเวลา และวิธีการวิเคราะห์ข้ามเขตทั้งสองพร้อมกันในการเก็บรวบรวมข้อมูล

### 5.2.5 การประมาณพารามิเตอร์

ขั้นตอนนี้อาจจะกล่าวได้ว่าเป็นหัวใจของการสำรวจเศรษฐมิติ ณ จุดๆ นี้ แบบจำลองเชิงทฤษฎีของเราและข้อมูลที่ได้จากการสังเกตและทดลองมาเผชิญหน้ากัน ยกตัวอย่าง แบบจำลองของเราอาจจะประกอบไปด้วยสมการอุปสงค์จากการสมมติฐาน คือ

$$Q = a + bP \quad (5.1)$$

เมื่อให้  $Q$  และ  $P$  เป็นปริมาณและราคาของสินค้าตามลำดับ ข้อมูลของเราประกอบไปด้วยชุดการสังเกตชุดหนึ่งของ  $Q$  และ  $P$  งานของเราตอนนี้ก็คือ เพื่อให้ได้ค่าประมาณที่เป็นตัวเลขจากพารามิเตอร์  $a$  และ  $b$

อาจจะใช้วิธีการได้หลายวิธีเพื่อการประมาณการนี้ วิธีที่นิยมมากที่สุด คือการใช้สมการถดถอยแบบเส้นตรง (Linear regression) ซึ่งเราอาจจะกล่าวถึงในรายละเอียดต่อไป

### 5.2.6 การทดสอบ

การทดสอบทางสถิติที่ประยุกต์ใช้กับการประมาณพารามิเตอร์ซึ่งได้มาจากขั้นตอนที่แล้วช่วยให้เราตัดสินใจได้ว่าสมมติฐานนี้ใช้ได้หรือไม่ได้กับการสังเกตจากการทดลองของเรา เราจะศึกษาการทดสอบเหล่านี้ในบทต่อไป ถ้าสมมติฐานนี้ไม่ผ่านการทดสอบการทดลองและการสังเกต เราจะต้องกลับไปยังขั้นตอนก่อนหน้านี้อีก ในลำดับขั้นที่แสดงในรูป 5.1 ยกตัวอย่างเช่น ค่าประมาณ  $a$  และ  $b$  จากสมการข้อ 5.1 ระบุว่าไม่มีความสัมพันธ์เชิงเส้นระหว่าง  $Q$  และ  $P$  ในกรณีนี้เราอาจจะคิดทบทวนสมมติฐานของเราอีกครั้ง บางทีอาจต้องนำตัวแปรใหม่มาอธิบายความหลากหลายใน  $Q$  ซึ่งก็จำเป็นที่จะต้องถอยกลับไปยังลำดับแรกใหม่ ตั้งแต่การสร้างสมมติฐาน

### 5.2.7 การพยากรณ์

ถ้าการทดสอบในส่วนก่อนหน้านี้นับสนับสนุนสมมติฐานภายใต้ข้อพิจารณาและเราสามารถที่จะใช้แบบจำลองนี้เป็นตัวแทนที่ได้จากการทดลองและการสังเกตของระบบภายใต้การศึกษาของเรา เราอาจจะต้องการสำรวจพฤติกรรมของแบบจำลองนี้ ภายใต้สถานการณ์ที่แตกต่างไปก็ได้ โดยเฉพาะเราอาจจะสนใจที่จะพยากรณ์ค่าตัวแปรบางตัว ในแบบจำลองโดยใช้ส่วนประกอบอื่นๆ ในระดับที่กำหนดให้ ตัวอย่างเช่น ชุดค่าของ  $a$  และ  $b$  ในสมการ 5.1 ขอมริบได้ เราอาจจะต้องการทำนายระดับอุปสงค์ที่จะเกิดขึ้นที่ระดับราคาแตกต่างกัน

จำเป็นที่ต้องเข้าใจในความแตกต่างของการทำนายและการพยากรณ์ การพยากรณ์หรือการคาดการณ์ในอนาคตเป็นกรณีพิเศษของการทำนาย นั่นก็คือเป็นคำอธิบายของเหตุการณ์ในอนาคตหรือการทำนายโดยในตัวของมันเอง ไม่เกี่ยวข้องกับอนาคต เป็นเพียงกระบวนการอนุมานอย่างง่ายว่าคาดว่าจะเป็นอะไรจะเกิดขึ้นมาจากผลลัพธ์ที่เป็นเหตุเป็นผลของการสันนิษฐานในบางเรื่อง ยกตัวอย่างเช่น ทฤษฎีอุปสงค์และอุปทาน ทำนายว่า เมื่ออุปสงค์คงที่ ภายใต้เงื่อนไขที่ว่าเมื่ออุปทานลดลงจะทำให้ราคาเพิ่มขึ้น การทำนายนี้เป็นเชิงคุณภาพโดยธรรมชาติ แต่แบบจำลองที่เกิดจากการทดลองและการสังเกตเป็นการทำนายเชิงปริมาณ ดังนั้น เมื่อกำหนดโครงสร้างของแบบจำลองเป็นการเฉพาะ แล้วกำหนดค่าของตัวแปรอิสระให้ จะสามารถทำนายค่าของตัวแปรตามได้

คอนท่ายขั้นตอนนี้อาจจะเห็นได้ว่ามีเครื่องหมายของงานขั้นสุดท้ายของการวิเคราะห์ที่เศรษฐศาสตร์จากการทดลองและการสังเกต แต่ผลลัพธ์ที่ได้จะต้องนำไปทำให้ละเอียดประณีตมากขึ้นกว่านี้ ด้วยการทำพร้อมกันไปกับสมมติฐานและแบบจำลองต่างๆ หรืออาจจะต้องเก็บข้อมูลใหม่ หรือประมาณค่าพารามิเตอร์ใหม่จากการใช้ข้อมูลที่มีอยู่เดิมด้วยเทคนิคการประมาณที่ต่างไปจากเดิม บางทีอาจจำเป็นต้องทดสอบแบบจำลองอีกครั้ง ความเป็นไปได้ของที่กล่าวมานี้ดูได้จากลูกศรที่ชี้ไปทางซ้ายของรูป 5.1

## 5.3 แบบจำลองถดถอยอย่างง่าย

ขณะนี้เราจะมุ่งไปที่ปัญหาการประมาณค่าพารามิเตอร์ของแบบจำลองโดยการใช้ชุดข้อมูลที่ได้มาจากการทดลองและการสังเกต และใช้เทคนิคการประมาณค่าโดยใช้สมการถดถอยแบบกำลังสองน้อยที่สุด เราจะตรวจปัญหานี้ในเนื้อหาโดยใช้ตัวอย่าง

### 5.3.1 ตัวอย่าง

สมมติว่าผู้ประกอบการรายเคียวต้องการหาสมการอุปสงค์ที่ได้จากการทดสอบและการสังเกตสำหรับสินค้าของเขา สมมติความสัมพันธ์ออกมาในรูปแบบนี้

$$Y = \alpha + \beta x$$

เมื่อ Y = ปริมาณที่ต้องการในหน่วยพันต่อเดือน

X = ราคาต่อหน่วยในหน่วยบาท

$\alpha, \beta$  = พารามิเตอร์ไม่ทราบค่าที่จะนำมาประมาณ เรียกว่าค่าสัมประสิทธิ์จากการถดถอย

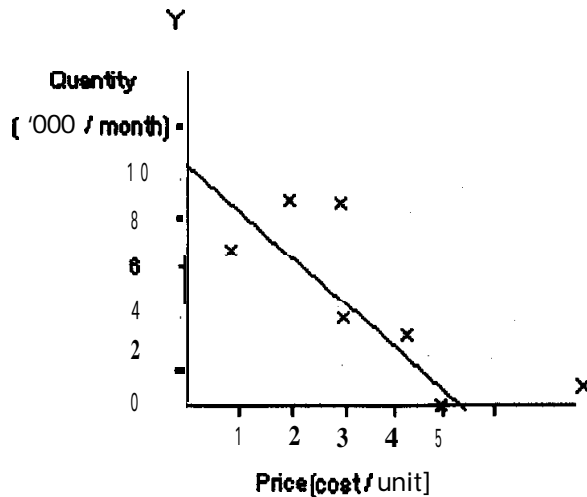
สมมติว่าราคาแตกต่างกันสำหรับสินค้า และปริมาณซื้อขายจะแตกต่างกันตามราคาที่แตกต่างกัน ภายใน 6 เดือน ชุดข้อมูลที่ได้จากการสังเกต ในตารางที่ 5.1

ตารางที่ 5.1 แสดงอุปสงค์ของตลาดผูกขาด (Monopolist's Demand)

Observation (i)	Month	Price (X) (บาท)	Sales (Y) (‘000)
1	Jan.	2	8
2	Feb.	4	3
3	Mar.	3	4
4	Apr.	1	7
5	May	3	8
6	June	5	0

รูปภาพแสดงความสัมพันธ์ระหว่าง Y และ X ได้มาจากการกำหนดจุดต่างๆ ของตัวเลขที่ได้จากตารางที่ 5.1 ในกราฟ ตัวเลขที่ได้จากการสังเกต จาก 1-6 ในตาราง 5.1 แสดงให้เห็นได้ในรูปที่ 5.2 จากการมองคร่าวๆ สามารถลากเส้นความสัมพันธ์ระหว่าง Y และ X ออกมาได้ ในรูปที่ 5.2 จากการใช้วิธีการตามที่ได้เรียนมาในบทที่ 4 เราสามารถวัดความชันของเส้นตรงนี้ได้ประมาณ -2 และมีจุดตัดที่ประมาณ 10.5 หรืออีกนัยหนึ่งค่าประมาณพารามิเตอร์จากสมการ 5.2 คือ  $\alpha = 10.5$  และ  $\beta = -2$  หรือเขียนสมการอธิบายข้อมูลเหล่านี้ในรูป

$$Y = 10.5 - 2X \quad (5.3)$$



รูปที่ 5.2 แสดงกราฟที่วาดโดยประมาณด้วยสายตา

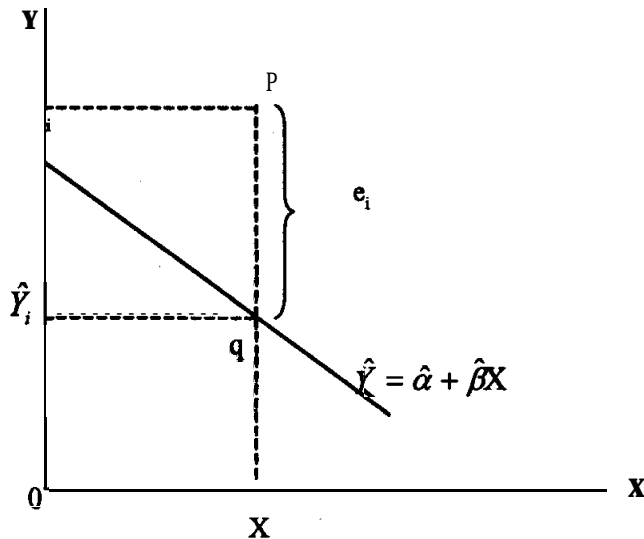
เป็นที่ชัดเจนว่าวิธีวาดกราฟเป็นการประมาณเท่านั้น เราต้องการวิธีการที่มีความถูกต้องมากกว่านี้ในการหาเส้นตรงที่เหมาะสมกับกลุ่มข้อมูล เราจะทำเช่นนั้นได้ด้วยการตัดสินใจว่า อะไรคือความหมายของความเหมาะสมที่สุด

### 5.3.2 ค่าความผิดพลาด (The Error Term)

สมมติเราให้ค่าสังเกต  $Y$  และ  $X$  เป็น  $Y_i$  และ  $X_i$  ตามลำดับ จากตัวอย่าง  $i=1,2,\dots,6$  ต่อไป ถ้าเป็นค่าตัวแปรอิสระที่ใช้ถูกทำนายโดยความสัมพันธ์ระหว่าง  $Y$  และ  $X$  สำหรับ  $X_i$  ที่กำหนดให้จากการใช้กราฟเลือก  $X_i$  มาตรฐานหนึ่ง ตามที่แสดงในรูป 5.3 ค่าสังเกตของ  $Y$  ที่ระดับ  $X$  ณ จุด  $p$  มันถูกบ่งชี้ในระดับของ  $Y_i$  บนแนวตั้ง เราพิจารณาเส้นความชันแนวตั้งลงในรูปที่ 5.3 จะเป็นตัวแทนความสัมพันธ์ระหว่าง  $Y$  และ  $X$

ความสัมพันธ์นี้ทำนายค่าของ  $Y$  ณ จุด  $q$  สำหรับค่า  $X_i$  ที่กำหนดให้ ค่าของ  $Y$  ที่ทำเครื่องหมายบนเส้นแนวตั้ง สุดท้ายให้  $e_i$  เท่ากับระยะทางจากค่าสังเกตและค่าทำนายของ  $Y$  ที่ระดับ  $i$  ของ  $X$  ยกตัวอย่างในรูปที่ 5.3 ความยาวของ  $e_i$  คือ ระยะทางระหว่าง  $p$  และ  $q$





รูปที่ 5.3 แสดงค่าผิดพลาด

ดังนั้นเราสามารถเขียนค่าพยากรณ์ของ Y เท่ากับ

$$\hat{Y}_i = \hat{\alpha} + \hat{\beta}X_i \quad (5.4)$$

หมวดที่อยู่บน  $\alpha$  และ  $\beta$  ใช้นี้ซึ่งถือว่าเป็นการประมาณจากค่าสัมประสิทธิ์จากการถดถอย (Regression coefficient) และ เราอาจจะเขียนค่าสังเกตของ Y นี้ว่า

$$Y_i = \hat{Y}_i + e_i \quad (5.5)$$

และเมื่อแทนสมการ (5.4) ลงสมการ (5.5) เราจะได้

$$Y_i = \hat{\alpha} + \hat{\beta}X_i + e_i \quad (5.6)$$

ความหมายของ  $e$  เรียกว่า ค่าความแตกต่างหรือความคลาดเคลื่อน (error term or residual) ถ้าเรามีค่าสังเกต  $n$  ตัวของ Y และ X (เช่น ถ้า  $i = 1, 2, \dots, n$ ) ก็จะมีค่าความแตกต่าง  $n$  ตัว สำหรับค่าสังเกตแต่ละตัว

### 5.3.3 หลักเบื้องต้นของวิธีกำลังสองน้อยสุด (The Principle of Least Squares)

เมื่อใช้จุดต่างๆ ที่กระจายตามรูปที่ 5.2 ก็ค่อนข้างชัดเจนว่าความหมายของเส้นตรงที่เหมาะสมที่สุดนั้น คือ เส้นตรงที่ทำให้ผลบวกของระยะห่างจากจุดต่างๆ มายังเส้นตรงนี้มีค่าน้อยที่สุดเท่าที่จะเป็นไปได้ เมื่อระยะห่างระหว่างจุดแต่ละจุด เราเรียกว่าค่าความคลาดเคลื่อน  $e_i$  ดังนั้น ถ้าเราเพียงแต่รวมเอา  $e_i$  ทั้งหมดไว้ด้วยกัน ค่าบวกกับค่าลบมันก็จะบวกลบตัวมันเอง ปัญหานี้แก้ไขได้ด้วยการยกกำลังสอง  $e_i$  เมื่อ  $e_i^2$  ก็จะได้ค่าเป็นบวก ไม่ว่า  $e_i$  จะเป็นบวกหรือลบ ตอนนี้เราก็สามารถ

ให้ความหมายของเส้นตรงที่เหมาะสมที่สุดว่าเป็น ผลบวกของความแตกต่างยกกำลังสองของจุดทุกจุดรวมกันแล้วมีค่าต่ำสุดเท่าที่จะเป็นไปได้ ตามความหมายนี้เราเรียกมันว่า หลักการกำลังสองน้อยที่สุด มันเป็นหลักการพื้นฐานของเทคนิคการถดถอยเชิงเส้น

ในทางคณิตศาสตร์ ผลรวมของกำลังสองของค่าผลต่างนี้ก็คือ ผลรวมของ  $e_i^2$  จากค่าสังเกต  $i = 1, 2, \dots, n$  เมื่อจัดลำดับใหม่สมการ 5.6 เราก็จะได้

$$e_i = Y_i - \hat{\alpha} - \hat{\beta}X_i \quad (5.7)$$

ยกกำลังสองทั้ง 2 ข้างจากสมการ 5.7 และบวกผลรวมของค่าสังเกตทั้งหมด ยกตัวอย่างว่า จาก  $i = 1, 2, \dots, n$  จะได้

$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n (Y_i - \hat{\alpha} - \hat{\beta}X_i)^2 \quad (5.8)$$

ปัญหาของเราก็คือ ต้องเลือก  $\hat{\alpha}$  และ  $\hat{\beta}$  ในสมการ 5.8 เพื่อให้ค่า  $\sum e^2$  นั้นต่ำที่สุด เราประยุกต์วิธีการนี้แบบตรงๆ ด้วยวิธีการ ทำให้น้อยที่สุดแบบไม่จำกัด โดยการดิฟเฟอเรนเชียลสมการ 5.8 เป็นส่วนๆ ออกมา ตามลำดับ  $\hat{\alpha}$  และ  $\hat{\beta}$  แล้วให้มีค่าเท่ากับ 0 เราจะได้

$$\frac{\partial}{\partial \hat{\alpha}} \sum e^2 = -2 \sum (Y - \hat{\alpha} - \hat{\beta}X) = 0 \quad (5.9a)$$

$$\frac{\partial}{\partial \hat{\beta}} \sum e^2 = -2 \sum X(Y - \hat{\alpha} - \hat{\beta}X) = 0 \quad (5.9b)$$

เมื่อทำให้สมการ 2 ข้อข้างต้นง่ายลง เราได้สมการดังต่อไปนี้

$$\sum Y = n\hat{\alpha} + \hat{\beta} \sum X \quad (5.10a)$$

$$\sum XY = \hat{\alpha} \sum X + \hat{\beta} \sum X^2 \quad (5.10b)$$

จากกลุ่มข้อมูลข้างต้นเราสามารถคำนวณ

$$\sum Y, n, \sum X, \sum XY \quad \text{และ} \quad \sum X^2$$

ดังนั้นเมื่อสิ่งที่เราไม่รู้จากสมการ 5.10 มีเพียง  $\hat{\alpha}$  และ  $\hat{\beta}$  สมการ 5.10 จึงเรียกว่า สมการปกติ (Normal Equation)

5.3.4 คำตอบสำหรับตัวอย่างที่เป็นจำนวนตัวเลข

ให้สมการปกติจากข้างต้นมา เราคำนวณ  $\hat{\alpha}$  และ  $\hat{\beta}$  สำหรับตัวอย่างจากข้อ 5.3.1 ได้ว่า เริ่มต้นเราต้องคำนวณผลรวม X และ Y กำลังสองของสินค้าหลายชนิด เช่น ผลรวมของ XY วิธีการคำนวณแสดงในตารางที่ 5.2

ตารางที่ 5.2 แสดงตัวอย่างการคำนวณค่าRegression

X	Y	X <sup>2</sup>	Y <sup>2</sup>	XY
2	8	4	64	16
4	3	16	9	12
3	4	9	16	12
1	7	1	49	7
3	8	9	64	24
5	0	25	0	0
$\Sigma X=18$ n = 6	$\Sigma Y=30$	$\Sigma X^2=64$	$\Sigma Y^2=202$	$\Sigma XY=71$

ต่อไปนี้จะแทนค่าจากตารางที่ 5.2 ในสมการ (5.10a และ 5.10 b) ได้ว่า

$$30 = 6\hat{\alpha} + 18\hat{\beta} \quad (5.11a)$$

$$71 = 18\hat{\alpha} + 64\hat{\beta} \quad (5.11b)$$

คำตอบของสมการ 2 สมการข้างต้น สำหรับค่าที่ไม่ทราบคือ  $\hat{\alpha}$  และ  $\hat{\beta}$  นำมาสู่ขั้นตอนต่อไปนี้เป็น

คูณสมการ (5.11 a) ด้วย 3

$$90 = 18\hat{\alpha} + 54\hat{\beta} \quad (5.11c)$$

ลบสมการ 5.11c จากสมการ 5.11b

$$-19 = 10\hat{\beta}$$

$$\hat{\beta} = -1.9$$

แล้วเอาไปแทนลงในสมการ (5.11a) จะได้

$$\hat{\alpha} = 10.7$$

ดังนั้นค่าประมาณสมการถดถอยจะอยู่ในรูป

$$Y = 10.7 - 1.9X \quad (5.12)$$

สมการนี้จะอธิบายได้ว่าเส้นตรงที่เหมาะสมที่สุดสำหรับข้อมูลในตารางที่ 5.1 เมื่อเปรียบเทียบสมการ (5.12) กับ (5.3) แสดงให้เห็นว่าการเดาด้วยกราฟก็ค่อนข้างใกล้เคียงกับวิธีนี้

## 5.4 ทางเลือกอีกทางหนึ่งสำหรับหาค่าสัมประสิทธิ์การถดถอยวิธีทั่วไป

### 5.4.1 วิธีทั่วไป (General Method)

อีกวิธีหนึ่งของการหาค่าสัมประสิทธิ์ถดถอยให้ไว้ต่อไปนี้จะไม่พิสูจน์ที่มาให้ดู แต่มีหลักการคล้ายกับวิธีสมการปกติ คือมีหลักการเบื้องต้น โดยการใช่วิธีกำลังสองน้อยที่สุด

$$\text{ให้ } \bar{X} = \frac{1}{n} \sum X \quad \text{และ} \quad \bar{Y} = \frac{1}{n} \sum Y \quad (5.13)$$

เมื่อ  $\bar{X}$  และ  $\bar{Y}$  คือค่าเฉลี่ยสำหรับ  $X$  และ  $Y$  ตามลำดับ ให้ค่าความแตกต่างของค่าสังเกตของ  $X$  และ  $Y$  จากค่าเฉลี่ยโดยใช้สัญลักษณ์  $x$  และ  $y$  นั้นก็คือจากค่าสังเกตของ  $i$  แต่ละตัวจะได้ว่า

$$x = X_i - \bar{X} \quad \text{และ} \quad y = Y_i - \bar{Y} \quad (5.14)$$

ทำให้ได้สัมประสิทธิ์การถดถอยด้วยสูตรต่อไปนี้

$$\hat{\beta} = \frac{\sum xy}{\sum x^2} \quad (5.15a)$$

$$\text{และ} \quad \hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X} \quad (5.15b)$$

$$\text{หรือแสดงได้ว่า} \quad \sum xy = \sum XY - n\bar{X}\bar{Y} \quad (5.16)$$

$$\text{และ} \quad \sum x^2 = \sum X^2 - n\bar{X}^2 \quad (5.17)$$

และสมการ (5.15a) อาจจะเขียนได้อีกวิธีหนึ่ง คือ

$$\hat{\beta} = \frac{\sum XY - n\bar{X}\bar{Y}}{\sum X^2 - n\bar{X}^2} \quad (5.18)$$

### 5.4.2 การประยุกต์กับตัวอย่างที่เป็นตัวเลข

การนำสมการ (5.15a) มาใช้คำนวณ  $\hat{\beta}$  จำเป็นต้องคำนวณความแตกต่างของค่าสังเกตแต่ละตัวของ  $X$  และ  $Y$  ตามลำดับ แสดงไว้ในตารางที่ 5.3 ผลลัพธ์จากตารางนี้ใส่เข้าไปในสมการ (5.15a)

$$\text{ได้ว่า} \quad \hat{\beta} = -\frac{19}{10} = -1.9 \quad (5.19)$$

$$\hat{\beta} = -1.9$$

ตารางที่ 5.3 แสดงการคำนวณค่า X และ Y ต่าง ๆ

X	Y	$x = (X - \bar{X})$	$y = (Y - \bar{Y})$	$x^2$	$y^2$	xy
2	8	-1	3	1	9	-3
4	3	1	-2	1	4	-2
3	4	0	-1	0	1	0
1	7	-2	2	4	4	-4
3	8	0	3	0	9	0
5	0	2	-5	4	25	-10
$\bar{X} = \frac{18}{6} = 3$	$\bar{Y} = \frac{30}{6} = 5$			$\sum x^2 = 10$	$\sum y^2 = 52$	$\sum xy = -19$

เพื่อเขียนสมการถดถอยที่ได้จากตารางที่แล้ว ผลลัพธ์จากสมการ 5.18 ทำให้ได้คำตอบเหมือนเดิม จึงเป็นวิธีการที่นิยมใช้ในการคำนวณวิธีหนึ่งจากข้อมูลในตาราง 5.2 และ 5.3 เราได้แทนค่าสมการ 5.18 ตามนี้

$$\hat{\beta} = \frac{71 - 6(3)(5)}{64 - 6(3)^2} = -1.9 \quad (5.20)$$

สุดท้ายเราคำนวณค่าของ  $\hat{\alpha}$  โดยเอาไปลบออกจากสมการ 5.18b ได้ว่า

$$\begin{aligned} \hat{\alpha} &= 5 - (-1.9)(3) \\ &= 10.7 \end{aligned}$$

ซึ่งก็จะได้ผลลัพธ์เท่ากัน

วิธีการคำนวณหาค่าสัมประสิทธิ์การถดถอยจะใช้แบบไหนขึ้นอยู่กับความพอใจของผู้วิเคราะห์และสถานการณ์ในการคำนวณ สำหรับปัญหาที่ยากซับซ้อน อาจต้องใช้คอมพิวเตอร์เพื่อลดระยะเวลาในการคำนวณได้ดี

### 5.5 การใช้รูปแบบของสมการถดถอย กับ ความสัมพันธ์แบบไม่เป็นเส้นตรง

ในบทนี้เราได้เห็นถึงการใช้งานของสมการถดถอยแบบเส้นตรงเพื่อจะประมาณค่าพารามิเตอร์ของสมการเส้นตรงที่มีตัวแปรเดียว จากการพิจารณาความกว้างและความหลากหลายของสมการที่เราได้เรียนในหนังสือ จะเห็นได้ว่ารูปแบบของสมการถดถอยอย่างง่าย ที่มีเพียง 2 ตัวแปรที่อยู่ในความสัมพันธ์แบบเส้นตรง ไม่ได้กล่าวถึงการประมาณค่าความสัมพันธ์ที่ได้จากประสบ

การณ์จริง ที่เรามักจะพบในเศรษฐศาสตร์ ตามที่ได้พูดถึงแล้วว่านักเศรษฐศาสตร์มักจะให้ความสำคัญสำคัญกับสมการของตัวแปรหลายตัวที่อยู่ในความสัมพันธ์แบบไม่เป็นเส้นตรง

เราต้องเข้าใจก่อนว่ารูปแบบของสมการถดถอยอย่างง่ายอาจจะถูกพัฒนาให้ใช้ในกรณีของความสัมพันธ์แบบเส้นตรงที่มีหลายตัวแปร ยกตัวอย่างเช่น สมมติเราพิจารณาความสัมพันธ์ระหว่างตัวแปรตาม  $Y$  และผลรวมของตัวแปรอิสระ  $k$  ตัว  $x_1, x_2, x_3, \dots, x_k$  สมการเส้นตรงอาจเขียนอยู่ในรูป

$$Y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k \quad (5.21)$$

นอกจากการประมาณสัมประสิทธิ์ถดถอยข้างต้นแล้ว เรายังสามารถคำนวณหาค่าของ  $\alpha, \beta_1, \beta_2, \dots, \beta_k$  ในสมการ (5.21) สำหรับค่าที่ได้จากการสังเกตใน  $Y$  และตัวแปร  $X$  เมื่อสมการถดถอยถูกใช้ในสมการที่มีตัวแปรตั้งแต่ 2 หรือมากกว่า จะเรียกว่า multiple regression

นอกจากนี้เมื่อความสัมพันธ์แบบไม่เป็นเส้นตรงถูกตั้งขึ้นระหว่างตัวแปรต่างๆ มันมีหลายวิธีที่สามารถใช้ในการประมาณค่าพารามิเตอร์ของรูปแบบที่เหมาะสม ตัวแปรในกระบวนการประมาณที่ไม่ใช่เส้นตรงอาจสามารถนำมาใช้ได้ หรือมีความเป็นไปได้ที่จะนำข้อมูลพื้นฐาน เช่น วิธีการถดถอยเส้นตรงแบบดั้งเดิมมาใช้

เพื่อที่จะแสดงถึงความเป็นไปได้ในกรณีหลัง สมมติเราใช้สมการ Cobb-Douglas

$$y = AK^b \quad (5.22)$$

เมื่อ  $y =$  ผลลัพธ์

$K =$  ทุน

เรามีข้อมูลของ  $y, K$  และต้องการประมาณค่าพารามิเตอร์  $A, b$  ในสมการที่ไม่ใช่เส้นตรง เราจะใช้วิธีใส่  $\log$  ในสมการ (5.22) เพื่อจะลดรูปให้เป็นสมการเส้นตรง

$$\log y = \log A + b \log K \quad (5.23)$$

ดังนั้นถ้าเราเปลี่ยนจากข้อมูลดั้งเดิมของ  $y$  และ  $K$  สู่ตัวแปรใหม่ เช่น  $Y, X$  ตามลำดับ เราสามารถใช้สมการเส้นตรงถดถอยเพื่อที่จะประมาณค่าพารามิเตอร์ของความสัมพันธ์

$$Y = \alpha + \beta X \quad (5.24)$$

โดยที่  $Y = \log y$

$X = \log K$

$a = \log A$

$\beta = b$

ในตัวอย่างต่อไป เราให้ความสัมพันธ์แบบกำลังสองเป็นสมมติฐานระหว่างตัวแปร 2 ตัว คือ  $z$  และ  $w$  อยู่ในรูป

$$z = a + bw + cw^2 \quad (5.25)$$

สำหรับข้อมูลของ  $Z$  และ  $w$  เราอาจใช้สมการถดถอยแบบ multiple เพื่อจะประมาณค่าพารามิเตอร์  $a, b, c$  ของสมการ (5.25) โดยใช้วิธีเปลี่ยนรูปของตัวแปรดั้งเดิม

$$\text{ให้ } X_1 = w$$

$$X_2 = w^2$$

หลังจากนั้น เราสามารถประมาณสมการเส้นตรง

$$z = a + bX_1 + cX_2 \quad (5.26)$$

โดย  $a, b, c =$  สัมประสิทธิ์ถดถอย

ไม่ใช่ความสัมพันธ์ที่ไม่เป็นเส้นตรงทั้งหมดจะสามารถใช้สมการถดถอยเส้นตรงในการประมาณค่าโดยการเปลี่ยนรูปตัวแปรเท่านั้น ยกตัวอย่างเช่น

$$Y = a/x + b \quad (5.27)$$

ไม่สามารถประมาณโดยใช้สมการถดถอยเส้นตรงด้วยการเปลี่ยนรูปข้อมูลใน  $Y, x$  ได้ ในกรณีเช่นนี้ การประมาณค่าสำหรับสมการที่ไม่ใช่เส้นตรงจะต้องถูกนำมาใช้

-----

## คำถามท้ายบทที่ 5

1. เขียนข้อความเกี่ยวกับสิ่งที่นักเศรษฐมิติ จำเป็นต้องระวังในการศึกษาถึงสาเหตุของเงินเฟ้อว่าเกิดจากการเพิ่มขึ้นของความต้องการหรือต้นทุนที่สูงขึ้น
2. พิสูจน์สูตรในสมการที่ (5.15) เพื่อใช้ในการประมาณค่าสัมประสิทธิ์การถดถอยโดยใช้หลักของกำลังสองน้อยที่สุด
3. แสดงให้เห็นว่า  $\sum y^2 = \sum Y^2 - n\bar{Y}^2$
4. พิจารณาข้อมูลต่อไปนี้ซึ่งเป็นข้อมูลเกี่ยวกับปริมาณเนื้อวัว ในตลาดส่งออกในช่วง 10 ปี กับราคาเนื้อวัวในตลาดโลก

ปี	2530	2531	2532	2533	2534	2535	2536	2537	2538	2539
ปริมาณ(ก.ก.)	5	4	3	4	7	9	8	10	8	2
ราคา (บาท/หน่วย)	2	4	2	3	8	7	6	8	7	3

- ก. วาดกราฟแสดงการกระจายของข้อมูลดังกล่าว
  - ข. ลากเส้นตรงด้วยตาเปล่าแล้วใช้เส้นตรงนั้นคำนวณค่ามากที่สุดสำหรับจุดตัดและความชันในความสัมพันธ์แบบเส้นตรงระหว่างปริมาณและราคา
  - ค. ใช้วิธีใดก็ได้คำนวณค่าประมาณของค่าพารามิเตอร์ของเส้นถดถอยกำลังสองน้อยที่สุดของปริมาณและราคา
  - ง. ตรวจสอบคำตอบกับข้อ ค. ด้วยการคำนวณสัมประสิทธิ์การถดถอยโดยใช้วิธีอื่น
5. คุณสามารถใช้การถดถอยแบบเส้นตรงในการประมาณค่าพารามิเตอร์ของความสัมพันธ์ต่อไปนี้ได้อย่างไร

ก.  $y = AL^aK^b$

ข.  $C = a + b_1W + b_2Z + b_3Z^2 + b_4WZ$