



บทที่ 2

งานเขียนที่เกี่ยวข้อง

ปัญหาการอ่านหนังสือไม่ออกการเขียนหนังสือไม่ได้ มีใช่เป็นปัญหาเฉพาะประเทศไทยเท่านั้น หากแต่เป็นปัญหาของประเทศต่าง ๆ เกือบทั่วโลก ตามประมาณการขององค์การศึกษาวิทยาศาสตร์และวัฒนธรรมแห่งสหประชาชาติ (ยูเนสโก) เมื่อปี 2513 ทั่วโลกมีประชากรราว 810 ล้านคน หรือร้อยละ 34.8 ที่อายุ 15 ปีขึ้นไปอ่านหนังสือไม่ออก¹ ซึ่งนับว่าเป็นปัญหาใหญ่จึงได้มีการประชุมปรึกษาหารือกันในระดับชาติหลายต่อหลายครั้ง เพื่อหาทางร่วมมือกันระหว่างประเทศในการที่จะหาทางแก้ไขปัญหานี้ การประชุมที่สำคัญที่สุดเกี่ยวกับการขจัดการไม่รู้หนังสือ ได้แก่ การประชุมรัฐมนตรีศึกษาธิการทั่วโลกที่กรุงเทพมหานคร ประเทศอิหร่าน (The World Congress of Ministers of Education on the Eradication of Literacy) เมื่อเดือนกันยายน 2508 ซึ่งยูเนสโกจัดขึ้นด้วยความร่วมมือของพระเจ้าชาห์แห่งอิหร่าน ประเทศไทยได้ส่งคณะผู้แทนเข้าร่วมในการประชุมสัมมนาครั้งนั้นด้วย ที่ประชุมได้พิจารณาและยอมรับว่า การสอนให้รู้หนังสืออย่างเกี่ยวไม่เพียงพอควรสอนให้การเรียนรู้หนังสือสัมพันธ์กับการพัฒนาเศรษฐกิจ สังคม และชีวิตความเป็นอยู่ของประชากรของประเทศด้วยโดยเสนอให้มีการสอนค่านาอาชีพควบคู่กันไปกับการสอนอ่านและเขียน เพื่อที่ผู้เรียนจะได้นำความรู้ที่ได้รับ ไปใช้ให้เป็นประโยชน์ในการปรับปรุงอาชีพ และความเป็นอยู่ของตนเองตลอดจนชุมชนที่ตนอาศัยอยู่ได้อย่างมีประสิทธิภาพ การสอนแบบนี้ที่ประชุมเรียกว่าการศึกษาแบบเบ็ดเสร็จ (Work Oriented Functional Literacy Program)²

¹ Unesco Regional office for Education in Asia, Study Visit and Seminar : Work Oriented Adult Literacy Pilot Project in Iran, Bangkok, 1970 . p. 29.

² ibid. pp. 16 - 22.

การศึกษาเพื่อแก้ไขการไม่รู้หนังสือในปัจจุบัน ได้มีบทบาทต่อบุคคลและสังคมมากขึ้นตามลำดับ ประเทศที่กำลังรีบเร่งพัฒนาตางมองเห็นว่าการแก้ไขผู้ใหญ่ที่ไม่รู้หนังสือเป็นเรื่องสำคัญ เพราะว่าผู้ใหญ่คือ ผู้ที่กำลังมีบทบาทในปัจจุบัน สามารถประกอบอาชีพในสังคมและพร้อมที่จะพัฒนาบ้านเมืองได้ การศึกษาภาคบังคับที่ให้แก่เด็กนั้นจะเห็นผลช้า และถ้าเด็กออกจากโรงเรียนไปแล้ว เมื่อไม่มีความจำเป็นจะต้องใช้หนังสือไม่ช้าก็ลืม จากการวิจัยของกรมสามัญศึกษาพบว่า "..... เด็กที่จบชั้นประถมศึกษาปีที่ 4 ถ้าไม่ได้อ่านหนังสือชั่วระยะเวลาหนึ่งก็มีโอกาสลืมหนังสือถึงร้อยละ 33.0"¹ แสดงว่าได้เกิดความสูญเสียไปทางการศึกษาอย่างมหาศาล นายเกรียง กิริติกกร เมื่อครั้งดำรงตำแหน่งอธิบดีกรมสามัญศึกษาได้ให้ทัศนะในเรื่องนี้ว่า "สำหรับผู้ที่รู้หนังสืออยู่บ้างแล้วหรือผู้ที่จบชั้นประถมศึกษาปีที่ 4 แล้วก็อาจลืมหนังสือ เนื่องจากเมื่อจบออกไปแล้วไม่ได้อ่านหนังสือต่อไป คงเป็นเพราะว่าไม่มีหนังสืออ่าน ถ้าเราไม่ช่วยคนพวกนี้ก็จะ เป็นปัญหาสำคัญต่อไปในอนาคต"²

โครงการแก้ไขการไม่รู้หนังสือแบบเบ็ดเสร็จแบบไทยนี้เริ่มโครงการทดลองที่จังหวัดลำปางและแพร่ เมื่อเดือนพฤษภาคม 2514 จังหวัดละ 10 โรงเรียน โดยมีเป้าหมายว่า ภายในระยะเวลา 6 เดือน หรือ 180 - 200 ชั่วโมง ผู้เรียนจะเกิดความคิดรวมนั้นในลักษณะของกลุ่ม รู้จักรับฟังความคิดเห็นของคนอื่น ตลอดจนใช้เหตุผลหลายอย่างและข้อมูลจากหลายด้านประกอบการตัดสินใจ อันเป็นการพัฒนาตัวบุคคลโดยตรงก่อนที่จะเข้าไปเป็นปัจจัยหลักในการพัฒนากิจกรรมอื่น ๆ อันจะนำมาซึ่งการพัฒนาตนเอง ชุมชน และประเทศชาติอันเป็นส่วนรวม ส่วนทักษะในการอ่านออกเขียนได้นั้น

¹ กรมสามัญศึกษา กระทรวงศึกษาธิการ, รายงานการวิจัย : สภาพการรู้หนังสือของผู้จบชั้นประถมศึกษาปีที่ 4 (พระนคร : คุรุสภา, 2512), หน้า 40.

² เกรียง กิริติกกร, "ปรัชญาและนโยบายในการจัดการศึกษาผู้ใหญ่", วารสารการศึกษาผู้ใหญ่, 9 : 3 (พฤษภาคม - มิถุนายน, 2515), หน้า 21.

เป็นเป้าหมายอันค้ำรอง¹ / การดำเนินการทดลองในสองจังหวัดนี้ เมื่อมีการวิจัยเพื่อ
 ประเมินผลโครงการทดลอง โดยใช้กลุ่มตัวอย่างของนักศึกษา 252 คนจากโรงเรียน
 ผู้ใหญ่แบบเบ็ดเสร็จ 20 โรงเรียน เป็นกลุ่มทดลองและกลุ่มตัวอย่างของนักศึกษาผู้ใหญ่จาก
 โรงเรียนที่สอนตามหลักสูตรการศึกษาผู้ใหญ่ขั้นพื้นฐาน ซึ่งเป็นหลักสูตรเดิม 2 โรงเรียน
 ในจังหวัดลำปาง และแพร่จังหวัดละ 1 โรงเรียน เป็นกลุ่มควบคุม มีการเก็บรวบรวมข้อมูล
 2 ครั้ง คือ ครั้งแรกเมื่อเริ่มเปิดสอนและครั้งหลังเมื่อสิ้นสุดการสอน โดยใช้แบบทดสอบ
 2 ฉบับ คือ แบบทดสอบความสามารถในการอ่านและแบบทดสอบความสามารถในการคิดเลข
 ซึ่งเรียงลำดับจากง่ายไปหายาก นอกจากนี้ยังมีแบบวัดทัศนคติและความคิดเห็นในเรื่อง
 การทำมาหากิน การอนามัย และการวางแผนครอบครัว เศรษฐกิจ และหน้าที่พลเมือง
 โดยสัมภาษณ์ให้นักศึกษาตอบ "เห็นด้วย" หรือ "ไม่เห็นด้วย" และวิเคราะห์ข้อมูล
 โดยวิธีหาค่าความแตกต่างระหว่างการสอบครั้งแรกและการสอบครั้งหลัง แล้วหาค่าเฉลี่ย
 ของทั้งสองกลุ่ม ทดสอบความแตกต่างของค่าเฉลี่ยด้วยการทดสอบค่าที (t - test)²
 ซึ่งพอสรุปผลการวิเคราะห์ได้ดังนี้³

ความสามารถในการอ่าน

1. นักศึกษาผู้ใหญ่ในโครงการ มีความสามารถในการอ่านเขียนในระดับ
 สูงกว่ากลุ่มควบคุม ซึ่งสอนตามหลักสูตรเดิม (การศึกษาขั้นมูลฐาน) แต่ความแตกต่างไม่มี

¹ กองการศึกษาผู้ใหญ่ กรมสามัญศึกษา, "โครงการจัดการไม่รู้นั่งสี่
 แบบเบ็ดเสร็จในประเทศไทย," 2514. (อักสำเนา)

² กองการศึกษาผู้ใหญ่ กรมสามัญศึกษา, รายงานการประเมินผล -
 โครงการการศึกษาผู้ใหญ่แบบเบ็ดเสร็จของจังหวัดลำปาง และแพร่, 2514.

(อักสำเนา)

³ ฉลอง บุญญานันท์, สรุปรายงานผลของการทดลองการสอนนักศึกษา-
 ผู้ใหญ่หลักสูตร เบ็ดเสร็จ, กองการศึกษาผู้ใหญ่ 2514. (อักสำเนา)



นัยสำคัญทางสถิติ

2. ความสามารถในการอ่านของนักศึกษาผู้ใหญ่ในโครงการ เพิ่มขึ้นจากระดับ 3.95 ในการทดสอบครั้งแรกเป็น 6.42 ในการทดสอบครั้งหลัง

3. ความสามารถในการอ่านโดยเฉลี่ย เมื่อสิ้นสุดโครงการแล้วอยู่ในระดับ 6.42 แสดงว่านักศึกษาผู้ใหญ่โดยเฉลี่ยพอมีความสามารถในการอ่านประโยคง่าย ๆ ได้เท่านั้น

ความสามารถในการคิดเลข

1. ความสามารถในการคิดเลขของนักศึกษาผู้ใหญ่ในโครงการสูงกว่ากลุ่มควบคุม คือ ระหว่าง 6.53 กับ 4.87 แต่ความแตกต่างไบนัยสำคัญทางสถิติ

2. การสอนตามหลักสูตรโครงการทำให้ความสามารถทางการคิดเลขสูงขึ้นคือ จากระดับ 3.79 เป็น 6.53

3. ความสามารถในการคิดเลข โดยเฉลี่ยเมื่อสิ้นสุดโครงการอยู่ในระดับ 6.53 หมายความว่า นักศึกษาผู้ใหญ่มีความสามารถคิดเลขโจทย์ บวก ลบ ง่าย ๆ ได้

ส่วนในด้านการศึกษาเปลี่ยนแปลงทัศนคติพบว่า นักศึกษาผู้ใหญ่ ส่วนใหญ่เปลี่ยนแปลงทัศนคติไปมาก กล่าวคือ ในจำนวนความคิดเห็นทั้งหมด 22 ข้อ ผู้เรียนเปลี่ยนแปลงทัศนคติไปในทางที่ต้องการอย่างมีนัยสำคัญถึง 18 ข้อ ส่วนอีก 4 ข้อ ผู้เรียนมีทัศนคติที่ถูกต้องอยู่แล้ว ตั้งแต่ต้น จึงมีการเปลี่ยนแปลงเล็กน้อย แต่ไม่มีนัยสำคัญทางสถิติ

และจากการวิจัยของ นางสาวสมทรง อิศวกุล¹ เกี่ยวกับระดับ การรู้หนังสือของผู้สำเร็จการศึกษาจากโครงการการศึกษาผู้ใหญ่แบบเบ็ดเสร็จ โดยใช้ ผู้สำเร็จการศึกษาจากโครงการแก้ไขการไม่รู้หนังสือแบบเบ็ดเสร็จจังหวัดขอนแก่น จำนวน 50 คน และสุ่มนักเรียนที่จบชั้นประถมศึกษาปีที่ 2, 3 และ 4 จากโรงเรียน- สังกัดองค์การบริหารส่วนจังหวัด อำเภอมืองขอนแก่น กลุ่มละ 200 คน ทดสอบ กล้วยแบบทดสอบมาตรฐานวิชาเลขคณิตฉบับที่ 1 ก. ฉบับที่ 2 ก. และวิชาภาษาไทย ของกองการประถมศึกษา กรมสามัญศึกษา ซึ่งมีสัมประสิทธิ์ของความเที่ยง 0.877, 0.749 และ 0.854 ตามลำดับ วิเคราะห์ข้อมูลโดยการหาคะแนนเฉลี่ย ขอบเขต ของคะแนนเฉลี่ยของประชากรที่ระดับความเชื่อมั่น 95 % และการทดสอบค่า ที่ (t - test) ผลปรากฏว่าระดับการรู้หนังสือแบบเบ็ดเสร็จมีขอบเขตของคะแนนเฉลี่ย เป็น 34.39 % ถึง 42.50 % ซึ่งต่ำกว่านักเรียนที่จบชั้นประถมศึกษาปีที่ 4 แต่สูงกว่านักเรียนที่จบชั้นประถมศึกษาปีที่ 3 และถ้าพิจารณาแยกเป็นรายวิชาพบว่า มีความรู้ในวิชาภาษาไทยใกล้เคียงกันกับนักเรียนที่จบชั้นประถมศึกษาปีที่ 4 แต่ความรู้ วิชาเลขคณิตต่ำกว่า

ขณะเดียวกัน นางกุหลาบ หวังคีรีกุล² ได้ทำศึกษาวิจัยเกี่ยวกับ

¹สมทรง อิศวกุล , " ระดับการรู้หนังสือของผู้สำเร็จการศึกษาจาก โครงการแก้ไขการไม่รู้หนังสือแบบเบ็ดเสร็จ , " วิทยานิพนธ์ ปริญญาโทบริหารศึกษาด้านศึกษาศาสตร์ คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, 2516. (อัครสำเนา)

²กุหลาบ หวังคีรีกุล , " ความคงอยู่ของการรู้หนังสือของผู้จบการศึกษา จากโครงการแก้ไขการไม่รู้หนังสือแบบเบ็ดเสร็จ " วิทยานิพนธ์ ปริญญาโทบริหารศึกษาด้านศึกษาศาสตร์ คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, 2517. (อัครสำเนา)

ความคงอยู่ของการรู้หนังสือของผู้ที่จบการศึกษาจากโครงการ แก่ไขการไม่รู้หนังสือแบบเบ็ดเสร็จไปแล้ว 3 ปี รวมทั้งการศึกษาถึงอิทธิพลของที่อ่านหนังสือพิมพ์ประจำหมู่บ้าน ทอความคงอยู่ของการรู้หนังสือดังกล่าวด้วย โดยศึกษาจากผู้จบการศึกษาจากโครงการแก่ไขการไม่รู้หนังสือแบบเบ็ดเสร็จ จังหวัดลำปางและแพร่ ปีการศึกษา 2514 จำนวน 140 คน เปรียบเทียบกับผู้จบชั้นประถมศึกษาปีที่ 4 จากโรงเรียนสังกัดองค์การบริหารส่วนจังหวัดลำปาง และแพร่ ปีการศึกษาเดียวกันจำนวนเท่ากัน ทดสอบควยวิชาเลขคณิต และภาษาไทย ซึ่งมีสัมประสิทธิ์แห่งความเที่ยงเป็น 0.74 และ 0.85 ตามลำดับ วิเคราะห์ข้อมูลโดยการหาค่าเฉลี่ย ร้อยละ ส่วนเบี่ยงเบนมาตรฐานและวิเคราะห์ความแปรปรวนแบบมี 3 ตัวประกอบ ($2 \times 2 \times 2$ Factorial Design) คือ ประเภทการศึกษา การมีที่อ่านหนังสือพิมพ์ และวิชาตามลำดับ ผลปรากฏว่า ผู้จบการศึกษาจากโครงการฯ มีความคงอยู่ของการรู้หนังสือค่ากว่าผู้จบชั้นประถมศึกษาปีที่ 4 อย่างมีนัยสำคัญ 0.05 ผู้จบการศึกษาจากโครงการฯ และผู้จบชั้นประถมปีที่ 4 ที่อยู่ในถิ่นที่มีที่อ่านหนังสือพิมพ์ประจำหมู่บ้าน มีความคงอยู่ของการรู้หนังสือสูงกว่าผู้อยู่ในถิ่นที่ไม่มีที่อ่านหนังสือพิมพ์ประจำหมู่บ้านอย่างมีนัยสำคัญ 0.001 นั่นคือ ที่อ่านหนังสือพิมพ์ประจำหมู่บ้านมีผลทำให้ความคงอยู่ของการรู้หนังสือของผู้จบการศึกษาจากโครงการฯ และผู้จบชั้นประถมศึกษาปีที่ 4 สูงขึ้น

* แมวการศึกษาผู้ใหญ่แบบเบ็ดเสร็จ จะสามารถช่วยเหลืองานด้านการศึกษาของประเทศได้ไม่มากนักเมื่อนักถึงจำนวนผู้เรียนในปัจจุบันและปัญหาที่จะต้องรณรงค์กันต่อไป แต่ก็เป็นแนวการจัดการศึกษาที่มีระบบอันสำคัญน่าเชื่อถือ เนื่องด้วยวิธีการเข้าถึงประชาชนแบบใหม่ในลักษณะของการศึกษานอกโรงเรียน และเป็นที่เด่นชัดว่าโครงการนี้จะสามารถสนองตอบหรือเพิ่มเติมความต้องการของประชาชน ในการนำความรู้ที่ไ้ไปใช้แก้ปัญหาในชีวิตประจำวันได้เป็นอย่างดี และสิ่งที่สำคัญที่สุดของการศึกษาผู้ใหญ่แบบเบ็ดเสร็จก็คือ เป็นวิธีการจัดการศึกษาที่ฝึกให้คนมีนิสัยเป็นคนคิด รู้จักคิดแก้ปัญหาโดยใช้กระบวนการ (process) ที่ถูกต้อง เป็นคนรู้จักวิถีชีวิตไม่เพิ่มพูนการศึกษาเพื่อหาแนวทางแก้ปัญหาของเขา ตลอดจนเป็นสมาชิกที่มีความรับผิดชอบต่องังคมประชาธิปไตย



การวิเคราะห์ถดถอยพหุคูณ

การวิเคราะห์ถดถอยพหุคูณ¹ เป็นเทคนิควิธีที่ขยายจากแนวความคิดเดิมของความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตามกับการพยากรณ์ค่าของตัวแปรดังกล่าวจากสมการถดถอยที่ได้ การถดถอยพหุคูณศึกษาถึงบทบาทหรืออิทธิพลและขนาดของอิทธิพลของตัวแปรอิสระที่มากกว่า 1 ตัวแปรขึ้นไปที่มีต่อตัวแปรตาม โดยใช้วิธีการหรือหลักความสัมพันธ์ (correlation) และการถดถอย (regression) หน้าที่ประการสำคัญของการถดถอยพหุคูณก็คือการช่วย "อธิบาย" ความแปรปรวนหรือการกระจายของตัวแปรตาม โดยประเมินจากการรู้ความแปรปรวนของตัวแปรอิสระ งานวิจัยทางการศึกษา ส่วนใหญ่พยายามที่จะแสวงหาค่าอธิบายความแปรปรวนของสัมฤทธิ์ผลทางการเรียน โดยศึกษาสิ่งที่มีความสัมพันธ์ต่อสัมฤทธิ์ผลนั้นหลายประการ เช่น เขวามันปัญญา ความถนัด สังคมในชั้นเรียน เชื้อชาติ สิ่งแวดล้อมทางบ้าน บรรยากาศในโรงเรียน คุณลักษณะของครู และอื่น ๆ ทางด้านรัฐศาสตร์ ในการศึกษาเพื่ออธิบายพฤติกรรม การลงคะแนนเสียงเลือกตั้งจำเป็นต้องศึกษาตัวแปรที่คิดว่าจะมีอิทธิพลต่อพฤติกรรมดังกล่าว เช่น อายุ เพศ รายได้ การศึกษา แรงจูงใจ ที่พำนัก เป็นต้น

วัตถุประสงค์ของการวิเคราะห์ถดถอยพหุคูณก็เพื่อที่จะคำนวณค่าน้ำหนัก (beta weight) หรือสัมประสิทธิ์การถดถอย (regression coefficient) ของตัวแปรอิสระแต่ละตัว ตลอดจนค่าคงที่เฉพาะของสมการ (beta zero หรือ intercept) เพื่อให้ผลที่คำนวณได้จากสมการถดถอย $Y = b_0 + b_1 X_1 + b_2 X_2 + \dots + b_k X_k$ เป็นค่าพยากรณ์ที่มีค่ากำลังสองของความคลาดเคลื่อนในการพยากรณ์น้อยที่สุด

¹ อารุง จันทวนิช , การวิเคราะห์ถดถอยพหุคูณ กองวิจัยการศึกษา สำนักงานคณะกรรมการการศึกษาแห่งชาติ , 2519. (อัคราเนา)

จากความรู้ค่าตัวแปรอิสระ $X_1, X_2, X_3, \dots, X_k$ ในกรณีที่มีตัวแปรอิสระ k ตัวแปร เราก็สามารถทำนายหรือหาค่าของตัวแปรตาม (Y) ได้ ตัวอย่าง เช่น งานวิจัยของ โฮลต์แมน และ บราวน์ (Holtzman and Brown, 1968)¹ ซึ่งทำการศึกษาเพื่อพยากรณ์คะแนนเฉลี่ยของนักเรียนระดับมัธยมศึกษา (GPA) โดยพิจารณาให้คะแนนลักษณะนิสัยในการเรียนและทัศนคติของนักเรียน (SHA) ตลอดจนคะแนนความถนัดทางการเรียน (SA) เป็นตัวแปรอิสระ ในการศึกษาพบว่าค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนเฉลี่ยในการเรียนระดับมัธยมศึกษา และคะแนนลักษณะนิสัย ตลอดจนทัศนคติของนักเรียนเป็น 0.55 และค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนเฉลี่ยของนักเรียนระดับมัธยมศึกษาและคะแนนความถนัดทางการเรียนเป็น 0.61 ผลจากการวิเคราะห์ที่ใช้ตัวแปรอิสระคะแนนลักษณะนิสัยในการเรียนและทัศนคติของนักเรียน (SHA) เพียงตัวแปรเดียว พบว่า สามารถอธิบายค่าความแปรปรวนของคะแนนเฉลี่ยนักเรียนมัธยมศึกษา (GPA) ได้ร้อยละ 30 ($r_1^2 = 0.30$) และถ้าใช้ตัวแปรความถนัดทางการเรียน (SA) ตัวแปรเดียว เช่นกันจะสามารถอธิบายค่าความแปรปรวนของคะแนนเฉลี่ยนักเรียนได้ร้อยละ 37 ($r_2^2 = 0.37$) ทั้งนี้ เนื่องจากคะแนนลักษณะนิสัยในการเรียนและทัศนคติของนักเรียน (SHA) ตลอดจนคะแนนความถนัดทางการเรียน (SA) มีการคาบเกี่ยวกันในค่าความแปรปรวนของการอธิบาย เพราะตัวแปรทั้งสองมีความสัมพันธ์กันเท่ากับ 0.30 ทำให้ไม่สามารถหาค่าความแปรปรวนที่ตัวแปรอิสระ SHA และ SA อธิบายร่วมกันโดยการรวมค่า ที่ได้ทั้งสองค่าเข้าด้วยกันโดยตรงจากการใช้วิธีการวิเคราะห์ถดถอยพหุคูณ โฮลต์แมน และ บราวน์ พบว่า คะแนนนิสัยในการเรียน และทัศนคติของนักเรียน (SHA) กับคะแนนความถนัดทางการเรียน (SA)

¹ W.H. Holtzman and W.F. Brown, "Evaluating Study Habits and Attitudes of High School Students", Journal of Educational Psychology, vol. 59, 1968, pp. 404-409.

รวมกันอธิบายค่าความแปรปรวนของคะแนนเฉลี่ยของนักเรียนระดับมัธยมศึกษา (GPA) ได้ถึงร้อยละ 52 ซึ่งจากการวิจัยนี้สามารถสรุปได้ว่า การใช้ตัวแปรอิสระร่วมในการพยากรณ์ตัวแปรตาม สามารถเพิ่มอำนาจในการพยากรณ์ให้สูงขึ้นจากเดิม อันแสดงให้เห็นถึงคุณประโยชน์ และการได้เปรียบในการใช้วิธีการวิเคราะห์หาคดถอยพหุคูณเพื่อพยากรณ์สัมฤทธิ์ผลทางการเรียนของนักเรียน

ลักษณะของการวิเคราะห์หาคดถอยพหุคูณ



ตัวแปรสำหรับการวิเคราะห์

การศึกษาคาดถอยพหุคูณเพื่อการพยากรณ์ต้องใช้ตัวแปรอิสระจำนวนหนึ่งจากตัวแปรที่ต้องการศึกษาทั้งหมด วิธีการนี้จะใช้โคถึกรายเท่าที่ข้อมูลที่ทำการศึกษาเป็นข้อมูลที่มีลักษณะต่อเนื่อง (continuous data) เป็นข้อมูลที่จัดเรียงลำดับ (ranked data) หรือเป็นข้อมูลที่จัดเป็นพวกเป็นหมู่ (categorical data) โดยที่ในแต่ละกลุ่มหรือระดับของข้อมูลที่จัดเป็นพวกเป็นหมู่นั้น ประกอบด้วยตัวเลข 2 ประเภท คือ เลข 1 และเลข 0 ในการศึกษาข้อมูลที่จัดเป็นพวกเป็นหมู่ โดยเฉพาะอย่างยิ่งข้อมูลที่สามารจัดแบ่งได้หลาย ๆ ระดับ อาจใช้ข้อมูลในแต่ละระดับเป็นตัวแปรอิสระสำหรับการวิเคราะห์หาคดถอยพหุคูณโดยเลือกระดับของตัวแปรจำนวนเท่ากับชั้นแห่งความเป็นอิสระ (degree of freedom) ของตัวแปรนั้น ๆ ตัวอย่างเช่น ศึกษาตัวแปรเกี่ยวกับอายุที่ประกอบด้วยกลุ่มอายุ 5 ระดับ ต้องใช้ระดับของกลุ่มอายุในการศึกษาจำนวน 4 ระดับ ใช้ข้อมูลทั้ง 4 ระดับนี้เป็นตัวแทนสำหรับตัวแปรที่เกี่ยวกับอายุทั้งหมดในการวิเคราะห์หาคดถอยพหุคูณ อีกวิธีหนึ่งอาจทำได้โดยการศึกษาคตัวแปรที่

006047

ต้องการศึกษาทุกระดับและกำหนดตัวเลขให้คุณลักษณะของระดับที่ต้องการศึกษาในแต่ละระดับมีค่าเป็น 1 ส่วนคุณลักษณะอื่นที่ไม่ต้องการศึกษามีค่าเป็น 0

สำหรับการศึกษาในครั้งนี้ ผู้วิจัยเลือกใช้วิธีการอันหลังเพื่อการเลือกตัวแปรอิสระที่เป็นพวกเป็นหมู่ เนื่องจากวิธีแรกจำเป็นต้องคัดระดับของตัวแปรบางระดับออกไปโดยที่ผลของระดับในตัวแปรที่คัดออกไปนั้นจะไปรวมอยู่ที่ค่าคงที่เฉพาะของสมการในสมการถดถอยที่ได้ และเนื่องจากการศึกษาระดับของตัวแปรที่คัดเลือกไปพร้อม ๆ กัน โดยการใช้กระบวนการเลือกตัวแปรนั้นมีความลำบากในการหาโปรแกรมการคำนวณสำเร็จรูป (software package) เป็นอย่างยิ่ง อีกประการหนึ่งการศึกษาระดับของตัวแปรที่ต้องการศึกษาทุกระดับในข้อมูลที่จัดเป็นพวกเป็นหมู่ โดยการกำหนดค่าของระดับที่ต้องการศึกษา (1 และ 0) นั้นเป็นวิธีการที่ใช้ได้ก็สำหรับการศึกษาตัวแปรอิสระที่มีลักษณะปนกันหลาย ๆ อย่าง (mixed mode independent variables) กล่าวคือ อาจเป็นตัวแปรที่มีลักษณะต่อเนื่อง ตัวแปรที่จัดอันดับ และตัวแปรที่เป็นพวกเป็นหมู่ได้ในเวลาเดียวกัน¹

ข้อตกลงเบื้องต้นของการวิเคราะห์ถดถอยพหุคูณ

ในการวิเคราะห์ถดถอยพหุคูณเรามีข้อตกลงเบื้องต้นหลายประการที่จะต้องพิจารณาประกอบก่อนลงมือทำการวิเคราะห์ คือ

1. การวิเคราะห์ถดถอยในทุกลักษณะ จะสมมุติว่า คะแนนของตัวแปรตาม (Y) มีการแจกแจงเป็นโค้งปกติในแต่ละค่าของตัวแปรอิสระ (X) ข้อสมมุตินี้เป็นสิ่งจำเป็นสำหรับการทดสอบนัยสำคัญทางสถิติ F

¹ E.F. Gocka, " Stepwise Regression for Mixed - Mode Predictor Variable" , Education and Psychology Measurement, 1973 , pp. 319 - 325.

2. คะแนนตัวแปรตาม (Y) ที่ไ้มีความแปรปรวน (variance) เท่ากันในทุก ๆ ค่าของตัวแปรอิสระ (X) กล่าวคือ ค่า Y ณ X ใดถือว่าเป็น ตัวแทนสุ่มมาจากประชากรปกติ โดยที่ทุก ๆ ประชากรมีการกระจายร่วมกันอยู่คือ $\sigma_{Y \cdot X}^2$

3. ค่าความผิดพลาดแต่ละค่า (error terms) ที่ไ้จากการคำนวณในสมการถดถอย เป็นค่าที่ไ้โดยการสุ่ม (random) และมีการแจกแจงเป็นโค้งปกติ ข้อสมมุตินี้ใช้ประกอบในการคาดคะเนทางสถิติ (statistical estimation)

คุณสมบัติของสมการถดถอยเส้นตรง

คุณสมบัติของสมการถดถอยแบบที่เป็นเส้นตรงมีดังนี้คือ

1. ค่าเฉลี่ยของตัวแปรอิสระ (\bar{X}) และค่าเฉลี่ยของตัวแปรตาม (\bar{Y}) จะอยู่บนเส้นถดถอย
2. ผลรวมของผลต่าง (deviation) ระหว่างจุดที่สังเกตไ้กับค่าเฉลี่ยของประชากรมีค่าเป็นศูนย์
3. ผลรวมของกำลังสองของผลต่างระหว่างจุดที่สังเกตไ้กับค่าเฉลี่ยมีค่าต่ำสุด

ข้อดีของการวิเคราะห์ถดถอยพหุคูณ

ประโยชน์ที่เห็นไ้ชัดของการวิเคราะห์ถดถอยพหุคูณ หรืออาจกล่าวไ้ว่าเป็นข้อดีที่สำคัญที่สุดของวิธีการนี้ก็คือ วิธีการดังกล่าวมีส่วนสัมพันธ์อย่างใกล้ชิดและสอดคล้องกับวัตถุประสงค์ทางวิทยาศาสตร์ นั่นคือ การอธิบายปรากฏการณ์ทาง - ชุมชนชาติในการวิจัยส่วนมากความพยายามอย่างน้อยที่สุดก็มุ่งที่จะอธิบายเหตุการณ์หรือปรากฏการณ์อันหนึ่งอันใดที่เกิดขึ้น แม้ว่าปรากฏการณ์นั้นจะมีลักษณะปะปนกันหรือสับสน -

เพียงใดก็ตาม ดังนั้น การวิเคราะห์ถดถอยพหุคูณ แมแต่ที่เกี่ยวข้องกับตัวแปรตาม เพียงตัวแปร ก็เกี่ยวข้องกับวิธีการทางวิทยาศาสตร์

ข้อดีประการที่สองของการวิเคราะห์ถดถอยพหุคูณ ก็คือ สามารถที่จะใช้กับตัวแปรอิสระ จำนวน หรือประเภทใดก็ได้ วิธีการวิเคราะห์ที่ให้ผลในทุกๆ ด้านตามที่การวิเคราะห์ความแปรปรวน (analysis of variance) จะทำได้และให้ผลมากกว่า ไม่ว่าตัวแปรอิสระจะมีลักษณะหรือมาตรการวัด (scale) ใดๆก็ตาม โดยเฉพาะอย่างยิ่ง การวิเคราะห์ถดถอยพหุคูณ เป็นวิธีที่ดีที่สุดสำหรับข้อมูลที่ไม่ได้ผลจากการทดลอง (nonexperimental data)

ข้อดีประการสุดท้ายของวิธีการวิเคราะห์แบบนี้ก็คือ การวิเคราะห์ถดถอยพหุคูณยังใช้วิธีการทางสถิติหลายวิธีในการพิจารณาแปลความหมายข้อมูล ตัวอย่างเช่น วัดความสัมพันธ์ (correlation) ทั้งหลายที่เกิดขึ้นในข้อมูลระหว่างตัวแปรอิสระและระหว่างตัวแปรอิสระกับตัวแปรตาม ค่า R^2 ใช้ในการคาดคะเนอัตราส่วนของความแปรปรวนในตัวแปรตาม ซึ่งเป็นผลมาจากการรู้ความแปรปรวนของตัวแปรอิสระ นอกจากนั้น การทดสอบนัยสำคัญทางสถิติ F ช่วยให้เห็นความแตกต่างในค่าของ R^2 หรือลักษณะความมีนัยสำคัญของการเพิ่มความรู้ในการกระจายของตัวแปรตาม อันเป็นผลมาจากการนำตัวแปรอิสระบางตัวตัวแปรเพิ่มในสมการถดถอย เป็นต้น

จุดอ่อนของการวิเคราะห์ถดถอยพหุคูณ

ในการวิเคราะห์ถดถอยพหุคูณ ความเชื่อถือได้ของผลการวิเคราะห์เป็นปัญหาที่สำคัญของวิธีการดังกล่าว ปัญหาประการแรกเกิดจากการที่ผู้วิจัยมีแนวโน้มที่จะโยนตัวแปรอิสระทั้งหมดที่ต้องการศึกษา หรือต้องการเลือกให้กับวิธีการวิเคราะห์ถดถอยพหุคูณโดยไม่มีการเลือกสรรตัวแปรอิสระเหล่านั้นเสียก่อน แต่ปล่อยให้วิธีการวิเคราะห์และเครื่องคำนวณไฟฟ้าทำหน้าที่เองตามลำพัง เปรียบเสมือนการยิงนกด้วยกระสุนลูกปราย โดยคิดว่าน่าจะถูกเข้าบ้าง การกระทำดังกล่าวไม่เหมาะสม -

อย่างยิ่งและชดกับหลักการที่ควรจะเป็นของการทำงานวิจัย นั่นก็คือ ก่อนที่ผู้วิจัย จะใช้ตัวแปรอิสระจำนวนมาก ๆ เพื่อทำการวิเคราะห์ถดถอยพหุคูณ ควรจะไต่พยายาม ลดจำนวนตัวแปรอิสระนั้นลงโดยอาศัยหลักทางทฤษฎี หรือวิธีการวิเคราะห์ที่ประกอบ เข้ารายชั้นหนึ่งก่อน

ข้อบกพร่องประการต่อไปของการวิเคราะห์ถดถอยพหุคูณอยู่ที่ความ ไม่เที่ยง (unreliability) ของค่าสัมประสิทธิ์การถดถอย (regression coefficient) เมื่อมีการนำตัวแปรอิสระเข้าในสมการถดถอยหรือออกจากสมการ ถดถอย ค่าสัมประสิทธิ์การถดถอยคงกล่าวจะไม่คงที่ และจะเปลี่ยนไปตามขนาด ของกลุ่มตัวอย่างและตัวแปรอิสระ แต่ปัญหานี้อาจลดลงได้หรือเกิดขึ้นน้อยหากใช้กลุ่มตัวอย่าง ขนาดใหญ่และใช้ตัวแปรอิสระที่ผ่านการคัดเลือกแล้วจำนวนน้อย ๆ สำหรับในกรณีนี้ ตัวแปรอิสระมีความสัมพันธ์ระหว่างกันและกันสูงนั้น ตัวแปรอิสระดังกล่าวจะวัดหรือ อธิบายในสิ่งเดียวกัน ผลที่ตามมาก็คือ ค่าความผิดพลาดในการวัด (standard error) ของสัมประสิทธิ์การถดถอยจะเพิ่มขึ้นอย่างเห็นได้ชัดและจะไม่มี ความหมายใด ๆ ในการแปรผลการอธิบายความแปรปรวนของตัวแปรตาม และผลอีกประการ หนึ่งของการที่ตัวแปรอิสระมีความสัมพันธ์ระหว่างกันสูง ก็คือ ทำให้เกิดปัญหา multicollinearity ซึ่งนอกจากจะก่อให้เกิดความลำบากในการอธิบายความ-แปรปรวนของตัวแปรตามแล้ว ยังทำให้ค่าสัมประสิทธิ์สหสัมพันธ์ส่วนย่อย (partial correlation coefficient) ที่ใช้สำหรับจะคัดเลือกตัวแปรอิสระเพื่อเข้าสู่ สมการถดถอยเกิดความคลาดเคลื่อนอีกด้วย

การวิเคราะห์ถดถอยพหุคูณแบบเพิ่มหรือลดตัวแปรเป็นขั้น ๆ

(Stepwise Multiple Regression Analysis)

ในการเลือกสมการถดถอยเพื่อการพยากรณ์ที่ดีที่สุดนั้น มีเกณฑ์การเลือก

๒. ประเภทที่เป็นข้อคิดสำหรับการพิจารณา¹ คือ

1. ให้สมการที่สร้างมีประสิทธิภาพสูง และเป็นประโยชน์ในการพยากรณ์มากที่สุด ซึ่งในสมการนั้นควรประกอบด้วยตัวแปรอิสระจำนวนมากที่สุดเท่าที่จะทำได้ เพื่อค่าตัวแปรตามที่คำนวณได้มีความเชื่อถือได้สูงสุด

๒. เนื่องจากปัญหาของการวิจัยทั้งหลายมักจะเนื่องมาจากองค์ประกอบของเวลาและค่าใช้จ่าย การวิเคราะห์ข้อมูลโดยใช้ตัวแปรอิสระจำนวนมาก ๆ เป็นการสิ้นเปลืองและแทบเป็นไปไม่ได้ ดังนั้น จึงควรที่จะใช้ตัวแปรอิสระให้น้อยที่สุดเท่าที่จะทำได้

เนื่องจากข้อพิจารณาทั้งสองประการดังกล่าวมีส่วนขัดแย้งกันในตัววิธีการที่จะประนีประนอมสำหรับเป็นข้อสรุปในการพิจารณาเลือกสมการพยากรณ์ก็คือ พยายามเลือกสมการลดน้อยเพื่อการพยากรณ์ที่ดีที่สุด เพื่อได้ตัวแปรอิสระที่เหมาะสมจำนวนน้อย ๆ แต่ให้ผลการพยากรณ์ที่เชื่อถือได้สูงเท่าที่จะเป็นไปได้ซึ่งการเลือกตัวแปรในสมการลดน้อยเพื่อการพยากรณ์ดังกล่าว สามารถกระทำได้หลายวิธี และในแต่ละวิธีความเห็น หรือการตัดสินใจของบุคคลที่เกี่ยวข้องก็มีส่วนสำคัญไม่น้อย

ในบรรดาวิธีการเลือกสมการลดน้อยเพื่อการพยากรณ์ต่าง ๆ นั้น วิธีการลดน้อยพหุคูณแบบเพิ่มหรือลดตัวแปรเป็นขั้น ๆ (Stepwise Multiple Regression Procedure) เป็นวิธีการที่ให้ความถูกต้องในการพยากรณ์สูงสุด และเป็นวิธีที่คัดเลือกกลุ่มของตัวแปรอิสระที่เหมาะสม จำนวนพอ (optimum set of independent variables) ในการพยากรณ์ตัวแปรตามหรือตัวแปรที่เป็นผล

¹ N.R. Draper and H. Smith, Applied Regression Analysis (Wiley and Sons, 1966), p. 163.

ของการศึกษา¹ วิธีการนี้ได้รับความนิยมและนำมาใช้อย่างแพร่หลาย ซึ่งในการวิเคราะห์และเลือกตัวแปรอิสระสำหรับการวิจัยครั้งนี้ก็ใช้วิธีการดังกล่าวด้วย

เพื่อที่จะแสดงการใช้การถดถอยพหุคูณแบบเพิ่มหรือลดตัวแปรเป็นขั้น ๆ ในการพิจารณาเลือกตัวแปร สมมุติว่ามีตัวแปรอิสระทั้งหมด 20 ตัวแปร และต้องการสร้างสมการพยากรณ์สมการหนึ่งที่ประกอบด้วยตัวแปรอิสระจำนวน 5 ตัวแปรที่สามารถให้การพยากรณ์ที่ดี วิธีการเลือกสมการที่ประกอบด้วยตัวแปร 5 ตัวในสมการพยากรณ์ที่ดีวิธีหนึ่งก็คือ การสร้างสมการพยากรณ์ที่มีตัวแปรทุกตัวอยู่ในสมการ โดยการจัดตัวแปรนั้น ๆ ให้มีความประสานสัมพันธ์ (combination) กันในลักษณะต่าง ๆ ซึ่งถ้าพิจารณาการจัดในลักษณะดังกล่าวแล้วอาจทำได้ถึง 15,000 วิธี และเป็นการแน่นอนว่าวิธีการเลือกตัวแปรจากลักษณะการจัดทั้งหมดโดยวิธีดังกล่าวย่อมทำได้ยากและสิ้นเปลืองวิธีการที่ง่ายและสะดวกกว่าจึงมีความจำเป็นอย่างยิ่ง การวิเคราะห์ถดถอยพหุคูณแบบเพิ่มหรือลดตัวแปรเป็นขั้น ๆ เป็นวิธีการที่มีประสิทธิภาพและทำได้รวดเร็ว วิธีนี้จะเลือกตัวแปรอิสระและจัดตัวแปรดังกล่าวในสมการพยากรณ์นี้ว่าเป็นคำตอบที่ค่อนข้างเหมาะสม

วิธีการเริ่มด้วยในตอนแรกสมการพยากรณ์หรือสมการถดถอย (regression equation) จะประกอบด้วยตัวแปรอิสระเพียงตัวเดียวที่มีค่าความสัมพันธ์สูงสุดกับตัวแปรตาม ถ้าเราให้ X_1 เป็นตัวแปรดังกล่าวที่มีค่าสัมประสิทธิ์สหสัมพันธ์สูงสุดกับ Y สมการแรกในกระบวนการวิเคราะห์ถดถอยพหุคูณแบบเพิ่มหรือลดตัวแปรเป็นขั้น ๆ จะเป็น

$$Y_i = b_{01} + b_{11} X_{1i} + e_{1i} \quad (i=1,2,\dots, N)$$

หรือ

$$\hat{Y}_i = b_{01} + b_{11} X_{1i}$$

¹ R.S. Halinski and L.S. Feldt, " The Selection of Variables in Multiple Regression Analysis," American Educational Research Journal; vol.8, No.3, 1970, pp.170-171.

- เมื่อ Y_i เป็นข้อมูลที่ i ของตัวแปรตาม Y
- \hat{Y}_i เป็นค่าพยากรณ์ที่คำนวณได้จากค่าตัวแปร X_1
- N เป็นจำนวนข้อมูลหรือจำนวนกลุ่มตัวอย่างทั้งหมด
- X_1 เป็นตัวแปรอิสระใช้ในการพยากรณ์
เป็นตัวแปรแรกซึ่งมาจากกลุ่มของตัวแปรอิสระทั้งหมด
(potential predictors) จำนวน p ตัวแปร
(X_1, X_2, \dots, X_p)
- X_{1i} เป็นข้อมูลที่ i ของตัวแปรอิสระตัวที่ 1
- b_{01} เป็นค่าค่าคงที่ค่าต่ำสุดของจุดตัดแกน Y
(intercept)
- b_{11} เป็นค่าของความลาด (slope) สำหรับสมการถดถอย
- e_{1i} เป็นส่วนของความแปรปรวนใน Y ที่ไม่สามารถรู้ได้จาก
ความสัมพันธ์ของ Y และ X_1

หลักค่าคงที่ค่าต่ำสุดในการวิเคราะห์ถดถอยกำหนดว่า ค่าค่าคงที่
ของความลาดและจุดตัดแกน Y ที่ได้เป็นค่าที่ใหญ่รวมของส่วนที่เหลือค่าคงที่ค่าต่ำสุด
($\sum_{i=1}^N e_{1i}^2$) มีค่าน้อยกว่าค่าที่ได้จากการใช้หลักการหรือวิธีการอื่น

ความสัมพันธ์ของความแปรปรวนในตัวแปรตาม ความแปรปรวนในส่วน
ที่รู้ และความแปรปรวนในส่วนที่เหลือหรือส่วนที่ไม่รู้ อาจเขียนได้ในรูปของ

$$\sum_{i=1}^N (Y_i - \bar{Y})^2 = \sum_{i=1}^N (\hat{Y}_i - \bar{Y})^2 + \sum_{i=1}^N e_{1i}^2$$

หรือ

$$\begin{pmatrix} \text{ความแปรปรวน} \\ \text{ในตัวแปรตาม} \end{pmatrix} = \begin{pmatrix} \text{ความแปรปรวนในตัวแปรตาม} \\ \text{ที่วัดได้จากผลของสมการ} - \\ \text{ถดถอย} \end{pmatrix} + \begin{pmatrix} \text{ความแปรปรวนที่} \\ \text{ที่ไม่สามารถรู้ได้} \\ \text{จากสมการถดถอย} \end{pmatrix}$$

หรือ อาจเขียนเป็นสัญลักษณ์ได้คือ

$$Q_y = Q_f + Q_r$$

เมื่อ Q_y = ความแปรปรวนในตัวแปรตาม -

Q_f = ความแปรปรวนในตัวแปรตามที่วัดได้จากสมการ
ถดถอย

Q_r = ความแปรปรวนที่ไม่สามารถรู้ได้จากสมการ
ถดถอย

เมื่อพิจารณาคุณสมบัติของความสัมพันธ์ ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่าง Y และ \hat{Y} อาจเรียกว่า สัมประสิทธิ์สหสัมพันธ์พหุคูณ (coefficient of Multiple correlation) ซึ่งเป็นค่าที่แสดงขนาดของความสัมพันธ์ระหว่าง ตัวแปรตามและค่าพยากรณ์ที่ได้จากความสัมพันธ์เชิงเส้นตรงของตัวแปรอิสระ แต่เราสนใจค่าที่แสดงความสัมพันธ์ระหว่าง Y และ \hat{Y} ที่เรียกว่าสัมประสิทธิ์การกำหนด (coefficient of determination = R^2) ซึ่งอาจคำนวณได้โดยการยกกำลังสองของค่าสัมประสิทธิ์สหสัมพันธ์พหุคูณ

อย่างไรก็ตาม ฟิชเชอร์¹ (Fisher) ระบุว่าลักษณะการกระจาย

¹ R.A.Fisher, "The General Sampling Distribution of the Multiple Correlation Coefficient," Proceedings of the Royal Society of London, Series A, vol. 121, 1928, pp.654-673.

ของ R^2 แตกต่างไปจากการกระจายของสัมประสิทธิ์สหสัมพันธ์พหุคูณกำลังสอง เว้นเสียแต่ว่าจะไม่มีความสัมพันธ์เชิงเส้นตรงระหว่างตัวแปรในสมการ

อนึ่ง สัมประสิทธิ์การถ่วงน้ำหนัก R^2 สามารถพิจารณาได้ในรูปของอัตราส่วนของความแปรปรวนตามที่สามารถวัดหรือรู้ได้จากความแปรปรวนของตัวแปรอิสระ และ $1 - R^2$ เป็นค่าของอัตราส่วนในความแปรปรวนของตัวแปรตามที่ไม่สามารถวัดได้ ซึ่ง เคนดอลล์ และ สจวร์ต¹ (Kendall and Stuart) ได้ให้ความหมายสัมประสิทธิ์การถ่วงน้ำหนักในรูปของอัตราส่วนของความแปรปรวนที่เราสามารถวัดได้ดังนี้

$$\begin{aligned} R^2 &= \frac{Q_f}{Q_y} \\ &= \frac{Q_y - Q_r}{Q_y} \\ &= 1 - \frac{Q_r}{Q_y} \end{aligned}$$

จากสมการแสดงความสัมพันธ์ดังกล่าวจะเห็นว่า Q_f และ R^2 จะมีค่าสูงสุดเมื่อ Q_r มีค่าน้อยที่สุด และ Q_y เป็นค่าคงที่เฉพาะสำหรับกลุ่มตัวอย่างแต่ละกลุ่ม

ขั้นตอนที่สองของกระบวนการคัดเลือกตัวแปรโดยวิธีการลดถอยสลับไปสลับจะเกี่ยวข้องกับวิธีการคัดเลือกตัวแปรตัวที่สองเข้าในสมการลดถอย โดยที่ตัวแปรดังกล่าวจะเลือกจากตัวแปรอิสระที่เหลือจำนวน $p - 1$ ตัวแปร ซึ่งยังไม่ได้รับการพิจารณาถึงเลือกเข้าสู่สมการลดถอยในขั้นตอนแรก เมื่อทำการคัดเลือกจะพิจารณาตัวแปรอิสระที่เพิ่มความรู้ในการพยากรณ์ตัวแปรตามมากที่สุดกล่าวอีกนัยหนึ่งก็คือ อากการเพิ่มของ X_2 ยังผลทำให้สมการลดถอยซึ่งประกอบด้วยตัวแปร 2 ตัวแปร ให้ค่าการเพิ่มของ R^2 สูงกว่าการเพิ่มตัวแปรอิสระอื่น ๆ ในสมการแล้ว X_2 จะได้รับการคัดเลือกให้เป็นตัวแปรที่เข้าร่วมการพยากรณ์ในสมการ

¹H.G.Kendall and A. Stuart, The Advance Theory of Statistics, New York : Hafner, vol.2, 1961, p.148.

$$Y_i = b_{02} + b_{12}X_{1i} + b_{22}X_{2i} + e_{2i}$$

ถ้าค่า R^2 จากขั้นตอนแรก และขั้นตอนที่สองเขียนแทนโดย R_1^2 และ R_2^2 ตามลำดับแล้ว $R_2^2 - R_1^2$ จะมีความมากกว่าค่าอื่น ๆ ที่ได้จากการเลือกตัวแปรอิสระที่ไม่ใช่ X_2 ในขั้นการเลือกดังกล่าว

อีกวิธีการหนึ่งในการหาว่าตัวแปรอิสระใดที่จะให้การเพิ่มค่า R^2 สูงกว่าในขั้นตอนที่สองนี้ อาจทำได้โดยการหาค่ากำลังสองของสหสัมพันธ์ส่วนย่อย (squared partial correlation coefficient) ระหว่างค่า Y และ ค่าแต่ละค่าของตัวแปรอิสระที่เหลือจำนวน $p - 1$

ตัวแปร โดยไม่พิจารณา X_1 (partialled out X_1) ถ้า X_2 ให้ค่าการเพิ่ม R^2 สูงสุดก็จะทำให้ค่ากำลังสองของสหสัมพันธ์ส่วนย่อยสูงด้วย ซึ่งค่ากำลังสองของสหสัมพันธ์ส่วนย่อยสามารถที่จะเขียนได้ในรูปอัตราส่วนของความแปรปรวนที่ไม่รู้ (unaccounted for variation) ในขั้นตอนแรกซึ่งสามารถจะวัดได้ (accounted for) โดยการเพิ่มตัวแปรอิสระตัวที่สองเข้าในสมการถดถอย ค่าสัมประสิทธิ์สหสัมพันธ์ส่วนย่อยนี้อาจเขียนโดยใช้สัญลักษณ์ $r_{yx_2 \cdot x_1}$ หรืออาจพิจารณาแปลความหมายในรูปของสหสัมพันธ์ระหว่างปริมาณส่วนตัวที่เหลือจากการพยากรณ์ (residuals) ของสมการพยากรณ์ 2 สมการตัวอย่างเช่น $r_{yx_2 \cdot x_1}$ อาจพิจารณาในรูปของ $R_{f_a f_b}$ เมื่อ f_a และ f_b เป็นส่วนที่เหลือจากการพยากรณ์ของสมการ

$$Y_{1i} = b_{0a} + b_{1a}X_{1i} + f_{ai}$$

และ

$$X_{2i} = b_{0b} + b_{1b}X_{1i} + f_{bi}$$

หลังจากการคัดเลือกตัวแปรอิสระเข้าในสมการพยากรณ์แต่ละครั้ง จะมีการทดสอบนัย-
สำคัญของตัวแปรอิสระที่รวมกันอยู่ในสมการถดถอย โดยการทดสอบสมมุติฐานว่าความ
แปรปรวนใน y จะสามารถรู้หรืออธิบายได้จากความแปรปรวนของตัวแปรอิสระที่อยู่ใน
สมการหรือไม่ หรืออีกนัยหนึ่งเป็นการทดสอบสมมุติฐานว่า อัตราส่วนของความ
แปรปรวนที่ได้จากการทำนาย (R^2) มีค่าแตกต่างจากศูนย์หรือไม่ ซึ่งการทดสอบ
สมมุติฐานนี้บางครั้งเรียกว่า การทดสอบสถิติส่วนรวม F (overall F -test)
โดยใช้สูตร -

$$F_{f, \nu_r} \sim \frac{R^2 Q_y / \nu_f}{(1-R^2) Q_y \nu_r} = \frac{\nu_r R^2}{\nu_f (1-R^2)}$$

เมื่อ ν_f เป็นจำนวนตัวแปรอิสระในสมการถดถอย (p)
 ν_r มีค่าเป็น $N - p - 1$

ขั้นตอนการทดสอบอีกประการหนึ่งเพื่อการพิจารณาว่าตัวแปรใดควร
พิจารณาเข้าหรือควรจะตัดออกจากสมการถดถอย อาจกระทำได้โดยการทดสอบสถิติ
ส่วนย่อย F (partial F -test) เนื่องจากอิทธิพลของตัวแปรอิสระหนึ่ง ๆ ที่มี
ต่อตัวแปรตามอาจจะมีมาก เมื่อตัวแปรนั้นเป็นตัวแปรพยากรณ์ (predictor)
เพียงตัวแปรเดียวในสมการถดถอย แต่ตัวแปรที่ได้รับการพิจารณาแล้วว่าเป็น-

ตัวแปร เกี่ยวที่ที่สุดได้รับการคัดเลือกเข้าสู่สมการถดถอยตั้งแต่ต้น อาจจะไม่มียุทธศาสตร์
ในการที่จะให้ข้อมูลเพิ่มเติมเมื่อพิจารณาประกอบในด้านความสัมพันธ์กับตัวแปรอิสระที่ได้
รับเลือกเข้าไปในสมการถดถอยใหม่ กล่าวคือ ตัวแปรอิสระเดียวกันนั้นอาจจะมีอิทธิพลต่อ
ตัวแปรตามน้อยเมื่อตัวแปรอิสระนั้นได้รับการพิจารณาเข้าสู่สมการตามหลังตัวแปรอิสระอื่น
ทั้งนี้ เนื่องจากความจริงเกี่ยวกับความสัมพันธ์ของตัวแปรอิสระที่ศึกษากับตัวแปรอิสระที่อยู่ใน
ในสมการพยากรณ์

การทดสอบสถิติส่วนย่อย F สามารถกระทำกับสัมประสิทธิ์การถดถอยทุกค่า
ไม่ว่าตัวแปรอิสระนั้นจะได้รับการเลือกในขั้นตอนใดของกระบวนการเลือกตัวแปร เข้าสู่สมการ
ถดถอย ทั้งนี้เพื่อจะศึกษาผลเกี่ยวของหรืออิทธิพลรวมของตัวแปรอิสระในการพยากรณ์เกี่ยว
เนื่องกับการศึกษาสถิติส่วนย่อย F การทดสอบนัยสำคัญของการเพิ่มความแปรปรวน อันเป็น
ผลของการเพิ่มตัวแปรอิสระตัวที่ n ในการสมการพยากรณ์ที่มีตัวแปรอิสระอยู่แล้ว m ตัวอาจ
ทำได้โดยการคำนวณอัตราส่วน

$$F_{1, N-n-1} \sim \frac{(R_n^2 - R_m^2)(N-n-1)}{(1-R_n^2)} = \frac{(R_n^2 - R_m^2)}{(1-R_n^2)/(N-n-1)}$$

เมื่อ R_n^2 คือสัมประสิทธิ์การพยากรณ์สำหรับสมการที่มีตัวแปรอิสระ n ตัว
 R_m^2 คือสัมประสิทธิ์การพยากรณ์สำหรับสมการที่มีตัวแปรอิสระ m ตัว

ผลต่างของ R_n^2 และ R_m^2 คืออัตราส่วนของความแปรปรวนที่เพิ่มขึ้นโดย
การนำตัวแปรที่ n เข้าสู่สมการพยากรณ์ อันเป็นการตรวจสอบโดยใช้สถิติ F ดังกล่าวจะกระทำ
กับตัวแปรทุกตัวที่พิจารณาคัดเลือกเข้าสู่สมการพยากรณ์ในทุก ๆ ขั้นตอน ตัวแปรใดที่ให้ผลการ
ทดสอบต่ำกว่าเกณฑ์ที่ตั้งไว้จะถูกกำจัดออกจากสมการพยากรณ์

กระบวนการคัดเลือกตัวแปร เข้าในสมการพยากรณ์และการทดสอบตัวแปร
นั้น ๆ จะสิ้นสุดลงเมื่อไม่มีตัวแปรใดที่ได้รับการพิจารณาเข้าสู่สมการพยากรณ์ และไม่มีตัว
แปรใดที่ถูกตัดออกจากสมการพยากรณ์

วิธีการคำนวณสำหรับการ ถดถอยพหุคูณแบบเพิ่มหรือลดตัวแปรเป็นชั้น ๆ ส่วนใหญ่ใช้หลักการแก้ระบบสมการเส้นตรง (system of linear equations) นั่นคือใช้วิธีเปลี่ยนแถว (row) และสัณฐาน (column) ของข้อมูลตัวแปรต่าง ๆ ตามวิธีของกอสส์ (Gauss Elimination) คอยวิธีการดังกล่าวจะให้ข้อมูลที่จำเป็นในการเลือกตัวแปรอิสระตัวต่อไปเพื่อที่จะนำเข้าสู่สมการพยากรณ์

ข้อสังเกตเกี่ยวกับการวิเคราะห์ถดถอยพหุคูณแบบเพิ่มหรือลดตัวแปรเป็นชั้น ๆ

ดังได้กล่าวแล้วว่า การวิเคราะห์ถดถอยพหุคูณแบบเพิ่มหรือลดตัวแปรเป็นชั้น ๆ เป็นวิธีการเลือกสรรตัวแปรอิสระหรือกลุ่มของตัวแปรอิสระสำหรับการพยากรณ์ตัวแปรตาม ซึ่งเป็นวิธีที่นิยมใช้กันอย่างแพร่หลายกว่าวิธีการอื่น วิธีการนี้ประหยัดทั้งกำลังแรงงาน เวลา ในการคำนวณและเลือกตัวแปร เพราะให้ผลการเลือกในรูปของสมการถดถอยแบบเส้นตรงเพียงสมการเดียว เมื่อเปรียบเทียบกับวิธีที่จะต้องศึกษาตัวแปรทั้งหมด และสร้างเป็นสมการถดถอยตามลักษณะการจับของกลุ่มตัวแปร (all possible regressions)

วิธีการวิเคราะห์ถดถอยพหุคูณแบบเพิ่มหรือลดตัวแปรเป็นชั้น ๆ นี้ มีส่วนคิดในการเปิดโอกาสสำหรับการทดสอบ และเลือกตัดตัวแปรอิสระที่มีความสัมพันธ์สูงกับตัวแปรที่เลือกรับเลือกก่อนในสมการถดถอย เพื่อให้ผลเพิ่มการพยากรณ์ การเลือกระดับนัยสำคัญสำหรับเลือกตัวแปรเข้าในสมการถดถอยหรือตัดตัวแปรออกจากสมการถดถอยนั้นมีขีดจำกัด F ไม่จำเป็นจะต้องกำหนดให้เท่าเทียมกันก็ได้ ข้อจำกัดของวิธีการวิเคราะห์ถดถอยพหุคูณแบบเพิ่มหรือลดตัวแปรเป็นชั้น ๆ ในการเลือกกลุ่มตัวแปรอิสระก็คือ วิธีการนี้สมมุติว่ากลุ่มตัวแปรอิสระที่ใช้สำหรับการอธิบายความแปรปรวนในตัวแปรตามนั้นจะมีเพียงกลุ่มเดียวที่ดีที่สุด และวิธีวิเคราะห์ดังกล่าวจะช่วยชี้ถึงกลุ่มตัวแปรอิสระนั้น แต่ถ้าตัวแปรอิสระมีความสัมพันธ์กันสูง วิธีการวิเคราะห์จะให้ผลที่เชื่อถือได้ยาก เพราะกลุ่มตัวแปรอิสระที่พิจารณาว่าดีที่สุดแล้วนั้นอาจจะไม่ใช่กลุ่มที่ดีที่สุดตามหลักเหตุผลที่ควรจะเป็น ดังนั้น ในการกำหนดว่าตัวแปรใดสมควรจะได้รับการพิจารณาคัดเลือกเพื่อการพยากรณ์ ควรจะพิจารณาโดยหลักเหตุผล (rational) คือพิจารณาในแง่ของการปฏิบัติและประหยัด ประการหนึ่ง กับทั้งจะต้องพิจารณาคด้วยหลักสถิติ (statistical) ประกอบกันไปอีกประการหนึ่งว่าตัวแปรไหนมีความสำคัญมากกว่ากันเพียงใด