

บทที่ 2

วรรณคดีที่เกี่ยวข้องกับการวิจัย

ในการวิจัยเรื่องการเปรียบเทียบประสิทธิภาพระหว่างการวิเคราะห์ถดถอยโลจิสติกพหุภาคกับการวิเคราะห์จำแนกในการศึกษาปัจจัยที่มีผลต่อระยะเวลาในการศึกษาและผล การประเมินวิทยานิพนธ์ ผู้วิจัยได้นำเสนอผลการศึกษาค้นคว้าและงานวิจัยที่เกี่ยวข้องแบ่งเป็น 4 ตอน คือ ตอนที่หนึ่ง การศึกษาระดับปริญญาโท สาขาสังคมศาสตร์ ตอนที่สอง ปัจจัยที่เกี่ยวข้องกับความสำเร็จในการทำวิทยานิพนธ์ ตอนที่สาม การวิเคราะห์ถดถอยโลจิสติกและการวิเคราะห์จำแนก และตอนที่สี่ งานวิจัยที่เปรียบเทียบผลการวิเคราะห์ข้อมูล ระหว่างการวิเคราะห์ถดถอยโลจิสติกกับการวิเคราะห์จำแนก

ตอนที่หนึ่ง การศึกษาระดับปริญญาโท สาขาสังคมศาสตร์

เนื้อหาในตอนนี้ ผู้วิจัยนำเสนอเป็นการศึกษาระดับปริญญาโท สาขา สังคมศาสตร์ ในความดูแลของบัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย โดยแยกเสนอเป็น 2 ส่วน คือ หลักสูตรปริญญาโท สาขาสังคมศาสตร์ และรูปแบบการทำวิทยานิพนธ์

1. หลักสูตรปริญญาโท สาขาสังคมศาสตร์

การศึกษาในระดับมหาบัณฑิต สาขาสังคมศาสตร์ อยู่ในความดูแลของบัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัยมีหลักสูตรปริญญาโท สาขาสังคมศาสตร์ตามระเบียบจุฬาลงกรณ์มหาวิทยาลัย ว่าด้วย การศึกษาในระดับบัณฑิตศึกษามหาบัณฑิตวิทยาลัย พ.ศ. 2534 (จุฬาลงกรณ์ มหาวิทยาลัย, บัณฑิตวิทยาลัย, 2539) ดังรายละเอียดของหลักสูตรรวม 4 ประเด็น ต่อไปนี้

1. โครงสร้างของหลักสูตร หลักสูตรนี้ประกอบด้วยจำนวนหน่วยกิตรวมตลอดหลักสูตร ไม่น้อยกว่า 36 หน่วยกิต และมี 2 แผน คือ

แผน ก เป็นแผนการศึกษาที่เน้นการวิจัยโดยมีการทำวิทยานิพนธ์ ซึ่งมีได้ 2 แบบ คือ

1) ประกอบด้วยหน่วยกิตวิทยานิพนธ์ไม่น้อยกว่า 36 หน่วยกิต และหลักสูตรอาจ กำหนดให้การเรียนรายวิชา หรือทำกิจกรรมวิชาการอื่นประกอบได้โดยไม่นับหน่วยกิต

2) ประกอบด้วยหน่วยกิตวิทยานิพนธ์ไม่น้อยกว่า 12 หน่วยกิต และหน่วยกิตรายวิชาไม่น้อยกว่า 12 หน่วยกิต

แผน ข เป็นแผนการศึกษาที่เน้นการศึกษารายวิชา โดยไม่ต้องทำวิทยานิพนธ์ แต่ต้องมีหน่วยกิตรายวิชาที่นิสิตทำการศึกษาค้นคว้าอิสระไม่น้อยกว่า 3 หน่วยกิต และไม่เกิน 6 หน่วยกิต สาขาวิชาใดเปิดสอนหลักสูตรแผน ก ไม่จำเป็นต้องเปิดสอนหลักสูตรแผน ข ด้วย แต่ถ้าเปิดสอนหลักสูตรแผน ข จะต้องมีหลักสูตรแผน ก ให้นิสิตเลือกศึกษาได้ด้วย

2.ระยะเวลาการศึกษาและการสำเร็จการศึกษา ตามระเบียบได้กำหนดระยะเวลาศึกษาในหลักสูตรปริญญาวิทยาศาสตรบัณฑิตได้ไม่เกิน 4 ปีการศึกษา นับจากภาคการศึกษาแรกที่รับเข้าศึกษาในหลักสูตร และนิสิตปริญญาวิทยาศาสตรบัณฑิตที่ศึกษาในหลักสูตรแผน ก จะสำเร็จการศึกษาได้ต้องสอบผ่านรายวิชาและวิทยานิพนธ์และอื่น ๆ ตามที่หลักสูตรกำหนดไว้ ส่วนนิสิตปริญญาวิทยาศาสตรบัณฑิตที่ศึกษาในหลักสูตรแผน ข จะสำเร็จการศึกษาได้ต้องสอบผ่านรายวิชาและสอบประมวลความรู้ (comprehensive examination) และอื่น ๆ ตามที่หลักสูตรกำหนดไว้

3.การลงทะเบียน การเสนอโครงร่าง อาจารย์ที่ปรึกษา คณะกรรมการสอบและการสอบวิทยานิพนธ์ในระดับปริญญาวิทยาศาสตรบัณฑิต ระเบียบได้กำหนดขั้นตอนการดำเนินการดังนี้

3.1 นิสิตจะลงทะเบียนวิทยานิพนธ์ตั้งแต่ภาคการศึกษาแรกที่เข้าศึกษาก็ได้

3.2 นิสิตจะต้องเสนอโครงร่างวิทยานิพนธ์ต่อบัณฑิตวิทยาลัยตามหลักเกณฑ์ และระยะเวลาที่บัณฑิตวิทยาลัยประกาศกำหนด ทั้งนี้คณะกรรมการบัณฑิตศึกษาประจำคณะหรือสหสาขาวิชาอาจกำหนดหลักเกณฑ์และระยะเวลาเพิ่มเติมได้เท่าที่ไม่ขัดหรือแย้งการกำหนดของบัณฑิตวิทยาลัย

3.3 เมื่อโครงร่างวิทยานิพนธ์ได้รับอนุมัติจากบัณฑิตวิทยาลัยแล้วให้หัวหน้าภาควิชาเสนอชื่ออาจารย์ที่ปรึกษาวิทยานิพนธ์ต่อคณบดีบัณฑิตวิทยาลัยโดยเร็วที่สุดเพื่อพิจารณาแต่งตั้งอาจารย์ที่ปรึกษาวิทยานิพนธ์จะต้องเป็นอาจารย์บัณฑิตศึกษาในภาควิชา หรืออาจมีบุคคลอื่นที่ไม่ได้เป็นอาจารย์บัณฑิตศึกษาในภาควิชา นั้นเป็นอาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วมได้ 1 คน

ในกรณีที่เป็นหลักสูตรสหสาขาวิชาหรือหลักสูตรที่มีภาควิชารับผิดชอบร่วมกันมากกว่า 1 ภาควิชาขึ้นไป ให้ประธานคณะกรรมการบัณฑิตศึกษาสหสาขาวิชา หรือประธานคณะกรรมการบัณฑิตศึกษาประจำคณะเป็นผู้เสนอชื่ออาจารย์ที่ปรึกษาวิทยานิพนธ์แล้วแต่กรณี

3.4 หัวหน้าภาควิชา หรือประธานคณะกรรมการบัณฑิตศึกษาสหสาขาวิชา หรือประธานคณะกรรมการบัณฑิตศึกษาประจำคณะ แล้วแต่กรณี เสนอชื่อกรรมการสอบวิทยานิพนธ์ให้

คณบดีบัณฑิตวิทยาลัยพิจารณาแต่งตั้ง ทั้งนี้จะต้องแต่งตั้งกรรมการสอบวิทยานิพนธ์อย่างซ้ำ
ในภาคการศึกษาที่ผลิตลงทะเบียนวิทยานิพนธ์ครบตามหลักสูตร

คณะกรรมการสอบวิทยานิพนธ์จะต้องมีจำนวนอย่างน้อย 3 คน แต่ไม่เกิน 5 คน ซึ่ง
ประกอบด้วยหัวหน้าภาควิชา หรือผู้ที่หัวหน้าภาควิชามอบหมายทำหน้าที่เป็นประธาน
อาจารย์ที่ปรึกษาวิทยานิพนธ์และผู้ทรงคุณวุฒิอย่างน้อย 1 คน เป็นกรรมการ ในกรณีหลักสูตร
แผน ก แบบที่ 1 ผู้ทรงคุณวุฒิจะต้องเป็นบุคคลภายนอกมหาวิทยาลัย

ในกรณีที่หัวหน้าภาควิชาเป็นอาจารย์ที่ปรึกษา หรืออาจารย์ที่ปรึกษาร่วม ให้หัวหน้า
ภาควิชามอบหมายให้ผู้อื่นทำหน้าที่เป็นประธาน

3.5 เมื่อนิสิตทำวิทยานิพนธ์เสร็จสมบูรณ์แล้ว ให้นิสิตยื่นคำขอนัดสอบวิทยานิพนธ์ตาม
ความเห็นชอบของอาจารย์ที่ปรึกษาวิทยานิพนธ์ต่อบัณฑิตวิทยาลัยก่อนวันนัดสอบอย่างน้อย 2
สัปดาห์

3.6 ในวันสอบวิทยานิพนธ์ คณะกรรมการสอบวิทยานิพนธ์จะต้องประกอบด้วยประธาน
กรรมการ อาจารย์ที่ปรึกษาวิทยานิพนธ์ และกรรมการอื่น ๆ อย่างน้อย 1 คน มาดำเนินการ
สอบจึงจะถือว่าการสอบนั้นมีผลสมบูรณ์ ถ้ากรรมการไม่ครบตามจำนวนดังกล่าวข้างต้นให้
เลื่อนการสอบออกไป ในกรณีที่จำเป็น ผู้มีอำนาจเสนอชื่อคณะกรรมการสอบวิทยานิพนธ์อาจ
เสนอให้คณบดีบัณฑิตวิทยาลัยแต่งตั้งซ่อมกรรมการสอบวิทยานิพนธ์ได้ ทั้งนี้จะต้องกำหนดวัน
สอบวิทยานิพนธ์ครั้งใหม่ให้พอสมควรแก่การที่กรรมการที่แต่งตั้งซ่อมใหม่มีเวลาตรวจอ่าน
วิทยานิพนธ์

การสอบวิทยานิพนธ์ต้องประกอบด้วย การตรวจอ่านวิทยานิพนธ์ การประชุมปรึกษา
ระหว่างกรรมการในวันสอบวิทยานิพนธ์ การทดสอบความรู้ นิสิตด้วยการซักถามหรือด้วยวิธี
ใด ๆ จึงจะถือว่าการสอบนั้นมีผลสมบูรณ์ โดยให้ประธานคณะกรรมการสอบวิทยานิพนธ์
รายงานผลการสอบไปยังบัณฑิตวิทยาลัยภายใน 2 สัปดาห์นับจากวันสอบ

4. การประเมินผลวิทยานิพนธ์การประเมินผลวิทยานิพนธ์ในหลักสูตรปริญญา
มหาบัณฑิตกำหนดไว้ 4 ระดับ คือ ดีมาก (very good) ดี (good) ผ่าน (pass) และตก (fail)

2. รูปแบบการทำวิทยานิพนธ์

บัณฑิตวิทยาลัยกำหนดรูปแบบในการจัดทำวิทยานิพนธ์ของนิสิตบัณฑิตศึกษาเพื่อให้เป็นไปในรูปแบบเดียวกัน โดยแบ่งเป็น 5 ส่วน คือ ส่วนนำ ส่วนเนื้อหา รายการอ้างอิง ภาคผนวก และประวัติผู้เขียน ซึ่งแต่ละส่วนประกอบด้วยรายละเอียดต่าง ๆ ดังนี้ (จุฬาลงกรณ์มหาวิทยาลัย, บัณฑิตวิทยาลัย, 2539)

ส่วนนำ ประกอบด้วย

1.หน้าปก เป็นหน้าที่ระบุชื่อวิทยานิพนธ์ ชื่อและสกุลของผู้เขียนพร้อมคำนำหน้านาม เช่น นาย ,นาง ,นางสาว แต่ไม่ต้องเขียนคุณวุฒิไว้ท้ายชื่อ ยกเว้นถ้ามียศ ฐานันดรศักดิ์ หรือ สมณศักดิ์ให้ใส่ไว้ด้วย หน้าปกให้จัดพิมพ์ 2 หน้า ถ้าเขียนวิทยานิพนธ์เป็นภาษาไทย หน้าแรกให้พิมพ์ชื่อความเป็นภาษาไทย หน้าที่สองเป็นภาษาอังกฤษ แต่ถ้าเขียนวิทยานิพนธ์เป็นภาษาอื่นที่ไม่ใช่ภาษาไทย หน้าแรกให้ใช้ภาษาไทย แต่หน้าที่สองให้ใช้ภาษาที่เขียนวิทยานิพนธ์

2.หน้าอนุมัติ ถ้าเขียนวิทยานิพนธ์เป็นภาษาไทย หน้าอนุมัติให้เขียนเป็นภาษาไทย แต่ถ้าเขียนวิทยานิพนธ์เป็นภาษาอื่น หน้าอนุมัติก็ให้เขียนด้วยภาษาที่ใช้เขียนวิทยานิพนธ์

3.บทคัดย่อ (abstract) เป็นข้อความสรุปผลการวิจัยที่จะทำให้ผู้อ่านทราบถึงเนื้อหาวิทยานิพนธ์อย่างรวดเร็ว ในบทคัดย่อวิทยานิพนธ์ควรประกอบด้วยส่วนต่าง ๆ ดังนี้

- วัตถุประสงค์ จุดมุ่งหมาย และขอบเขตของการวิจัย
- วิธีการวิจัย รวมถึงเครื่องมือที่ใช้ วิธีการเก็บข้อมูล จำนวนและลักษณะของกลุ่มที่ศึกษา
- ผลการวิจัย รวมถึงระดับนัยสำคัญทางสถิติ (ถ้ามีการทดสอบ)

ในการเขียนบทคัดย่อต้องคำนึงถึงเรื่องต่อไปนี้

- ความถูกต้อง ต้องระบุวัตถุประสงค์และเนื้อหาของเรื่องตามที่ปรากฏในวิทยานิพนธ์
- ความสมบูรณ์ ในการเขียนคำย่อ หรือคำที่ไม่คุ้นเคย ให้เขียนด้วยคำเต็ม เมื่อกล่าวถึงครั้งแรก ไม่จำเป็นจะต้องอ้างเอกสาร ยกตัวอย่าง ยกข้อความ สมการ หรือภาพวาด คำที่ใช้ในบทคัดย่อเป็นคำสำคัญ เพื่อประโยชน์ในการทำดัชนีเพื่อการสืบค้น นิสิตสามารถใช้พจนานุกรมฉบับราชบัณฑิตยสถาน หรือ Dictionary ซึ่งเป็นที่ยอมรับกันทั่วไปเป็นแนวทางในการสะกดคำ

- ความเฉพาะ กระชับ ชัดเจน ในการเขียนบทคัดย่อต้องพยายามเขียนให้ได้ใจความและสั้นที่สุด

- ความน่าอ่านและราบรื่น ควรคำนึงถึงหลักการใช้ภาษาตามความเหมาะสม ในกรณีที่เขียนบทความย่อเป็นภาษาต่างประเทศให้คำนึงถึงการใช้กาล (tenses) ให้ถูกต้องตามมาตรฐานที่ใช้กันในแต่ละสาขาวิชา

4. กิตติกรรมประกาศ (acknowledgements) เป็นข้อความกล่าวขอบคุณผู้ให้ความช่วยเหลือและให้ความร่วมมือในการทำวิทยานิพนธ์ เพื่อแสดงถึงจรรยาบรรณทางวิชาการที่ผู้วิจัยควรถือปฏิบัติ

5. สารบัญ เป็นรายการที่แสดงถึงหัวข้อสำคัญทั้งหมดของวิทยานิพนธ์ โดยเรียงลำดับตามเลขหน้า และให้เขียนด้วยภาษาที่เขียนวิทยานิพนธ์

6. สารบัญอื่น ๆ (ถ้ามี) เช่น สารบัญตาราง สารบัญภาพ เป็นต้น ในส่วนนี้จะเป็นส่วนที่บอกชื่อตารางพร้อมทั้งเลขหน้าของตาราง ภาพ ทั้งหมดที่มีอยู่ในวิทยานิพนธ์

7. คำอธิบายสัญลักษณ์และคำย่อ (ถ้ามี) เป็นส่วนอธิบายถึงสัญลักษณ์และคำย่อต่าง ๆ ที่ใช้ในวิทยานิพนธ์

ส่วนเนื้อหา ในส่วนนี้นิสิตอาจนำเสนอได้ใน 2 แนวทางหลัก คือ

1. เสนอเป็นงานเรื่องเดียวทั้งเล่ม โดยแบ่งเนื้อหาเป็น 3 ตอน คือ

- บทนำ เป็นขอบทที่ 1 กล่าวถึงความสำคัญและความเป็นมาของปัญหา วัตถุประสงค์ ขอบเขตของการวิจัย วิธีดำเนินการวิจัยโดยย่อ และประโยชน์ที่จะได้จากการวิจัย

รายละเอียดอื่น ๆ ให้อยู่ในดุลพินิจของอาจารย์ที่ปรึกษาวิทยานิพนธ์และภาควิชา นั้น ๆ

- ตัวเรื่อง เนื้อหาในตัวเรื่องวิทยานิพนธ์จะแบ่งออกเป็นกี่บทก็ได้ตามความจำเป็น ในส่วนนี้ควรมีบทหนึ่งเกี่ยวกับปริทัศน์วรรณกรรม (literature review) ที่เกี่ยวข้องกับ รายงานทฤษฎี แนวคิด งานวิจัยที่เกี่ยวข้องกับเรื่องที่ศึกษา อภิปรายบรรยายวิธีการที่ใช้ในการวิจัยโดยละเอียดว่ามีขั้นตอนในการทำอย่างไร แต่ละขั้นตอนใช้เอกสารข้อมูลหรือเครื่องมือใด และเอกสารข้อมูลหรือเครื่องมือ นั้นได้มาอย่างไร โดยวิธีใด รวมถึงการวิเคราะห์ รายงานผล และอภิปรายผลการวิจัย ในบางสาขาวิชา บทหลังนี้อาจแยกเป็น 2 บท คือ บทหนึ่งรายงานวิธีการที่ใช้ และอีกบทหนึ่งเป็นการวิเคราะห์ รายงานผล และอภิปรายผลการวิจัย

ขั้นตอนโดยละเอียดของส่วนตัวเรื่อง ควรจะได้ดำเนินการตามแบบแผนการวิจัยที่เป็นที่ยอมรับในแต่ละสาขาวิชา

- ข้อสรุปและ / หรือข้อเสนอแนะ เป็นตอนสรุปเนื้อหาในวิทยานิพนธ์ทั้งหมด พร้อมทั้งชี้ถึงข้อจำกัดในการวิจัย (ถ้ามี) ข้อเสนอแนะเกี่ยวกับการวิจัยต่อไปและประโยชน์ที่ได้จากการประยุกต์ผลการวิจัยที่ได้

2.เสนอเป็นบทความวิจัยมากกว่าหนึ่งบทความ บทความวิจัยที่เสนอจะต้องครอบคลุมเนื้อหาเรื่องเดียวกันที่โยงใยกันเป็นวิทยานิพนธ์ 1 ฉบับ ได้ บทความวิจัยที่เสนออาจอยู่ในรูปแบบดังต่อไปนี้หนึ่งหรือมากกว่าหนึ่งรูปแบบ คือ บทความวิจัยที่ได้รับการตีพิมพ์แล้ว บทความวิจัยที่ได้รับการตอบรับที่จะตีพิมพ์แล้ว บทความที่นิสิตเสนอให้วารสารตีพิมพ์แล้ว บทความวิจัยที่ได้นำเสนอในการประชุมทางวิชาการแล้ว บทความวิจัยที่ได้รับการตอบรับให้นำเสนอในการประชุมทางวิชาการแล้ว หรือบทความวิจัยที่นิสิตได้เสนอที่จะเสนอในการประชุมทางวิชาการต่อผู้จัดประชุมแล้ว การนำเสนอเนื้อหาแบ่งออกเป็น 3 ตอน เช่นกัน คือ

-บทนำ ให้เขียนความเป็นมาและสำคัญของปัญหา วัตถุประสงค์ ขอบข่ายของบทความวิจัยแต่ละเรื่องที่จะนำเสนอในตัวเรื่อง นิสิตต้องเขียนให้เห็นถึงความเชื่อมโยงของบทความวิจัยที่จะนำเสนอในตัวเรื่องอย่างชัดเจน

-ตัวเรื่อง นิสิตสามารถนำเสนอในรูปแบบที่เผยแพร่บทความวิจัยนั้น ๆ ได้เลยโดยไม่ต้องเปลี่ยนแปลงในตัวเรื่องอย่างชัดเจน

-ข้อสรุป และหรือข้อเสนอแนะ ให้เขียนข้อสรุปและหรือข้อเสนอแนะโดยแสดงความเชื่อมโยงของบทความวิจัยที่นำเสนอในตัวเรื่อง

รายการอ้างอิง (references) ในวิทยานิพนธ์แต่ละเรื่องจะต้องมีรายการอ้างอิง ได้แก่ รายชื่อหนังสือ สิ่งพิมพ์ ใตศทัศน์วัสดุ แหล่งอ้างอิงอื่น ๆ ตลอดจนวิธีการที่ได้ข้อมูลมาเพื่อประกอบการเขียนวิทยานิพนธ์เรื่องนั้น ๆ รายการอ้างอิงจะอยู่ต่อจากส่วนเนื้อเรื่องและก่อนภาคผนวก การลงรายการอ้างอิงให้พิมพ์เฉพาะเอกสารที่มีการอ้างอิงถึงในเนื้อหาของวิทยานิพนธ์เท่านั้น ถ้ามีเอกสารอื่นไม่ได้อ้างอิง แต่มีความเกี่ยวข้องและผู้วิจัยต้องการนำมาพิมพ์ด้วย เช่น เอกสารอื่นที่เกี่ยวข้อง ให้ระบุต่อรายการอ้างอิงโดยขึ้นหน้าใหม่และใช้คำว่า บรรณานุกรม

(หมายเหตุ ในกรณีที่นำเสนอเป็นบทความวิจัยมากกว่าหนึ่งบทความนิสิตสามารถระบุรายการอ้างอิงไว้ท้ายบทแต่ละบทได้)

ภาคผนวก (appendix) เป็นข้อความที่ไม่สามารถบรรจุอยู่ในส่วนเนื้อหา แต่เป็นส่วนที่เสริมให้เกิดความเข้าใจชัดเจนขึ้น ภาคผนวกเป็นข้อมูลที่ใช้ในวิทยานิพนธ์แต่ไม่ได้อ้างอิงโดยตรงหรือไม่สมควรอ้างอิงในส่วนเนื้อหา เพราะมีความยาวมากหรือไม่เหมาะสมแก่การพิมพ์ไว้ในส่วนเนื้อหา หน้าแรกของภาคผนวกให้ขึ้นหน้าใหม่มี คำว่า ภาคผนวก อยู่กลางหน้ากระดาษ ในกรณีที่มียหลายภาคผนวกให้ใช้เป็นภาคผนวก ก ภาคผนวก ข ภาคผนวก ค ตามลำดับ โดยแต่ละภาคผนวกให้ขึ้นหน้าใหม่

ประวัติผู้เขียน (biography) ประวัติผู้เขียนให้เขียนเป็นความเรียง โดยไม่ต้องแยกเป็นข้อ ๆ ซึ่งมีรายละเอียดเกี่ยวกับ ชื่อ สกุล พร้อมคำนำหน้า ได้แก่ นาย, นาง, นางสาว ถ้ามียศ ฐานันดรศักดิ์ สมณศักดิ์ หรือตำแหน่งทางวิชาการ พร้อมทั้งวัน เดือน ปี และสถานที่เกิด วุฒิการศึกษาตั้งแต่ชั้นปริญญาบัณฑิตหรือเทียบเท่าขึ้นไป สถานศึกษาและปีการศึกษาที่สำเร็จการศึกษา รวมทั้งประสบการณ์ ผลงานทางวิชาการ รางวัลทางวิชาการ รางวัลหรือทุนการศึกษาเฉพาะที่สำคัญ ตำแหน่งและสถานที่ทำงานของผู้เขียน ความยาวไม่เกิน 1 หน้า

ตอนที่สอง ปัจจัยที่เกี่ยวข้องกับการทำวิทยานิพนธ์ (thesis or dissertation)

วิทยานิพนธ์หรือปริญญาวิทยานิพนธ์นั้นเป็นเอกสารวิชาการที่เรียบเรียงจากการศึกษาค้นคว้าอย่างเป็นระบบ เพื่อตอบปัญหาวิจัยเรื่องใดเรื่องหนึ่งโดยเฉพาะ การทำวิทยานิพนธ์ถือว่าเป็นส่วนหนึ่งในการศึกษาระดับบัณฑิตศึกษา ซึ่งนิสิตส่วนใหญ่จัดทำขึ้นเพื่อสิทธิในปริญญาบัตรตามที่สถานศึกษาได้กำหนดไว้ จุดมุ่งหมายในการทำวิทยานิพนธ์คือ การฝึกให้นิสิตได้เรียนรู้การศึกษาค้นคว้าด้วยตนเอง รู้จักวิธีการค้นคว้าอย่างมีเหตุผลและมีระเบียบ สามารถกำหนดแนวทางการศึกษาอย่างกว้างขวางและลึกซึ้ง การทำวิทยานิพนธ์เป็นกิจกรรมสำคัญที่ส่งเสริมให้เกิดทักษะในการเรียบเรียงและเสนอความคิดอย่างเป็นระบบ รวมทั้งสามารถใช้ภาษาได้อย่างถูกต้อง กระชับ ภายใต้คำแนะนำของอาจารย์ที่ปรึกษา ดังนั้นวิทยานิพนธ์จึงเป็นผลผลิตที่แสดงถึงความสามารถในการศึกษาค้นคว้าด้วยตนเองของนิสิตและแสดงถึงความสร้างสรรค์ทางวิชาการ (มหาวิทยาลัยศรีนครินทรวิโรฒ, 2527 อ้างถึงใน กริสนา นกสกุล, 2531; สิริรัตน์ คุณจักร, 2539)

การทำวิทยานิพนธ์ประกอบด้วยกิจกรรมหลายอย่าง เช่น การหาและกำหนดหัวข้อวิทยานิพนธ์ การกำหนดวัตถุประสงค์การวิจัย การศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง การเสนอโครงร่างวิทยานิพนธ์ การพัฒนาเครื่องมือวิจัย การเก็บรวบรวมข้อมูล การวิเคราะห์ข้อมูลและการเสนอรายงานวิจัยในรูปวิทยานิพนธ์ กิจกรรมเหล่านี้เป็นกิจกรรมที่นิสิตต้องทำอย่างต่อเนื่องเพื่อให้บรรลุเป้าหมาย 2 ประการ คือ วิทยานิพนธ์เสร็จตามกำหนดเวลา และเป็นงานวิจัยที่มีคุณภาพ องค์ประกอบที่เกี่ยวข้องกับความสำเร็จในการทำวิทยานิพนธ์แบ่งออกได้เป็น 4 ด้าน คือ ด้านคุณลักษณะนิสิต ด้านคุณลักษณะอาจารย์ที่ปรึกษาวิทยานิพนธ์

ด้านการให้บริการวิชาการของมหาวิทยาลัย และด้านการบริหารวิชาการของหน่วยงานที่เกี่ยวข้อง ดังรายละเอียดต่อไปนี้

3.1 คำนคุณลักษณะนิสิต

คุณลักษณะของนิสิตที่เกี่ยวข้องกับการทำวิทยานิพนธ์ มี 4 ด้าน คือ ด้านทัศนคติและนิสัย ด้านความรู้ความสามารถในการวิจัย ด้านสุขภาพ และด้านสถานภาพส่วนตัว

3.1.1 ด้านทัศนคติและนิสัย ที่ประชุม Pan Pacific Science Congress ในปี 1961 สหรัฐอเมริกา (พวงรัตน์ ทวีรัตน์, 2535) สรุปคุณลักษณะสำคัญที่ทำให้นักวิจัยประสบความสำเร็จในการทำวิจัย ไว้ว่า ต้องเป็นบุคคลที่มีความอยากรู้อยากเห็น มีทัศนคติที่ดีต่อการทำวิจัย สามารถติดต่อประสานงานกับผู้อื่นได้ดี มีความสุขและความกระตือรือร้นในการทำวิจัย มีความรอบคอบมีเหตุผล มีความเชื่อมั่นในตนเองและผลงานของตนเอง เข้าใจศักยภาพของตนเองและสามารถประเมินสิ่งต่าง ๆ ได้ถูกต้อง มีความขยันอดทนต่อการแสวงหาความรู้ กระตือรือร้นที่จะเห็นผลงานวิจัยอยู่เสมอ หมั่นฝึกเป็นคนช่างสังเกต และจดจำสิ่งต่าง ๆ มีความซื่อสัตย์และไม่เข้าข้างตนเอง

งานวิจัยของ จารีก อาจวารินทร์ (2529) ให้ผลการวิจัยสอดคล้องกับผลสรุปของที่ประชุม Pan Pacific Science Congress เมื่อทำการศึกษาสมรรถภาพของนักวิจัยทางการศึกษา ผลการวิจัยพบว่า องค์ประกอบด้านทัศนคติเชิงวิทยาศาสตร์และลักษณะที่เอื้อต่อการวิจัย เป็นคุณลักษณะที่สำคัญด้านหนึ่งของสมรรถภาพนักวิจัย ที่จะส่งผลทำให้ประสบความสำเร็จในการทำวิจัย

3.1.2 ด้านความรู้ความสามารถในการวิจัย การทำวิทยานิพนธ์นั้นนอกจากเป็นผู้มีทัศนคติและนิสัยที่ช่วยส่งเสริมในการทำวิจัยแล้ว นิสิตจะต้องมีความรู้ความสามารถในการทำวิจัย จึงจะทำให้ประสบความสำเร็จในการทำวิทยานิพนธ์ ดังที่ประชุม Pan Pacific Science Congress ในปี 1961 สหรัฐอเมริกา (พวงรัตน์ ทวีรัตน์, 2535) สรุปคุณลักษณะด้านความรู้ของนักวิจัยที่ประสบความสำเร็จในการทำวิจัยไว้ว่า จะต้องเป็นผู้มีความสามารถในการวิเคราะห์คัดเลือกความรู้จากเอกสารต่าง ๆ ได้เร็วและเหมาะสม สามารถเลือกใช้งานวิจัยได้อย่างถูกต้อง วางแผนการทำงานได้อย่างมีประสิทธิภาพ มีความรอบรู้เกี่ยวกับระเบียบวิธีการวิจัย มีความสามารถในการวิพากษ์วิจารณ์และการสังเคราะห์ สามารถนำเสนอผลการวิจัยที่เข้าใจง่ายและชัดเจน

งานวิจัยของจาร์ก อัจวารินทร์ (2529) ที่ศึกษาองค์ประกอบสมรรถภาพของนักวิจัย ให้ผลการวิจัยสอดคล้องกับผลสรุปจากที่ประชุม Pan Pacific Science Congress ผลการวิจัยพบว่า สมรรถภาพนักวิจัยทางการศึกษาในด้านความรู้ ประกอบด้วย ความรู้ความสามารถในระเบียบวิธีวิจัย ความสามารถในการเลือกและพัฒนาเครื่องมือในการเก็บรวบรวมข้อมูล ความสามารถในการใช้คอมพิวเตอร์ ความสามารถในการรายงานผลการวิจัย ความสามารถในการเลือกใช้เทคนิคการวิเคราะห์ทางสถิติ ความสามารถในการเลือกแบบวิจัย

เมื่อ สิริรัตน์ คุณจักร (2539) ศึกษาปัจจัยที่ส่งผลต่อความสำเร็จในการทำวิทยานิพนธ์ของนิสิต 2 กลุ่มที่จบการศึกษาภายในกำหนดเวลาและหลังกำหนด ผลการวิจัยพบว่านอกจากสมรรถภาพนักวิจัยด้านความรู้ที่ได้จากผลการวิจัยของ จาร์ก อัจวารินทร์แล้ว ความสามารถในการค้นคว้าเอกสารที่เกี่ยวข้องทั้งภาษาไทยและต่างประเทศ และความรู้ในเนื้อหาสาระที่ทำการวิจัย ก็เป็นสมรรถภาพในการทำวิทยานิพนธ์เช่นกัน

3.1.3 ด้านสุขภาพ การทำวิทยานิพนธ์เป็นการทำงานที่นิสิตจะต้องทุ่มเททั้งแรงกายแรงใจเพื่อให้เสร็จทันตามกำหนดเวลา และได้งานวิจัยที่มีคุณภาพ สุขภาพกายและสุขภาพจิตของนิสิตจึงเป็นส่วนสำคัญในการทำวิทยานิพนธ์ บุคคลที่มีร่างกายสมบูรณ์ มีจิตใจที่มั่นคงหนักแน่น ก็จะสามารถทำงานได้ดี ดังที่ สุชาติ โสมประยูร (2521) ได้สรุปว่า สุขภาพกับการศึกษาเป็นสิ่งที่ส่งเสริมซึ่งกันและกัน นั่นคือ ถ้าคนเรามีสุขภาพที่สมบูรณ์ดีแล้ว จะช่วยให้สามารถศึกษาเล่าเรียนได้ผลเต็มที่มากขึ้น

3.1.4 สถานภาพส่วนตัว ปัจจัยที่เกี่ยวข้องกับการทำวิทยานิพนธ์ของนิสิต คือ ลักษณะการมาศึกษาต่อ อายุ และสภาพเศรษฐกิจและการเงิน

ลักษณะการมาศึกษาต่อระดับปริญญาโท นิสิตที่ศึกษาต่อในระดับบัณฑิตศึกษามีทั้งผู้ทำงานแล้วกลับมาศึกษาต่อจนทำวิทยานิพนธ์เสร็จ หรือลามาเฉพาะที่เรียนเนื้อหาวิชาแล้วกลับไปทำงานควบคู่กับการทำวิทยานิพนธ์หรือมาเรียนโดยไม่ได้ลาศึกษาต่อ หรือเป็นผู้ที่ยังไม่ได้ทำงาน ลักษณะการมาศึกษาต่อที่แตกต่างกันเช่นนี้ก็เป็นปัจจัยส่วนหนึ่งให้นิสิตประสบความสำเร็จในการทำวิทยานิพนธ์ต่างกัน ดังที่ ศศิธร แม้นสงวน (2523) ศึกษาเปรียบเทียบภูมิลำเนาของบัณฑิตระหว่างกลุ่มนิสิตที่มีสัมฤทธิ์ผลทางการเรียนต่างกันของนักศึกษามหาวิทยาลัยรามคำแหง ผลการวิจัยพบว่า นิสิตกลุ่มที่มีสัมฤทธิ์ผลต่ำกว่าความสามารถและกลุ่มที่มีสัมฤทธิ์ผลสูงกว่าความสามารถ ในคณะเศรษฐศาสตร์ ศึกษาศาสตร์ รัฐศาสตร์ และนิติศาสตร์ มีลักษณะการมาศึกษาแตกต่างกัน โดยที่บัณฑิตที่มีสัมฤทธิ์ผลทางการเรียนสูงส่วนใหญ่จะมาเรียนอย่างสม่ำเสมอที่มหาวิทยาลัย และอนงค์ ปิยะกมลานนท์ (2530) ศึกษา

ปัจจัยที่จำแนกระหว่างนิสิตมหาบัณฑิต กลุ่มผู้ใช้เวลาดำสุดและสูงสุดของหลักสูตร ผลการวิจัยพบว่า กลุ่มนิสิตผู้ใช้เวลาดำสุด จะลามาศึกษาต่อจนสำเร็จการศึกษา ส่วนผู้ใช้เวลาสูงสุดมาเรียนขณะทำงานอยู่ หรือลาศึกษาเฉพาะที่เรียนเนื้อหาวิชาแล้วกลับไปทำงานพร้อมกับทำวิทยานิพนธ์

อายุ นิสิตที่ศึกษาในระดับบัณฑิตศึกษาส่วนใหญ่จะมีอายุแตกต่างกัน เพราะเมื่อสำเร็จการศึกษาระดับปริญญาตรี นิสิตบางส่วนจะศึกษาต่อระดับปริญญาโททันที แต่อีกบางส่วนจะทำงานก่อนแล้วจึงกลับมาเรียนต่อ ลักษณะอายุที่ต่างกันน่าจะเป็นปัจจัยที่ทำให้ นิสิต ประสบความสำเร็จในการทำวิทยานิพนธ์แตกต่างกัน จากงานวิจัยของ กฤษณีย์ อุทุมพร (2527) ที่ศึกษาองค์ประกอบที่สัมพันธ์กับสัมฤทธิ์ผลทางการศึกษาของนิสิตมหาบัณฑิต ผลการวิจัยพบว่า องค์ประกอบในด้านสถานภาพส่วนตัว อายุ เป็นปัจจัยที่มีความสัมพันธ์กับสัมฤทธิ์ผลทางการศึกษา โดยที่นิสิตที่มีอายุน้อยจะมีสัมฤทธิ์ผลทางการศึกษาสูงส่วนนิสิตที่มีอายุมากจะมีสัมฤทธิ์ผลทางการศึกษาดำ และการวิจัยของ สิริรัตน์ คุณจักร (2539) ศึกษาปัจจัยที่จำแนกกลุ่มนิสิตทำวิทยานิพนธ์ได้สำเร็จแตกต่าง 2 กลุ่ม ผลการวิจัยพบว่า กลุ่มนิสิตที่สำเร็จการศึกษาภายใน 2 ปี มีแนวโน้มว่าจะเป็นผู้ที่มีอายุมากกว่ากลุ่มที่สำเร็จการศึกษาหลัง 2 ปี

สภาพเศรษฐกิจและการเงิน เป็นปัจจัยที่เกี่ยวข้องกับความสำเร็จในการทำวิทยานิพนธ์ของนิสิต ดังที่ อนุช ปิยะกมลานนท์ (2530) ศึกษาปัจจัยที่จำแนกระหว่างนิสิตมหาบัณฑิต กลุ่มผู้ใช้เวลาดำสุดและสูงสุดของหลักสูตร ผลการวิจัยพบว่า กลุ่มนิสิตที่ทำวิทยานิพนธ์เสร็จหลังกำหนด 2 ปี ในคณะรัฐศาสตร์และคณะนิเทศศาสตร์ มีปัญหาเรื่องการเงินที่ต้องนำมาเป็นค่าใช้จ่ายในการเรียน

3.2 ด้านคุณลักษณะอาจารย์ที่ปรึกษาวิทยานิพนธ์

อาจารย์ที่ปรึกษาวิทยานิพนธ์เป็นบุคคลที่เกี่ยวข้องกับการทำวิทยานิพนธ์ของนิสิตมากที่สุด เพราะต้องเป็นผู้ที่ดูแล ให้คำปรึกษาแก่นิสิต ดังที่บัณฑิตวิทยาลัยเสนอแนะว่าอาจารย์ที่ปรึกษาวิทยานิพนธ์ควรร่วมมือวางแผนในการทำวิทยานิพนธ์กับนิสิต ให้คำแนะนำเกี่ยวกับแหล่งค้นคว้ารวบรวมข้อมูลเบื้องต้นและแหล่งเงินทุนอุดหนุนงานวิจัย ทำการตรวจสอบ แก้ไข และให้ข้อเสนอแนะการเขียนโครงการเสนอวิทยานิพนธ์ และควรให้เวลาในการปรึกษาเกี่ยวกับวิทยานิพนธ์แก่นิสิตอย่างน้อย 1 ครั้งต่อสัปดาห์ (บัณฑิตวิทยาลัย, แผนกมาตรฐานการศึกษา, 2534 อ้างถึงในขณะ กองไทรย์, 2537) คุณลักษณะอาจารย์ที่ปรึกษาวิทยานิพนธ์จึงเป็นแรงสนับสนุนสำคัญที่ทำให้ นิสิตสามารถทำวิทยานิพนธ์ได้รวดเร็ว ถูกต้อง ตรงประเด็น ยิ่งขึ้น

การที่อาจารย์ที่ปรึกษาวิทยานิพนธ์จะให้การสนับสนุนนิสิตเกี่ยวกับการทำวิทยานิพนธ์ได้ดีมากน้อยเพียงใดขึ้นอยู่กับ คุณลักษณะ 4 ด้าน คือ ด้านความรู้ของอาจารย์ที่ปรึกษาวิทยานิพนธ์ ด้านภาระงานคุมวิทยานิพนธ์ ด้านการให้คำปรึกษา และด้านความสัมพันธ์ระหว่างอาจารย์กับนิสิต

3.2.1 ด้านความรู้ของอาจารย์ที่ปรึกษาวิทยานิพนธ์ การทำวิทยานิพนธ์เป็นการฝึกให้นิสิตได้เรียนรู้การทำวิจัย ดังนั้นอาจารย์ที่ปรึกษาวิทยานิพนธ์จึงต้องเป็นผู้มีความรู้ในเรื่องระเบียบวิธีวิจัยและเรื่องที่ทำ เพื่อจะได้ให้คำแนะนำแก่นิสิต Easterby-Smith, Thorpe และ Lowe (1991) สรุปว่า การทำงานวิจัยหรือวิทยานิพนธ์จะประสบความสำเร็จได้ ต้องอาศัยการสนับสนุนจากอาจารย์ที่ปรึกษา และอาจารย์ที่ปรึกษาจะต้องมีความเชี่ยวชาญทางเทคนิควิธีและมีความรู้เกี่ยวกับการวิจัย

3.2.2 ด้านภาระงานคุมวิทยานิพนธ์ การที่นิสิตจะสามารถทำวิทยานิพนธ์ได้ประสบผลสำเร็จนั้น ส่วนหนึ่งเป็นผลมาจากอาจารย์ที่ปรึกษาวิทยานิพนธ์ให้ความช่วยเหลือเอาใจใส่ตลอดจนให้คำแนะนำปรึกษาต่าง ๆ และการที่อาจารย์ที่ปรึกษาจะสามารถให้การดูแล และให้คำแนะนำปรึกษาได้อย่างเพียงพอหรือไม่ ส่วนหนึ่งขึ้นอยู่กับภาระงานคุมวิทยานิพนธ์ หรือจำนวนนิสิตที่อาจารย์รับเป็นที่ปรึกษาวิทยานิพนธ์ ดังที่ อ่ำไพ สุจริตกุล (2523 : 145) กล่าวว่า จำนวนนิสิตที่มากเกินไปไม่เป็นสัดส่วนที่เหมาะสมกับจำนวนอาจารย์ อาจารย์บางคนจึงต้องรับเป็นที่ปรึกษาวิทยานิพนธ์แก่นิสิตหลายคน ก็อาจจะทำให้แบ่งเวลาให้คำปรึกษาแก่นิสิตได้ไม่ทั่วถึง

3.2.3 ด้านการให้คำแนะนำปรึกษา การทำวิทยานิพนธ์มีอยู่หลายขั้นตอนเป็นกระบวนการต่อเนื่อง นิสิตจะผ่านแต่ละขั้นตอนได้ ส่วนหนึ่งเป็นเพราะได้รับคำแนะนำจากอาจารย์ที่ปรึกษาวิทยานิพนธ์ การสำรวจความคิดเห็นของนิสิตเกี่ยวกับบทบาทของอาจารย์ที่ปรึกษาวิทยานิพนธ์ พบว่า การที่อาจารย์ที่ปรึกษาจัดเวลาให้นิสิตพบอย่างเพียงพอ ให้ข้อแนะนำที่มีคุณภาพตรงประเด็นที่นิสิตศึกษา จะช่วยให้การทำวิทยานิพนธ์มีประสิทธิภาพยิ่งขึ้น (จุฬาลงกรณ์มหาวิทยาลัย, บัณฑิตวิทยาลัย, 2521) ความคิดเห็นดังกล่าวของนิสิตสอดคล้องกับที่ อ่ำไพ สุจริตกุล (2523:146) กล่าวว่า การให้คำแนะนำปรึกษา การแลกเปลี่ยนความคิดเห็นและประสบการณ์ ตลอดจนทัศนคติต่าง ๆ ระหว่างนิสิตนักศึกษา กับอาจารย์ จะช่วยเพิ่มพูนคุณภาพของนักศึกษาในด้านวิชาการ สร้างความเชื่อมั่นในตนเอง และช่วยให้นักศึกษาเกิดแรงบันดาลใจในการก้าวขึ้นเป็นนักวิชาการชั้นสูง

3.2.4 ด้านความสัมพันธ์ระหว่างอาจารย์กับนิสิต การได้รับการดูแลเอาใจใส่ให้ความสนิทสนมเป็นกันเอง รวมทั้งการช่วยเหลือและให้กำลังใจจากอาจารย์ที่ปรึกษาวิทยานิพนธ์ จะเป็นแรงสนับสนุนให้นิสิตประสบความสำเร็จในการทำวิทยานิพนธ์ ดังที่ จินตนา มาพวง (2520) ศึกษาปัจจัยที่เกี่ยวข้องกับประสิทธิภาพการสอนในระดับอุดมศึกษาสาขาสังคมศาสตร์ ผลการวิจัยพบว่า ด้านความสัมพันธ์ระหว่างอาจารย์กับนิสิต การให้ความเป็นกันเองกับนิสิต และการช่วยเหลือทั้งด้านวิชาการและกิจกรรมเสริมหลักสูตร จะทำให้ประสิทธิภาพการเรียนการสอนดีขึ้น

เมื่อ กริสนา นกสกุล (2531) ศึกษาองค์ประกอบที่เกี่ยวข้องกับความสำเร็จในการทำวิทยานิพนธ์ของนิสิตบัณฑิตศึกษา ผลการวิจัยพบว่า ความสัมพันธ์ที่ดีระหว่างอาจารย์กับนิสิต และการให้กำลังใจของอาจารย์ เป็นคุณลักษณะหนึ่งของอาจารย์ที่ปรึกษาวิทยานิพนธ์ที่มีความเกี่ยวข้องกับความสำเร็จในการทำวิทยานิพนธ์

3.3 ด้านการให้บริการทางวิชาการของมหาวิทยาลัย

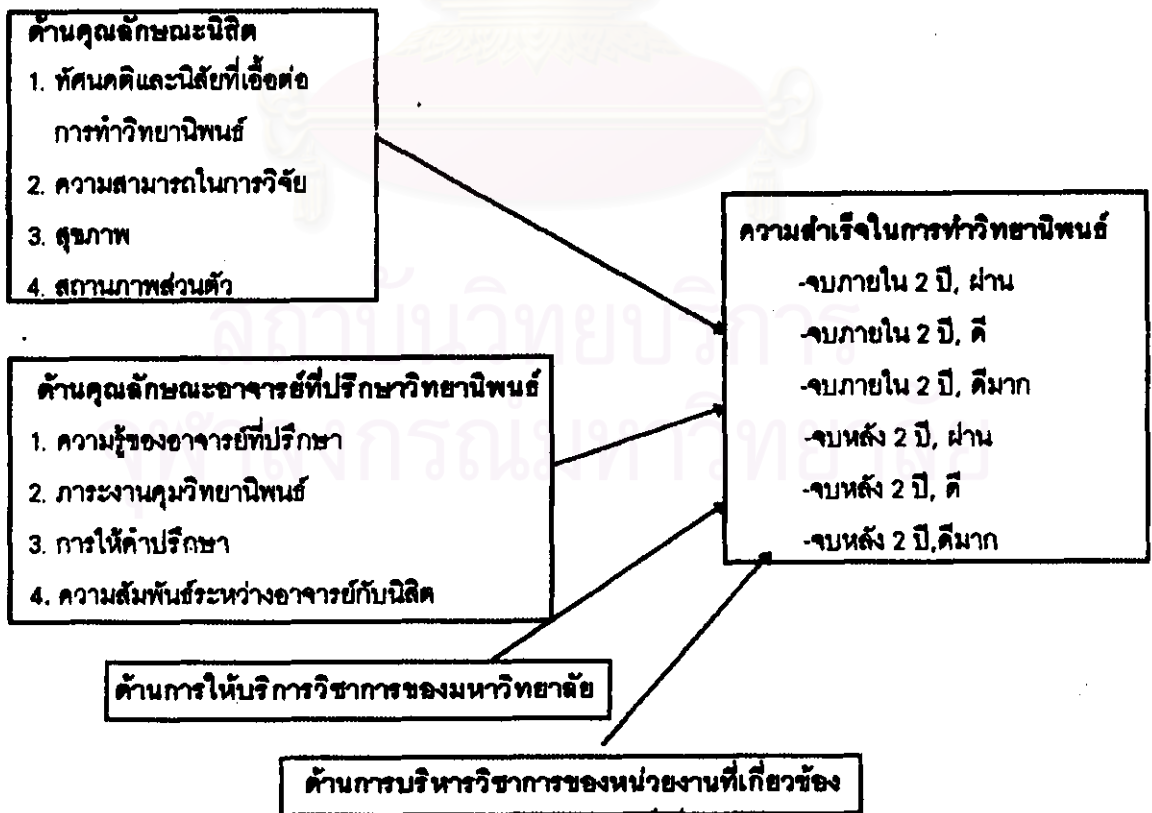
ห้องสมุดและแหล่งค้นคว้าต่าง ๆ มีความจำเป็นต่อการทำวิทยานิพนธ์มาก เพราะในการจัดทำวิทยานิพนธ์นิสิตจะต้องศึกษาค้นคว้าเอกสารตำราที่เกี่ยวข้องอย่างละเอียด เพื่อให้เกิดความเข้าใจอย่างลึกซึ้ง ดังนั้นความพร้อมของห้องสมุด ในด้านการให้บริการ และด้านปริมาณความเพียงพอของเอกสาร วัสดุ และอุปกรณ์ต่าง ๆ ในห้องสมุด จึงเป็นส่วนช่วยให้นิสิตมีความสะดวกในหาข้อมูลสนับสนุนการวิจัย ซึ่งเป็นผลทำให้นิสิตประสบความสำเร็จในการทำวิทยานิพนธ์

ณพิศร กฤตติกากุล (2521) ได้ศึกษาปัญหาในการใช้ห้องสมุดของนิสิตบัณฑิตวิทยาลัย ผลการวิจัยพบว่า นิสิตต้องการให้ห้องสมุดปรับปรุงในด้านการให้บริการ ควรขยายเวลาทำการห้องสมุดให้มากขึ้น และจัดหาบรรณารักษ์ที่มีความรู้ในการให้บริการห้องสมุด และในด้านหนังสือ เอกสาร สิ่งพิมพ์ต่าง ๆ ควรจัดหาให้มีปริมาณมากขึ้นทั้งภาษาไทยและภาษาต่างประเทศ งานวิจัยของ สิริรัตน์ คุณจักร (2539) ให้ผลสอดคล้องกับ ผลการวิจัยของ ณพิศร กฤตติกากุล เมื่อสัมภาษณ์นิสิตที่เรียนแผน ก เน้นการทำวิทยานิพนธ์ที่สำเร็จการศึกษาด้วยระยะเวลาที่ต่างกัน นิสิตทั้ง 2 กลุ่ม มีความเห็นตรงกันว่า การจัดบริการระบบการค้นหาข้อมูลที่ทันสมัย และความเพียงพอและทันสมัยของ หนังสือ ตำรา เอกสาร และงานวิจัยที่เกี่ยวข้อง เป็นปัจจัยสำคัญที่จะทำให้นิสิตค้นคว้าเอกสารและงานวิจัยที่เกี่ยวข้องได้เร็วขึ้น

3.4 ด้านการบริหารวิชาการของหน่วยงานที่เกี่ยวข้อง

ภาควิชา คณะวิชา และบัณฑิตวิทยาลัย เป็นหน่วยงานที่มีความสำคัญต่อการทำวิทยานิพนธ์ของนิสิต เพราะเป็นผู้อนุมัติโครงร่างวิทยานิพนธ์และวิทยานิพนธ์ฉบับสมบูรณ์ และเสนอแนะปรับหัวข้อวิทยานิพนธ์ให้มีความสมบูรณ์เป็นไปได้อย่างมากที่สุด ตลอดจนให้ความสะดวกในการออกหนังสือติดต่อราชการ จากที่ สมหวัง พิธิยานุวัฒน์ และ ศิริชัย กาญจนวาสี (2523) ศึกษาสภาพการทำวิทยานิพนธ์ครุศาสตรมหาบัณฑิต พบว่า กลุ่มตัวอย่างนิสิตมีความเห็นว่าบัณฑิตวิทยาลัย ควรปรับปรุงการอนุมัติหัวข้อวิทยานิพนธ์และปริญญาให้เร็วขึ้น ส่วนภาควิชา ควรมีการปฐมนิเทศและปัจฉิมนิเทศในการทำวิทยานิพนธ์แก่นิสิต ช่วยชี้แนะสำหรับโครงร่างวิทยานิพนธ์ที่ไม่ผ่าน ซึ่งถ้านิสิตได้รับการสนับสนุนที่ดีจากหน่วยงานดังกล่าว น่าจะช่วยให้นิสิตทำวิทยานิพนธ์เสร็จตามกำหนดเวลาและได้งานวิจัยที่มีคุณภาพ

สรุป ปัจจัยที่เกี่ยวข้องกับความสำเร็จในการทำวิทยานิพนธ์ประกอบด้วยองค์ประกอบ 4 ด้าน ซึ่งลักษณะความสัมพันธ์เขียนเป็นโมเดลได้ดังภาพ 1



ภาพ 1 โมเดลแสดงความสัมพันธ์ระหว่างปัจจัยที่เกี่ยวข้องกับความสำเร็จในการทำวิทยานิพนธ์ กับความสำเร็จในการทำวิทยานิพนธ์

สมมติฐานการวิจัยข้อที่ 1

จากการศึกษางานวิจัยที่เกี่ยวข้องกับความสำเร็จในการทำวิทยานิพนธ์ ผู้วิจัยตั้งสมมติฐานการวิจัยข้อที่หนึ่งว่า กลุ่มตัวแปรในองค์ประกอบทั้ง 4 ด้าน คือ ด้านคุณลักษณะนิสัย ด้านคุณลักษณะอาจารย์ที่ปรึกษาวิทยานิพนธ์ ด้านการให้บริการวิชาการของมหาวิทยาลัย และด้านการบริหารวิชาการของหน่วยงานที่เกี่ยวข้อง น่าจะเป็นตัวแปรที่สามารถจำแนกกลุ่มนิสิตมหาบัณฑิต สาขาสังคมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ที่สำเร็จการศึกษา โดยใช้ระยะเวลาในการศึกษาและมีผลการประเมินวิทยานิพนธ์แตกต่างกัน 6 กลุ่ม

ตอนที่สาม การวิเคราะห์ถดถอยโลจิสติกและการวิเคราะห์จำแนก

เนื้อหาในตอนที่สาม ผู้วิจัยนำเสนอเป็น 2 ส่วน คือ

1. การวิเคราะห์ถดถอยโลจิสติก รายละเอียดในส่วนนี้มี 4 หัวข้อ คือ แนวความคิดและหลักการวิเคราะห์ถดถอยโลจิสติก ข้อตกลงเบื้องต้น การทดสอบความกลมกลืนของโมเดล และการวิเคราะห์ถดถอยโลจิสติกด้วยโปรแกรมคอมพิวเตอร์
2. การวิเคราะห์จำแนก รายละเอียดที่นำเสนอแบ่งออกเป็น 4 หัวข้อ คือ แนวคิดและหลักการวิเคราะห์จำแนก ข้อตกลงเบื้องต้น การประเมินฟังก์ชันจำแนก และการวิเคราะห์จำแนกด้วยโปรแกรมคอมพิวเตอร์

1. การวิเคราะห์ถดถอยโลจิสติก (Logistic Regression Analysis)

1.1 แนวคิดและหลักการวิเคราะห์ถดถอยโลจิสติก

การวิเคราะห์ถดถอยโลจิสติกมีเป้าหมายคล้ายคลึงกับการวิเคราะห์ถดถอยเชิงเส้นคือ ต้องการศึกษความสัมพันธ์ระหว่างตัวแปร และพยากรณ์ค่าเฉลี่ยของตัวแปรตามจากชุดของตัวแปรต้น โดยหาโมเดลที่มีความสอดคล้องกับข้อมูลและมีความประหยัดมากที่สุด แต่การวิเคราะห์ถดถอยโลจิสติกใช้กับข้อมูลที่ตัวแปรตามเป็นตัวแปรทวิภาค มี 2 ค่า คือ 0 และ 1 ส่วนการวิเคราะห์ถดถอยเชิงเส้นที่ใช้กับข้อมูลที่ตัวแปรตามเป็นตัวแปรต่อเนื่อง (continuous variable) (สุชาติ ประสิทธิ์รัฐสินธุ์, 2537 ; ศิริเดช สุชีวะ, 2539 ; DeMaris, 1992 ; Hosmer and Lemeshow, 1989 ; Menard, 1995) Hosmer และ Lemeshow (1989) ได้เสนอตัวอย่างงาน

วิจัย เรื่องการศึกษาความสัมพันธ์ระหว่างการเป็นโรคหลอดเลือดหัวใจกับอายุ โดยได้นำเสนอข้อมูลและผลการวิเคราะห์ด้วยการถดถอยโลจิสติกเปรียบเทียบกับการวิเคราะห์ถดถอยดังต่อไปนี้

ข้อมูลในตาราง 1 ประกอบด้วยตัวแปร 4 ตัว จากกลุ่มตัวอย่างซึ่งมีประวัติการเป็นและไม่เป็นโรคหลอดเลือดหัวใจในคนที่มีอายุแตกต่างกันจำนวน 100 คน โดยกำหนดรหัส ID แทนเลขประจำตัว จาก 1-100 AGE แทนอายุ มีค่าตั้งแต่ 20 ถึง 69 AGRP แทนกลุ่มอายุ มีค่าตั้งแต่ 1 ถึง 8 และ CHD แทนการเป็นโรคหลอดเลือดหัวใจ มี 2 ค่า คือ 0 เมื่อไม่เป็นโรคหลอดเลือดหัวใจ และ 1 เมื่อเป็นโรคหลอดเลือดหัวใจ

การหาความสัมพันธ์ระหว่างตัวแปรต้นกับตัวแปรตาม อาจจะเริ่มจากการสร้างแผนภาพกระจัดกระจาย (scatterplot) เพื่อดูลักษณะและขนาดของความสัมพันธ์ ซึ่งในที่นี้เป็นแผนภาพกระจัดกระจายแสดงความสัมพันธ์ระหว่าง AGE กับ CHD ดังภาพ 2

ตามแผนภาพกระจัดกระจายภาพ 2 ทุกค่าของ AGE จะตกอยู่บนเส้นขนาน 2 เส้น ที่เป็นตัวแทนการเป็น CHD ($Y=1$) และการไม่เป็น CHD ($Y=0$) ตามแนวโน้มแล้วแต่คนที่ไม่เป็น CHD จะมีอายุน้อยกว่าคนที่ เป็น CHD แผนภาพกระจัดกระจายที่ได้นี้จะแสดงถึงลักษณะที่เป็นตัวแปรทวิภาคของตัวแปรตาม CHD อย่างชัดเจน แต่ไม่สามารถแสดงความสัมพันธ์ระหว่าง AGE กับ CHD ได้อย่างเด่นชัด เนื่องจากความผันแปรของ CHD ในทุกค่าอายุมีมาก

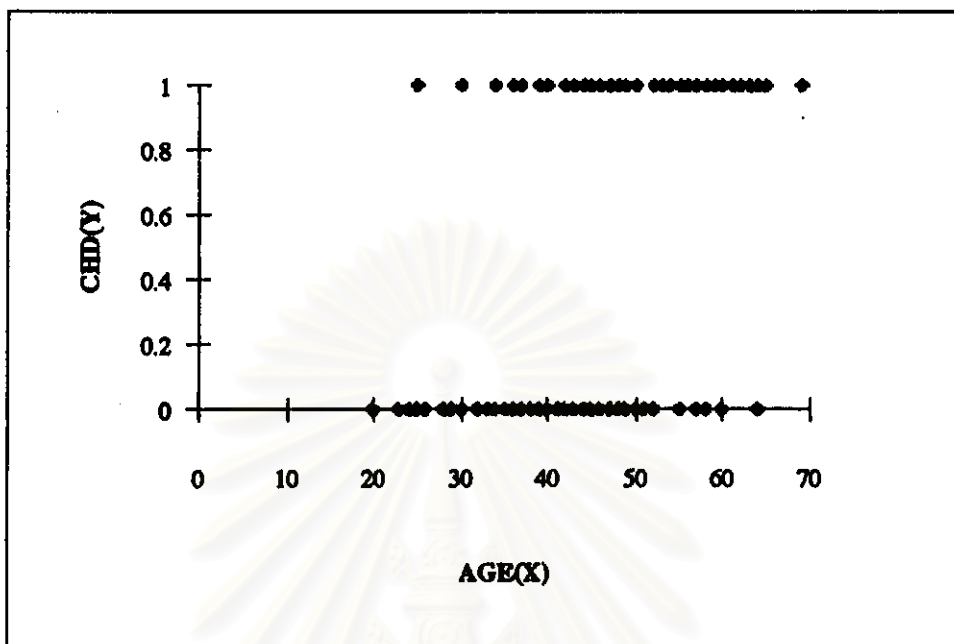
วิธีการหนึ่งที่ลดความผันแปรทำได้โดยโครงสร้างความสัมพันธ์ระหว่างตัวแปรต้นและตัวแปรตามมีลักษณะเช่นเดิม คือ สร้างอันตรภาคสำหรับตัวแปรต้น และหาค่าเฉลี่ยของตัวแปรตาม ในแต่ละชั้น ดังข้อมูลในตาราง 2 ที่ใช้ AGRP มาจัดเป็นอันตรภาคของอายุ และใช้สัดส่วนการเป็น CHD แทนค่าเฉลี่ยในแต่ละชั้น เมื่อนำมาสร้างแผนภาพกระจัดกระจายจะเห็นความสัมพันธ์ระหว่างอายุกับค่าเฉลี่ย หรือสัดส่วนการเป็น CHD ชัดเจนขึ้น นั่นคือ เมื่ออายุเพิ่มขึ้นสัดส่วนของแต่ละคนที่ เป็น CHD ก็เพิ่มขึ้นด้วย ดังภาพ 3 ซึ่งลักษณะความสัมพันธ์เช่นนี้สามารถดำเนินการวิเคราะห์ตามหลักการวิเคราะห์ถดถอยเชิงเส้นต่อไปได้

แผนภาพกระจัดกระจายตามภาพ 2 และ 3 มีลักษณะแตกต่างกัน ความสัมพันธ์ระหว่างตัวแปร AGE กับ CHD ตามภาพ 2 เป็นความสัมพันธ์แบบโมเดลความน่าจะเป็นเชิงเส้น (linear probability model) ส่วนความสัมพันธ์ระหว่างตัวแปร AGE กับสัดส่วน CHD ตามภาพ 3 เป็นความสัมพันธ์แบบโมเดลโลจิสติก (logistic model) (Menard, 1995)

ตาราง 1 อายุและการเป็นโรคหลอดเลือดหัวใจ (CHD) ของกลุ่มตัวอย่าง 100 คน

ID	AGRP	AGE	CHD	ID	AGRP	AGE	CHD	ID	AGRP	AGE	CHD
1	1	20	0	35	3	38	0	68	6	51	0
2	1	23	0	36	3	39	0	69	6	52	0
3	1	24	0	37	3	39	1	70	6	52	1
4	1	25	0	38	4	40	0	71	6	53	1
5	1	25	1	39	4	40	1	72	6	53	1
6	1	26	0	40	4	41	0	73	6	54	1
7	1	26	0	41	4	41	0	74	7	55	0
8	1	28	0	42	4	42	0	75	7	55	1
9	1	28	0	43	4	42	0	76	7	55	1
10	1	29	0	44	4	42	0	77	7	56	1
11	2	30	0	45	4	42	1	78	7	56	1
12	2	30	0	46	4	43	0	79	7	56	1
13	2	30	0	47	4	43	0	80	7	57	0
14	2	30	0	48	4	43	1	81	7	57	0
15	2	30	0	49	4	44	0	82	7	57	1
16	2	30	1	50	4	44	0	83	7	57	1
17	2	32	0	51	4	44	1	84	7	57	1
18	2	32	0	52	4	44	1	85	7	57	1
19	2	33	0	53	5	45	0	86	7	58	0
20	2	33	0	54	5	45	1	87	7	58	1
21	2	34	0	55	5	46	0	88	7	58	1
22	2	34	0	56	5	46	1	89	7	59	1
23	2	34	1	57	5	47	0	90	7	59	1
24	2	34	0	58	5	47	0	91	8	60	0
25	2	34	0	59	5	47	1	92	8	60	1
26	3	35	0	60	5	48	0	93	8	61	1
27	3	35	0	61	5	48	1	94	8	62	1
28	3	36	0	62	5	48	1	95	8	62	1
29	3	36	1	63	5	49	0	96	8	63	1
30	3	36	0	64	5	49	0	97	8	64	0
31	3	37	0	65	5	49	1	98	8	64	1
32	3	37	1	66	6	50	0	99	8	65	1
33	3	37	0	67	6	50	1	100	8	69	1
34	3	38	0								

ที่มา : Hosmer, D. W., and Lemeshow, S. (1989).



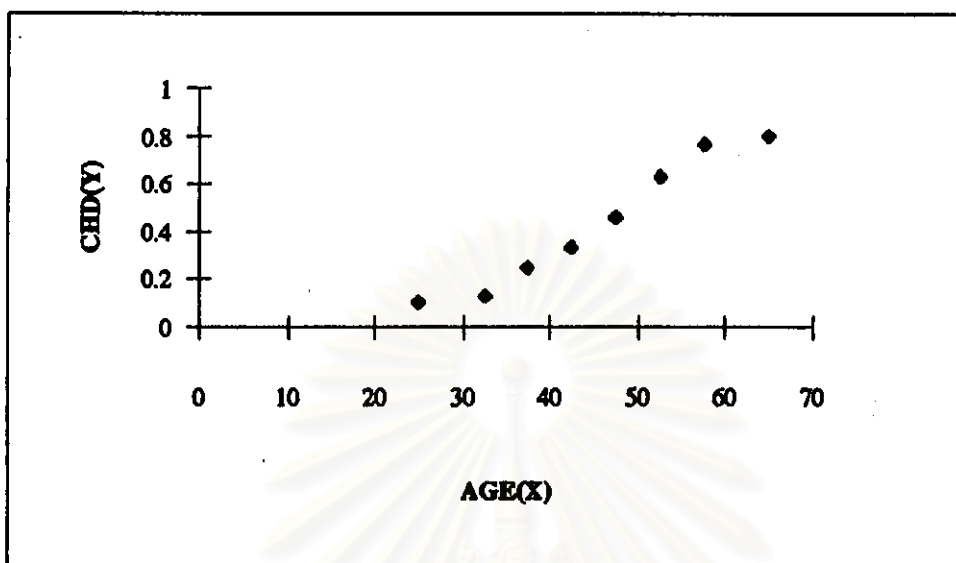
ภาพ 2 แผนภาพการจัดกระจายของ CHD ตามอายุ

ที่มา : Hosmer, D. W., and Lemeshow, S. (1989).

ตาราง 2 ตารางความถี่ของ CHD ตามกลุ่มอายุ

กลุ่มอายุ	การเป็นโรคหลอดเลือดหัวใจ			ค่าเฉลี่ย(สัดส่วน)
	จำนวน	ไม่เป็น	เป็น	
20-29	10	9	1	0.10
30-34	15	13	2	0.13
35-39	12	9	3	0.25
40-44	15	10	5	0.33
45-49	13	7	6	0.46
50-54	8	3	5	0.63
55-59	17	4	13	0.76
60-69	10	2	8	0.80
รวม	100	57	43	0.43

ที่มา : Hosmer, D. W., and Lemeshow, S. (1989).



ภาพ 3 แผนภาพการจัดกระจายของค่าเฉลี่ย CHD ในแต่ละกลุ่มอายุ

ที่มา : Hosmer, D. W., and Lemeshow, S. (1989).

การวิเคราะห์ถดถอยโลจิสติกมีความแตกต่างจากการวิเคราะห์ถดถอยเชิงเส้นที่สำคัญ 2 ประการ (วีรพันธ์ พงศาภักดิ์, 2537 ; ศิริเดช สุชีวะ, 2539 ; Hosmer and Lemeshow, 1989) คือ

ประการแรก ค่าเฉลี่ยแบบมีเงื่อนไข (conditional mean) ซึ่งเป็นค่าเฉลี่ยของตัวแปรตามที่ถูกกำหนดโดยตัวแปรต้น ค่าเฉลี่ยแบบมีเงื่อนไขใช้สัญลักษณ์ $E(y | x)$ เมื่อ Y เป็นตัวแปรตาม และ X เป็นค่าของตัวแปรต้น ในการวิเคราะห์ถดถอยเชิงเส้นค่าเฉลี่ยนี้สามารถแสดงในรูปของสมการเชิงเส้นของ x ได้

$$E(y | x) = \beta_0 + \beta_1 x$$

หมายความว่า $E(y | x)$ อาจจะเป็นค่าใดๆ ณ ค่า $-\infty < x < +\infty$ ที่กำหนดให้

แต่เมื่อตัวแปรตามเป็นตัวแปรตัวแปรทวิภาค ค่า $E(y | x)$ จะต้องมามีค่าตั้งแต่ 0 ถึง 1 ($0 < E(y | x) < 1$) ตามตัวอย่างข้างต้น ค่าประมาณของ $E(y | x)$ ก็คือ ค่าสัดส่วนการเป็น CHD เมื่อพิจารณาจากภาพ 3 จะเห็นว่าค่าสัดส่วนการเป็น CHD เมื่อตัวแปรอายุเปลี่ยนไป 1 หน่วย จะเข้าใกล้ 0 และ 1 ที่ละน้อย กราฟที่ได้จะมีลักษณะเป็นรูปตัวเอส (S-shaped) คล้ายกับรูปของการแจกแจงแบบสะสมของตัวแปรสุ่มซึ่งเป็นการแจกแจงโลจิสติก

เมื่อกำหนด $\pi(x) = E(y | x)$ โมเดลถดถอยโลจิสติกสามารถเขียนได้ดังนี้

$$\pi(x) = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}}$$

ในการวิเคราะห์ถดถอยโลจิสติก นักวิจัยจะแปลงค่า $\pi(x)$ ให้อยู่ในรูปของแอดัมต่อ (odds) ได้ และแปลงค่าแอดัมต่อให้อยู่ในรูปของโลจิต (logit) เพื่อให้ได้สมการถดถอยที่มีความสัมพันธ์แบบเส้นตรงได้ดังสมการ

$$\begin{aligned} \text{odds } (Y=1) &= \frac{\pi(x)}{1 - \pi(x)} \\ &= e^{\beta_0 + \beta_1 x} \end{aligned}$$

สมการในรูปโลจิต

$$\begin{aligned} g(x) &= \ln \left| \frac{\pi(x)}{1 - \pi(x)} \right| \\ &= \beta_0 + \beta_1 x \end{aligned}$$

ในที่นี้ logit หรือ $g(x)$ ที่ได้เป็นฟังก์ชันเชิงเส้นที่อาจมีค่าต่อเนื่องตั้งแต่ $-\infty$ ถึง $+\infty$ ก็ได้ขึ้นอยู่กับพิสัยของ x

ความแตกต่างประการที่สองคือ การแจกแจงแบบมีเงื่อนไขของตัวแปรตาม ในการวิเคราะห์ถดถอยเชิงเส้น ค่าสังเกตได้ของตัวแปรตามแสดงได้ตามสมการ $y = E(y | x) + \varepsilon$ เมื่อ ε เป็นความคลาดเคลื่อน เทอมความคลาดเคลื่อนต้องมีการแจกแจงแบบปกติด้วยค่าเฉลี่ยเป็นศูนย์และความแปรปรวนคงที่ในทุกตำแหน่งของตัวแปรต้น ซึ่งเป็นผลจากการแจกแจงแบบมีเงื่อนไขของตัวแปรตามที่ถูกกำหนดโดยค่า X เป็นแบบปกติด้วยค่าเฉลี่ย $E(y | x)$ และความแปรปรวนคงที่ แต่ในการวิเคราะห์ถดถอยโลจิสติก ค่าสังเกตได้ของตัวแปรตามแสดงตามสมการ $y = \pi(x) + \varepsilon$ เทอมความคลาดเคลื่อนจะมี 2 ค่า คือ ถ้า $Y=1$ จะทำให้ $\varepsilon = 1 - \pi(x)$ ถ้า $Y=0$ จะทำให้ $\varepsilon = -\pi(x)$ ดังนั้นเทอมความคลาดเคลื่อนจึงมีการแจกแจงแบบทวินาม (binomial distribution) มีค่าเฉลี่ยเป็นศูนย์และความแปรปรวนเท่ากับ $\pi(x)[1 - \pi(x)]$ ตามการแจกแจงแบบมีเงื่อนไขของตัวแปรตามที่เป็นแบบทวินามด้วยค่าความน่าจะเป็นที่กำหนดโดยค่าเฉลี่ยแบบมีเงื่อนไข

โดยสรุป ในการวิเคราะห์ถดถอยที่มีตัวแปรตามเป็นตัวแปรทวิภาค มีลักษณะสำคัญ 3 ประการ คือ ประการที่หนึ่ง ค่าเฉลี่ยแบบมีเงื่อนไขของสมการถดถอยต้องปรับให้มีค่าตั้งแต่ 0 ถึง 1 เพื่อที่จะใช้การวิเคราะห์ถดถอยโลจิสติก ประการที่สอง เทอมความคลาดเคลื่อนต้องมีการแจกแจงแบบทวินาม ประการที่สาม หลักการวิเคราะห์ที่ใช้ในการวิเคราะห์ถดถอยเชิงเส้นสามารถนำมาใช้ในการวิเคราะห์ถดถอยโลจิสติก

ในกรณีที่มีตัวแปรต้นหลายตัวโมเดลการวิเคราะห์ถดถอยโลจิสติก สามารถเขียนได้ดังนี้ (Hosmer and Lemeshow, 1989: 25-26)

$$g(x_1, x_2, \dots, x_p) = \ln \left| \frac{\pi(x_1, x_2, \dots, x_p)}{1 - \pi(x_1, x_2, \dots, x_p)} \right|$$

$$= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p$$

และเทอมของค่าเฉลี่ยแบบมีเงื่อนไขแสดงได้ดังสมการ

$$\pi(x_1, x_2, \dots, x_p) = \frac{e^{g(x_1, x_2, \dots, x_p)}}{1 + e^{g(x_1, x_2, \dots, x_p)}}$$

การทดสอบว่าค่าสัมประสิทธิ์ของตัวแปรต้นแต่ละตัวว่าเท่ากับ 0 หรือไม่ พิจารณาได้จากค่า Wald statistics ซึ่งเป็นสถิติที่มีการแจกแจงแบบไคสแควร์ สามารถคำนวณได้จากกำลังสองของอัตราส่วนระหว่างค่าสัมประสิทธิ์กับค่าความคลาดเคลื่อนมาตรฐานของค่าสัมประสิทธิ์ $(b/SE)^2$

การแปลความหมายสัมประสิทธิ์ถดถอยที่ได้จากการวิเคราะห์ถดถอยโลจิสติกจะแตกต่างจากการแปลความหมายในการวิเคราะห์ถดถอยเชิงเส้น เนื่องจากในการวิเคราะห์ถดถอยเชิงเส้นเป็นการประมาณค่าของตัวแปรตามโดยตรง ค่าสัมประสิทธิ์ที่ได้จึงบอกถึงการเปลี่ยนแปลงค่าของตัวแปรตาม เมื่อตัวแปรต้นเปลี่ยนแปลงไป 1 หน่วย แต่การวิเคราะห์ถดถอยโลจิสติกใช้ทำนายโอกาสในการเกิดหรือไม่เกิดเหตุการณ์ของตัวแปรตาม ซึ่งอยู่ในรูปของแอดด์ส (odds) ในที่นี้ แอดด์ส หมายถึงสัดส่วนระหว่างความน่าจะเป็นในการเกิดเหตุการณ์กับความน่าจะเป็นในการไม่เกิดเหตุการณ์ ดังนั้นโมเดลโลจิสติกจึงสามารถเขียนได้ตามสมการโลจิสติกได้ดังนี้

$$\ln \left| \frac{\text{prob(event)}_i}{\text{prob(no event)}} \right| = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p$$

การแปลความหมายจากสมการข้างบน ค่าสัมประสิทธิ์ถดถอยโลจิสติกจะบอกถึงการเปลี่ยนแปลงของ ค่าลอการิทึมของแอดัมต่อ เมื่อตัวแปรต้นเปลี่ยนแปลงไปหนึ่งหน่วย

ในบางครั้งโมเดลโลจิสติกก็อาจเขียนในรูปของ แอดัมต่อ ดังนี้

$$\frac{\text{prob(event)}_i}{\text{prob(no event)}} = e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p} = e^{\beta_0} e^{\beta_1 x_1} e^{\beta_2 x_2} \dots e^{\beta_p x_p}$$

การแปลความหมายตามสมการนี้ก็จะแตกต่างจากการแปลค่าตามสมการที่แอดัมต่ออยู่ในรูปลอการิทึม คือ ค่า e ยกกำลัง β_i เป็นค่าแอดัมต่อ ที่เปลี่ยนแปลง เมื่อตัวแปรต้นตัวที่ i มีค่าเปลี่ยนแปลงไป 1 หน่วย ถ้าค่า β_i เป็นบวก เทอมนี้จะมีค่ามากกว่า 1 แอดัมต่อจะเพิ่มขึ้น แต่ถ้าค่า β_i เป็นลบ เทอมนี้จะมีค่าน้อยกว่า 1 แอดัมต่อจะลดลง ถ้าค่า $\beta_i = 0$ เทอมนี้จะมีค่าเท่ากับ 1 แอดัมต่อก็จะไม่เปลี่ยนแปลง (ศิริเดช สุชีวะ, 2539 ; Menard, 1995 ; Ziek , Beardsley, Deren and Tortu, 1996)

1.1.1 การวิเคราะห์ถดถอยโลจิสติกพหุวิภาค (Polytomous Logistic Regression Analysis)

จากที่ได้เสนอมมาแล้วเป็นแนวคิดในการวิเคราะห์ถดถอยโลจิสติกที่มีตัวแปรตามเป็นตัวแปรพหุวิภาค แต่การวิเคราะห์ถดถอยโลจิสติกยังสามารถประยุกต์ใช้กับตัวแปรตามที่เป็นตัวแปรพหุวิภาค คือ มีมากกว่า 2 ค่าและมีระดับการวัดแบบนามบัญญัติ การวิเคราะห์ถดถอยโลจิสติก ที่นำมาใช้กับข้อมูลลักษณะนี้เรียกว่า การวิเคราะห์ถดถอยโลจิสติกพหุวิภาค (polytomous logistic regression analysis หรือ multinomial logistic regression analysis) โดยกำหนดให้ตัวแปรตามเป็นตัวแปรดัมมี่ (dummy variable) เมื่อตัวแปรตามมี k กลุ่ม (ค่า) ต้องมีตัวแปรดัมมี่ $k-1$ ตัว ที่แต่ละตัวจะมีค่าเท่ากับ 1 เมื่อมีหน่วยตัวอย่างเป็นสมาชิกในกลุ่มนั้น ทำการเปรียบเทียบตัวแปรตามแต่ละกลุ่มกับกลุ่มที่เลือกเป็นกลุ่มอ้างอิง (reference group) โดยใช้สมการโลจิส ซึ่งจำนวนสมการโลจิสจะมีได้เท่ากับจำนวนกลุ่มตัวแปรตามลบ 1 หรือ $k-1$ (สุชาติ ประสิทธิ์รัฐสินธุ์, 2537 ; Begg and Gray, 1984 ; DeMaris, 1992 ; Hosmer and Lemeshow, 1989 ; Menard, 1995)

สมการโลจิสติกสามารถเขียนได้ดังนี้ (Menard, 1995 : 80)

$$g_h(x_1, x_2, \dots, x_p) = e^{\beta_{h0} + \beta_{h1}x_1 + \beta_{h2}x_2 + \dots + \beta_{hp}x_p}$$

$$h = 1, 2, \dots, k-1$$

h แทนค่าของตัวแปรตาม (Y) เมื่อมีการวิเคราะห์เปรียบเทียบกับกลุ่มอ้างอิง

$$g_0(x_1, x_2, \dots, x_p) = 1$$

สำหรับค่าความน่าจะเป็นแบบมีเงื่อนไขของตัวแปรตาม เท่ากับค่าใด ๆ ของ h ที่หักออกจากค่าของกลุ่มอื่น และกลุ่มอ้างอิง เขียนได้ตามสมการ (Menard, 1995 : 81)

$$P(Y=h | x_1, x_2, \dots, x_p) = \frac{e^{\beta_{h0} + \beta_{h1}x_1 + \beta_{h2}x_2 + \dots + \beta_{hp}x_p}}{1 + \sum e^{\beta_{h0} + \beta_{h1}x_1 + \beta_{h2}x_2 + \dots + \beta_{hp}x_p}}$$

ค่าความน่าจะเป็นแบบมีเงื่อนไขของตัวแปรตามที่เป็นกลุ่มอ้างอิง ($h_0 = k$ หรือ $h_0 = 0$)

$$P(Y=h_0 | x_1, x_2, \dots, x_p) = \frac{1}{1 + \sum e^{\beta_{h0} + \beta_{h1}x_1 + \beta_{h2}x_2 + \dots + \beta_{hp}x_p}}$$

ตัวอย่างเช่น ตัวแปรตามมี 3 กลุ่ม ให้รหัสเป็น $Y=0$, $Y=1$ และ $Y=2$ ตามลำดับ กำหนดให้ กลุ่ม $Y=0$ เป็นกลุ่มอ้างอิง ตามตัวอย่างนี้จะมีสมการโลจิสติกอยู่ 2 สมการ คือ สมการที่ 1 เป็นสมการโลจิสติกระหว่าง $Y=1$ กับ $Y=0$ ซึ่งสามารถเขียนได้ดังนี้

$$g_1(x_1, x_2, \dots, x_p) = \ln \left| \frac{P(Y=1 | x_1, x_2, \dots, x_p)}{P(Y=0 | x_1, x_2, \dots, x_p)} \right|$$

$$= \beta_{10} + \beta_{11}x_1 + \beta_{12}x_2 + \dots + \beta_{1p}x_p$$

สมการที่ 2 เป็นสมการโลจิสติกระหว่าง $Y=2$ กับ $Y=0$ แสดงได้ดังนี้

$$g_2(x_1, x_2, \dots, x_p) = \ln \left| \frac{P(Y=2 | x_1, x_2, \dots, x_p)}{P(Y=0 | x_1, x_2, \dots, x_p)} \right|$$

$$= \beta_{20} + \beta_{21}x_1 + \beta_{22}x_2 + \dots + \beta_{2p}x_p$$

โดยที่ค่าของ x_i ในสมการอาจจะเป็นตัวแปรต่อเนื่องหรือตัวแปรคัมมีก็ได้

สำหรับสมการโลจิสที่ใช้เปรียบเทียบระหว่าง $Y = 2$ กับ $Y = 1$ หาได้จากความแตกต่างของค่าสัมประสิทธิ์ถดถอยโลจิสติกระหว่างสมการโลจิสที่ 2 กับสมการโลจิสที่ 1 (Hosmer and Lemeshow, 1989)

การประมาณค่าความน่าจะเป็นแบบมีเงื่อนไข (conditional probabilities) ตามกลุ่มของตัวแปรตามสามารถคำนวณได้จากสมการ (Hosmer and Lemeshow, 1989)

$$P(Y=0|x_1, x_2, \dots, x_p) = \frac{1}{1 + e^{g_1(x_1, x_2, \dots, x_p)} + e^{g_2(x_1, x_2, \dots, x_p)}}$$

$$P(Y=1|x_1, x_2, \dots, x_p) = \frac{e^{g_1(x_1, x_2, \dots, x_p)}}{1 + e^{g_1(x_1, x_2, \dots, x_p)} + e^{g_2(x_1, x_2, \dots, x_p)}}$$

$$P(Y=2|x_1, x_2, \dots, x_p) = \frac{e^{g_2(x_1, x_2, \dots, x_p)}}{1 + e^{g_1(x_1, x_2, \dots, x_p)} + e^{g_2(x_1, x_2, \dots, x_p)}}$$

การแปลความหมายค่าสัมประสิทธิ์ถดถอยโลจิสติกพหุภาคในแต่ละสมการโลจิสมีลักษณะคล้ายกับการแปลความหมายค่าสัมประสิทธิ์ถดถอยโลจิสติกโดยทั่วไป เพียงแต่ค่าสัมประสิทธิ์ถดถอยโลจิสติกพหุภาคที่ได้จะบอกถึงการเปลี่ยนแปลงของค่าลอการิทึมของดั้มต่อ (ในที่นี้ ดั้มต่อ จะหมายถึงสัดส่วนระหว่างความน่าจะเป็นการเป็นสมาชิกกลุ่มนั้นกับความน่าจะเป็นการเป็นสมาชิกในกลุ่มที่เลือกเป็นกลุ่มอ้างอิง) เมื่อตัวแปรต้นเปลี่ยนแปลงไปหนึ่งหน่วย (สุชาติ ประสิทธิ์รัฐสินธ์ , 2537 : 210 ; Begg and Gray, 1984 ; Menard, 1995 : 81)

1.2 ข้อตกลงเบื้องต้น

การวิเคราะห์ถดถอยโลจิสติกมีข้อตกลงเบื้องต้นเกี่ยวกับลักษณะของตัวแปร 2 ประการ คือ ประการแรก ตัวแปรตามเป็นตัวแปรทวิภาค หรือตัวแปรพหุภาค ที่มีการแจกแจงแบบทวินาม (binomial distribution) ประการที่สอง ในแต่ละกลุ่มที่แบ่งตามค่าตัวแปรตาม ตัวแปรต้นมีการแจกแจงแบบปกติพหุด้วยเมทริกซ์ความแปรปรวนร่วมเท่ากัน (multivariate normally distributed with equal covariance matrices) หรือเป็นตัวแปรทวิภาคที่เป็นอิสระต่อกัน

(multivariate independent dichotomous variables) หรืออาจจะเป็นตัวแปรไม่ต่อเนื่องที่มีความสัมพันธ์เชิงเส้นกับตัวแปรตามแบบล็อกลิเนียร์ชนิดที่มีอิทธิพลลำดับที่สองหรือสูงกว่าในแต่ละกลุ่มประชากร (discrete variable following a loglinear model with second and higher-order effects in each population) หรือเป็นตัวแปรที่มีลักษณะตามที่กล่าวมารวมกัน (ธวัชชัย วรพงศธร, 2530 อ้างถึงในพัชรี เพ็งประโคน, 2539 ; Anderson, 1972)

1.3 การทดสอบความกลมกลืน (goodness of fit) ของโมเดล

การประมาณค่าพารามิเตอร์ในการวิเคราะห์ถดถอยโลจิสติกใช้วิธีความเป็นไปได้สูงสุด (maximum likelihood) ซึ่งเป็นการคำนวณทวนซ้ำ เพื่อให้ค่าประมาณพารามิเตอร์มีค่าใกล้เคียงกับข้อมูลเชิงประจักษ์มากที่สุด การทดสอบความสอดคล้องกลมกลืนระหว่างโมเดลกับข้อมูลเชิงประจักษ์สามารถพิจารณาได้จากค่าสถิติที่มีการกระจายแบบไคสแควร์ (Teiatia,1980) ค่าสถิติที่สำคัญ ได้แก่ (ศิริเดช สุชีวะ, 2539 ; DeMaris, 1992 ; Hosmer and Lemeshow, 1989 ; Menard, 1995)

1. ค่า -2 LL likelihood มีลักษณะคล้ายกับค่าผลรวมความคลาดเคลื่อนกำลังสอง (error sum of squares, SSE) ในการวิเคราะห์ถดถอยพหุคูณ ใช้ในการทดสอบว่าโมเดลมีความสอดคล้องกับข้อมูลหรือไม่ ภายใต้สมมติฐานศูนย์ว่า โมเดลมีความสอดคล้องกับข้อมูล

$$D_0 = -2 \{ n_{y=1} \ln [P(Y=1)] + (N - n_{y=1}) \ln [1 - P(Y=1)] \}$$

$$= -2 \{ (n_{y=1}) \ln [P(Y=1)] + (n_{y=0}) \ln [P(Y=0)] \}$$

2. ค่า Model Chi-Square (G_M) เป็นค่าสถิติที่ใช้ทดสอบว่าสัมประสิทธิ์ของตัวแปรทุกตัวในโมเดลยกเว้นค่าคงที่มีค่าเท่ากับศูนย์หรือไม่ สถิตินี้มีลักษณะคล้ายกับ F-ratio ในการวิเคราะห์ถดถอยพหุคูณ สามารถคำนวณได้จากการนำ ค่า -2 LL likelihood ของโมเดลที่มีเฉพาะค่าคงที่ (D_0) ลบด้วยค่า -2 LL likelihood ของโมเดลที่นำตัวแปรต้นเข้าวิเคราะห์ด้วย (D_M) นั่นคือ

$$G_M = D_0 - D_M$$

3. ค่า R^2_L เป็นค่าสถิติที่แสดงถึงสัดส่วนความแปรปรวนในตัวแปรตามที่สามารถอธิบายได้ด้วยชุดตัวแปรต้น คล้ายกับค่า R^2 ในการวิเคราะห์ถดถอยพหุคูณ สามารถคำนวณได้จากสูตร

$$R^2_L = G_M / D_0 = G_M / (G_M + D_M)$$

นอกจากนี้ความสัมพันธ์ระหว่างโมเดลกับข้อมูลเชิงประจักษ์ยังสามารถทำได้ด้วยการพิจารณาจากร้อยละของจำนวนหน่วยวิเคราะห์ที่โมเดลสามารถทำนายได้ถูกต้อง

1.4 การวิเคราะห์ถดถอยโลจิสติกด้วยโปรแกรมคอมพิวเตอร์

1.4.1 การวิเคราะห์ถดถอยโลจิสติกสามารถทำได้โดยตรงด้วยโปรแกรมสำเร็จ SPSS SAS ที่มีวิธีการเลือกตัวแปรต้นเข้าสมการ 2 วิธี คือ วิธีโดยตรงที่นำตัวแปรต้นทุกตัวเข้าสมการ และวิธีวิเคราะห์แบบมีลำดับขั้นที่พิจารณาเลือกตัวแปรเข้าสมการทีละตัว และพิจารณาผลของตัวแปรที่นำเข้าสมการแต่ละขั้นตอนด้วย ซึ่งสามารถทำได้ 2 แบบ คือ แบบ forward stepwise และแบบ backward stepwise (Marija, 1992)

1.4.2 การวิเคราะห์ถดถอยโลจิสติกพหุวิภาคในโปรแกรมคอมพิวเตอร์ เนื่องจากโปรแกรมคอมพิวเตอร์ไม่สามารถที่จะทำการวิเคราะห์ถดถอยโลจิสติกพหุวิภาค (กรณีตัวแปรตาม เป็นตัวแปรที่ระดับการวัดแบบนามบัญญัติ) ได้โดยตรง ดังนั้นในการวิเคราะห์ถดถอยโลจิสติกพหุวิภาคจึงต้องประยุกต์ใช้โปรแกรมการวิเคราะห์ถดถอยโลจิสติกเข้ามาช่วยวิเคราะห์ โดยกำหนดให้กลุ่มใดกลุ่มหนึ่งเป็นกลุ่มอ้างอิง แล้วทำการวิเคราะห์ทีละโมเดล เช่น กรณีที่ตัวแปรตามมี 3 กลุ่ม คือ กลุ่ม 1, 2 และ 3 ในการวิเคราะห์อาจกำหนดให้กลุ่ม 3 เป็นกลุ่มอ้างอิง ก็จะได้โมเดลการวิเคราะห์ 2 โมเดล คือ โมเดลที่ 1 สร้างสมการโลจิสระหว่างกลุ่ม 1 กับกลุ่ม 3 โมเดลที่ 2 สร้างสมการโลจิสระหว่างกลุ่ม 2 กับกลุ่ม 3 ส่วนอัตราส่วนแค้นต่อระหว่างกลุ่ม 1 กับกลุ่ม 3 หาได้จากผลต่างค่าสัมประสิทธิ์ถดถอยโลจิสติกระหว่างโมเดลที่ 1 กับโมเดลที่ 2 ซึ่งวิธีนี้จะทำให้สูญเสียประสิทธิภาพในการวิเคราะห์น้อย และความสอดคล้องของโมเดลที่ได้ใกล้เคียงกับการวิเคราะห์ที่ให้ตัวแปรตามเป็นตัวแปรพหุวิภาค (Begg and Gray, 1984 ; Hosmer and Lemeshow, 1989) ส่วน Menard (1995) ได้นำเสนอการวิเคราะห์โดยการประยุกต์ใช้โปรแกรม SPSS LOGLINEAR และ SAS PROC CATMOD สำหรับโมเดลโลจิสต์ในเมียล (multinomial logit model) ในการคำนวณโมเดลถดถอยโลจิสติก

พหุภาค ซึ่งจะได้ค่า G_{χ^2} , R^2 , ค่าสัมประสิทธิ์ประสิทธิ์ถดถอยโลจิสติก และตารางการทำนาย แต่ขั้นตอนการวิเคราะห์จะซับซ้อนกว่าการวิเคราะห์ถดถอยโลจิสติกโดยทั่วไป และการวิเคราะห์ด้วยโปรแกรม SPSS LOGLINEAR มีข้อจำกัด คือ สามารถวิเคราะห์ตัวแปรต้นได้ไม่เกิน 10 ตัวแปร (Marija, 1992)

อย่างไรก็ตามการใช้โปรแกรมคอมพิวเตอร์คำนวณจะได้ผลที่ดีและมีความน่าเชื่อถือวิจัยควรพิจารณาให้รอบคอบในประเด็นต่อไปนี้ ประเด็นที่หนึ่ง ต้องพิจารณาถึงความเหมาะสมของโมเดลโลจิสติกกับลักษณะความสัมพันธ์ระหว่างตัวแปรต้นและตัวแปรตาม ประเด็นที่สอง การเลือกตัวแปรอิสระควรพิจารณาเฉพาะตัวแปรที่มีความสัมพันธ์เชิงสาเหตุกับตัวแปรตามมากกว่าที่จะให้โปรแกรมคอมพิวเตอร์ช่วยเลือกให้อย่างเดียว ประเด็นที่สาม ลักษณะของข้อมูลควรเป็นไปตามข้อตกลงเบื้องต้น ถ้ามีการละเมิดข้อตกลงเบื้องต้นมากจะมีผลกระทบต่อโมเดลที่ใช้ ซึ่งมีผลต่อการประมาณค่า β_j และมีผลกระทบทำให้การแปลผลผิดพลาดด้วย ประเด็นที่สี่ ตัวแปรต้นไม่ควรมีความสัมพันธ์กันมากเพราะจะทำให้เกิดภาวะร่วมเส้นตรงพหุ ประเด็นที่ห้า จำนวนหน่วยวิเคราะห์มีน้อยเกินไปจะมีผลต่อความน่าเชื่อถือของการประมาณค่าสถิติและการทดสอบค่าไคสแควร์ (พัชรี เพ็งประโคน, 2539) ควรจะต้องมีจำนวนหน่วยวิเคราะห์มากกว่าตัวแปร 2 ถึง 3 เท่า (สุชาติ ประสิทธิ์รัฐสินธุ์, 2537: 207)

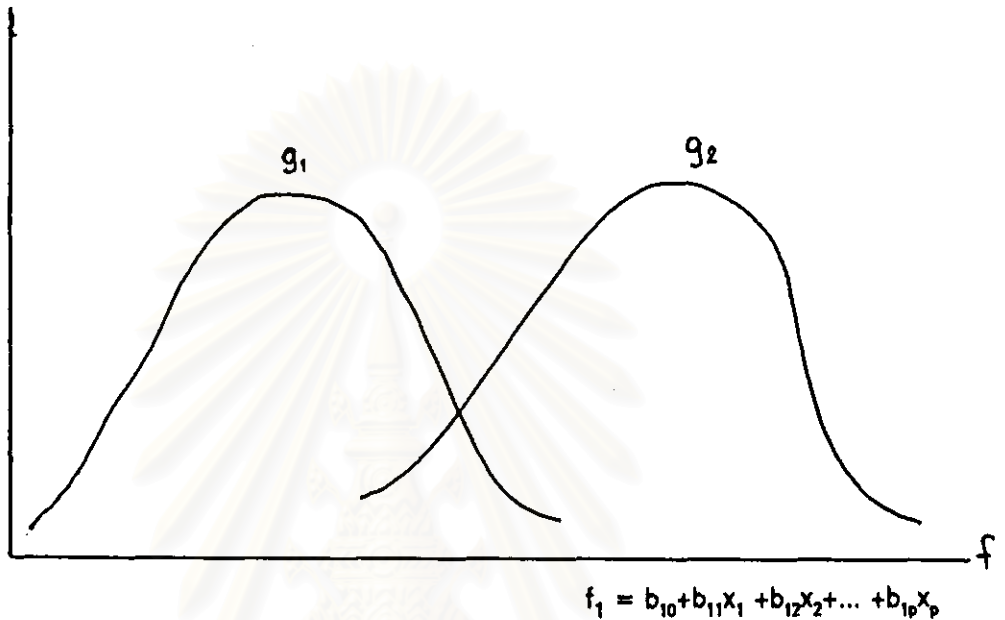
2. การวิเคราะห์จำแนก (Discriminant Analysis)

2.1 แนวคิดและหลักการวิเคราะห์จำแนก

2.1.1 แนวคิดในการวิเคราะห์จำแนก

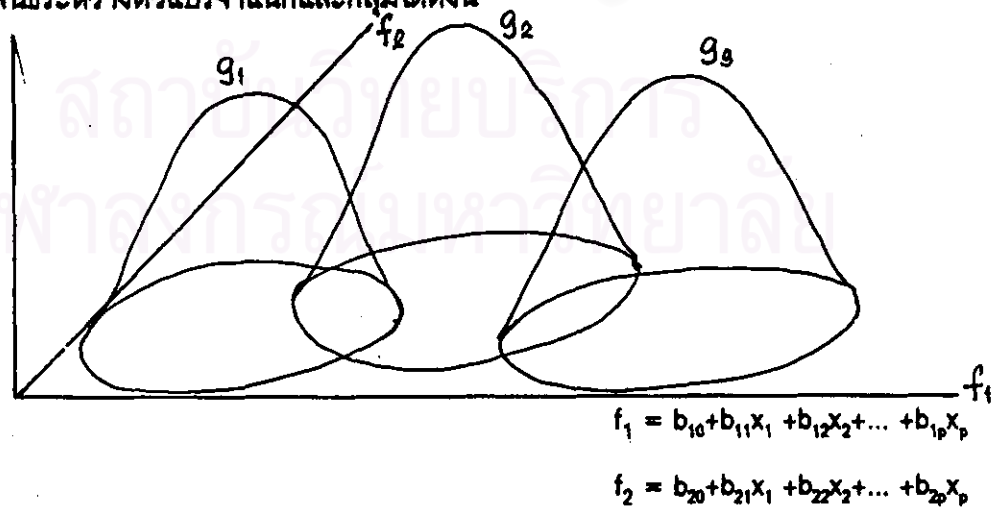
การวิเคราะห์จำแนกเป็นวิธีการทางสถิติที่ใช้ศึกษาความแตกต่างระหว่างกลุ่มตั้งแต่ 2 กลุ่มขึ้นไปที่เป็นผลมาจากตัวแปรหลายตัว โดยมุ่งหาชุดตัวแปรจำแนก (discriminating variables) ที่ทำให้ความแตกต่างระหว่างกลุ่มมีค่าสูงสุด และจำแนกหน่วยวิเคราะห์เข้ากลุ่ม (มาณี ไชยธีรานุกัมศิริ, 2528 ; สาราวุฒิ ตูวินันท์, 2529 ; สุชาติ ประสิทธิ์รัฐสินธุ์, 2537 ; Klecka, 1987 ; Lindeman, Merenda and Gold, 1980)

ในกรณีที่มี 2 กลุ่ม จะได้ฟังก์ชันจำแนก 1 ฟังก์ชัน ซึ่งเขียนเป็นแผนภาพแสดงความสัมพันธ์ระหว่างฟังก์ชันจำแนกและกลุ่มได้ดังนี้



ภาพ 4 แผนภาพความสัมพันธ์ระหว่างฟังก์ชันจำแนกและกลุ่มในกรณีที่ตัวแปรตามมี 2 ค่า

ในกรณีที่มี 3 กลุ่ม จะได้ฟังก์ชันจำแนก 2 ฟังก์ชัน ซึ่งเขียนเป็นแผนภาพแสดงความสัมพันธ์ระหว่างตัวแปรจำแนกและกลุ่มได้ดังนี้



ภาพ 5 แผนภาพความสัมพันธ์ระหว่างฟังก์ชันจำแนกและกลุ่มในกรณีที่ตัวแปรตามมี 3 ค่า

หลักการที่ใช้วิเคราะห์หาความแตกต่างระหว่างกลุ่ม ทำได้โดยนำความแปรปรวนของตัวแปรตามมาแบ่งเป็นส่วน ๆ แล้วนำมาเปรียบเทียบกัน การวิเคราะห์จำแนกและการวิเคราะห์ความแปรปรวนทางเดียวใช้หลักการเดียวกัน คือ แบ่งความแปรปรวนในตัวแปรตามออกเป็นความแปรปรวนระหว่างกลุ่มกับความแปรปรวนภายในกลุ่ม และทำการเปรียบเทียบความแปรปรวนระหว่างกลุ่มกับความแปรปรวนภายในกลุ่ม

การวิเคราะห์ความแปรปรวนมีข้อตกลงเบื้องต้นเรื่อง ตัวแปรต้นต้องมีระดับการวัดแบบนามบัญญัติ ส่วนตัวแปรตามมีระดับการวัดแบบอันตรภาค ดังนั้นถ้าความแปรปรวนระหว่างกลุ่มมีมากกว่าความแปรปรวนภายในกลุ่ม ความแปรปรวนที่เกิดขึ้นจึงเป็นความแปรปรวนของตัวแปรตามแยกตามกลุ่มตัวแปรต้น ส่วนการวิเคราะห์จำแนกมีข้อตกลงเบื้องต้นเรื่อง ตัวแปรต้นต้องมีระดับการวัดแบบอันตรภาค และตัวแปรตามมีระดับการวัดแบบนามบัญญัติ จึงมีลักษณะเป็นส่วนกลับของการวิเคราะห์ความแปรปรวน ความแปรปรวนที่ได้จึงเป็นความแปรปรวนของตัวแปรต้นแยกตามกลุ่มตัวแปรตาม

นอกจากต้องการศึกษาว่ามีความแตกต่างระหว่างกลุ่มหรือไม่ ในการวิเคราะห์จำแนกยังต้องการทราบว่าตัวแปรจำแนกใดมีความสำคัญทำให้เกิดความแตกต่างระหว่างกลุ่มมากที่สุด การวิเคราะห์จำแนกจึงสร้างรูปแบบความสัมพันธ์ระหว่างตัวแปรต้นและกลุ่มคล้ายกับสมการถดถอยเชิงเส้นในการวิเคราะห์ถดถอยพหุคูณ (สราวุฒิ ตูวิพันธ์, 2529 ; สุชาติ ประสิทธิ์รัฐสินธุ์, 2537)

2.1.2 หลักการวิเคราะห์

การวิเคราะห์จำแนกมีขั้นตอนที่สำคัญ 2 ขั้นตอนคือ การสร้างฟังก์ชันจำแนก และการจำแนกหน่วยวิเคราะห์เข้ากลุ่ม (Klecka, 1987) ดังรายละเอียดต่อไปนี้

1.การสร้างฟังก์ชันจำแนก เป็นการหารูปแบบความสัมพันธ์ระหว่างตัวแปรจำแนกและกลุ่มที่ทำให้แต่ละกลุ่มมีความแตกต่างกันมากที่สุด ฟังก์ชันที่ได้ในขั้นตอนนี้ คือ ฟังก์ชันจำแนกคาโนนิคัล (canonical discriminant function) ใช้ในการศึกษาลักษณะความแตกต่างระหว่างกลุ่มว่า กลุ่มตัวแปรจำแนกมีความสามารถในการแยกความแตกต่างระหว่างกลุ่มมากน้อยเพียงใด และตัวแปรจำแนกตัวใดมีความสำคัญที่ทำให้กลุ่มมีความแตกต่างกัน

ฟังก์ชันจำแนกคาโนนิคัล เป็นฟังก์ชันที่เกิดจากการรวมเชิงเส้น (linear combination) ของตัวแปรจำแนก สามารถแสดงได้ตามสมการ

$$f_{ikm} = u_0 + u_1 x_{1km} + u_2 x_{2km} + \dots + u_p x_{pkm}$$

เมื่อ

f_{ikm} เป็นคะแนนบนฟังก์ชันจำแนกคาโนนิกัลของหน่วยวิเคราะห์ m ในกลุ่ม k

x_{ikm} เป็นค่าของตัวแปรจำแนก i ของหน่วยวิเคราะห์ m ในกลุ่ม k

u_i เป็นสัมประสิทธิ์ที่แสดงคุณลักษณะที่ต้องการในฟังก์ชัน

โดยที่ u_i คำนวณได้จาก

$$u_i = v_i \sqrt{n-g}$$

และ
$$u_0 = - \sum_{i=1}^p u_i x_{iL}$$

ค่าสัมประสิทธิ์มาตรฐาน (standardized coefficients)

$$c_i = u_i \sqrt{w_{ii} / n - g}$$

เมื่อ

w_{ii} เป็นผลรวมกำลังสองของตัวแปรจำแนก i

สมการที่ได้จากการวิเคราะห์มี 2 ลักษณะ คือ สมการคะแนนดิบ และสมการมาตรฐาน สมการคะแนนดิบนั้นเหมาะในการใช้ประมาณค่าของหน่วยวิเคราะห์ ส่วนสมการมาตรฐานเหมาะสำหรับใช้เปรียบเทียบอิทธิพลของตัวแปร เพราะเป็นสมการที่ได้ปรับค่าของตัวแปรจำแนกของแต่ละหน่วยวิเคราะห์ให้เป็นคะแนนมาตรฐาน

นอกจากนี้เรายังสามารถคำนวณหาค่าสัมประสิทธิ์โครงสร้าง (structure coefficients) ซึ่งเป็นค่าที่แสดงถึงความสัมพันธ์ระหว่างคะแนนจำแนก (discriminant score) กับตัวแปรต้นแต่ละตัว เมื่อยกกำลังสองของค่าสัมประสิทธิ์โครงสร้าง จะหมายถึงสัดส่วนของความแปรปรวนในตัวแปรต้นที่สามารถอธิบายโดยฟังก์ชันจำแนก มีสูตรในการคำนวณดังนี้

$$S'_{ij} = \sum_{k=1}^p r_{ik} c_{ik} = \sum_{k=1}^p \frac{w_{ik} c_{kj}}{\sqrt{w_{ik} c_{kj}}}$$

เมื่อ S'_{ij} = ค่าสัมประสิทธิ์โครงสร้างภายในกลุ่มสำหรับตัวแปร i กับฟังก์ชันจำแนก j

r_{ik} = ค่าสัมประสิทธิ์สหสัมพันธ์ภายในกลุ่มระหว่างตัวแปร i ถึง k

c_{jk} = ค่าสัมประสิทธิ์มาตรฐานของตัวแปร k ในฟังก์ชันจำแนก j

Pedhazur (1982 : 704) ให้ข้อเสนอแนะว่า ค่าสัมประสิทธิ์โครงสร้างที่ถือว่ามีความสำคัญควรมีค่าตั้งแต่ .30 ขึ้นไป

ในการวิเคราะห์จำแนกนั้นจะต้องมีฟังก์ชันจำแนกน้อยกว่าจำนวนกลุ่มย่อยของตัวแปรตามอยู่ 1 เสมอ แต่ถ้าจำนวนตัวแปรจำแนกมีน้อยกว่าจำนวนกลุ่ม ฟังก์ชันจำแนกจะมีได้เท่ากับจำนวนตัวแปรจำแนก โดยที่ฟังก์ชันแรกจะต้องมีอัตราส่วนของผลรวมกำลังสองระหว่างกลุ่มต่อภายในกลุ่มสูงสุด และฟังก์ชันที่ 2 ต้องไม่สัมพันธ์กับฟังก์ชันแรก จะมีอัตราส่วนดังกล่าวรองลงมา ฟังก์ชันที่ 3 จะต้องไม่สัมพันธ์กับฟังก์ชันที่ 1 และฟังก์ชันที่ 2 และมีอัตราส่วนของผลรวมกำลังสองระหว่างกลุ่มต่อภายในกลุ่มรองลงมาอีก จะเป็นลักษณะเช่นนี้จนถึงฟังก์ชันสุดท้าย ซึ่งทำให้มีอัตราส่วนของผลรวมกำลังสองระหว่างกลุ่มต่อภายในกลุ่มสูงสุด แต่ในทางปฏิบัติไม่จำเป็นที่จะต้องใช้ฟังก์ชันจำแนกทุกฟังก์ชันที่เป็นไปได้ ถ้าฟังก์ชันที่ได้ครั้งแรกมีความสามารถในการอธิบายความแตกต่างระหว่างกลุ่มได้ดี

2 การจำแนกหน่วยวิเคราะห์เข้ากลุ่ม เมื่อได้รูปแบบความสัมพันธ์ระหว่างตัวแปรจำแนกที่ทำให้แต่ละกลุ่มมีความแตกต่างกันมากที่สุดแล้ว เราสามารถนำรูปแบบดังกล่าวมาพยากรณ์ความเป็นสมาชิกของกลุ่มของหน่วยวิเคราะห์ และจำแนกหน่วยวิเคราะห์เข้ากลุ่มได้ ซึ่งการจำแนกกลุ่มสามารถทำได้ 3 วิธี (Klecka, 1987) คือ

2.1) ฟังก์ชันการจำแนกกลุ่ม

ฟังก์ชันที่ใช้ในการจำแนกประเภทมีชื่อเรียกภาษาอังกฤษว่า discriminant function หรือ Fisher's Linear Discriminant Function หรือเรียกตามลักษณะใช้งานว่า classification function ฟังก์ชันจำแนกเป็นฟังก์ชันที่ได้จากการรวมเชิงเส้นของตัวแปรจำแนก ซึ่งทำให้ความแตกต่างระหว่างกลุ่มมีค่าสูงสุดและความแตกต่างภายในกลุ่มมีค่าต่ำสุด สามารถเขียนในรูปสมการเชิงเส้นของแต่ละกลุ่มย่อยได้เป็น

$$h_k = b_{k0} + b_{k1}x_1 + b_{k2}x_2 + \dots + b_{kp}x_p$$

เมื่อ

h_k เป็นคะแนนของกลุ่ม k

$b_{k0}, b_{k1}, b_{k2}, \dots, b_{kp}$ เป็นค่าสัมประสิทธิ์การจำแนกกลุ่ม

$$D^2(x|G_k) = (n - g) \sum_{i=1}^g \sum_{j=1}^g a_{ij} (x_i - x_{ik})(x_j - x_{jk})$$

เมื่อ

$D^2(x|G_k)$ = ระยะห่างกำลังสองจากจุด x (หน่วยวิเคราะห์แต่ละหน่วย) ถึงศูนย์กลางของกลุ่ม k

n = จำนวนหน่วยวิเคราะห์ทั้งหมด

g = จำนวนกลุ่ม

หลังจากที่คำนวณ D^2 ของหน่วยวิเคราะห์แต่ละหน่วยจนครบทุกกลุ่มแล้ว เราสามารถจำแนกหน่วยวิเคราะห์เข้ากลุ่มด้วย D^2 ที่น้อยที่สุด และกลุ่มนั้นก็จะเป็นกลุ่มที่มีคุณลักษณะที่ใกล้เคียงคล้ายคลึงกับหน่วยวิเคราะห์ (Klecka, 1987)

2.3) ความน่าจะเป็นของสมาชิกภาพกลุ่ม (probability of group membership)

เป็นการพิจารณาโอกาสหรือความน่าจะเป็นที่หน่วยวิเคราะห์แต่ละหน่วยจะเป็นสมาชิกในกลุ่มใดกลุ่มหนึ่ง โดยอาศัยกฎของเบย์ส์ (Bayes' Rule)

$$P(G_k|x) = \frac{P(x|G_k)}{\sum_{i=1}^g P(x|G_i)}$$

ค่าความน่าจะเป็นที่ได้นี้เรียกว่า ความน่าจะเป็นภายหลัง (posterior probability) โดยที่ค่าความน่าจะเป็นทุกกลุ่มรวมกันแล้วจะมีค่าเท่ากับ 1 เมื่อหาความน่าจะเป็นภายหลังครบทุกกลุ่มแล้ว หน่วยวิเคราะห์แต่ละหน่วยจะถูกจำแนกเข้ากลุ่มที่มีค่าความน่าจะเป็นภายหลังใหญ่ที่สุด (สุชาติ ประสิทธิ์รัฐสินธุ์ และกรรณิการ์ สุรเกษม, 2533 ; Klecka, 1987)

2.2 ข้อตกลงเบื้องต้น

ในการวิเคราะห์จำแนกมีข้อตกลงเบื้องต้นที่สำคัญ 7 ประการด้วยกัน คือ ประการที่หนึ่ง ตัวแปรตามจะต้องมีระดับการวัดแบบนามบัญญัติ ซึ่งอาจมี 2 ค่า หรือมากกว่า 2 ค่าก็ได้ ประการที่สอง แต่ละกลุ่มย่อยของตัวแปรตามนั้นจะต้องมีหน่วยวิเคราะห์ไม่ต่ำกว่า 2 หน่วย ประการที่สาม จำนวนตัวแปรจำแนก (ตัวแปรต้น) ต้องมีอย่างน้อยที่สุด 1 ตัว และมากที่สุดจะต้องน้อยกว่าจำนวนหน่วยวิเคราะห์ทั้งหมดลบ 2 ($0 < p < (n-2)$) ประการที่สี่ ตัวแปรจำแนกจะต้องมีระดับการวัดแบบอันตรภาค ประการที่ห้า ตัวแปรจำแนกจะต้องไม่มีภาวะ

ร่วมเส้นตรงพหุ ประการที่หก เมทริกซ์ความแปรปรวนร่วมในแต่ละกลุ่มจะต้องเท่ากัน ประการสุดท้าย ตัวแปรจำแนกจะต้องมีการแจกแจงปกติพหุ (Klecka, 1987)

2.3 การประเมินฟังก์ชันจำแนก

การประเมินฟังก์ชันจำแนกที่ได้ว่ามีความเหมาะสมใช้ได้ดีหรือไม่ สามารถพิจารณาได้จากค่าสถิติสำคัญ (สุชาติ ประสิทธิ์รัฐสินธุ์ และ กรรณิการ์ สุขเกษม, 2533 ; Klecka, 1987) ดังนี้

1. ค่าไอเกน (eigenvalue) เป็นอัตราส่วนความแปรปรวนระหว่างกลุ่มกับความแปรปรวนภายในกลุ่ม หรือค่ารวมของความแปรปรวนของตัวแปรมาตรฐาน ซึ่งแต่ละตัวจะมีความแปรปรวนเท่ากับ 1 ถ้าค่าดังกล่าวสูงแสดงว่าฟังก์ชันจำแนกมีความเหมาะสมใช้ได้ดี ค่าไอเกนสามารถคำนวณจากสูตร

$$\sum b_{pi} v_i = \lambda \sum w_{pi} v_i$$

เมื่อ λ = ค่าไอเกน , V = ชุดของค่าสัมประสิทธิ์ p , b และ w = ผลรวมของค่ากำลังสองและผลคูณไขว้ของคะแนนเบี่ยงเบน (sum of squares and crossproducts of deviations) ระหว่างกลุ่มกับภายในกลุ่ม ตามลำดับ

2. ค่าสหสัมพันธ์คาโนนิกัล (canonical correlation) เป็น ค่ารากที่สองของความแปรปรวนระหว่างกลุ่มต่อความแปรปรวนทั้งหมด ถ้านำค่าสหสัมพันธ์คาโนนิกัลมายกกำลังสองก็จะเป็นสัดส่วนความแปรปรวนของฟังก์ชันที่อธิบายได้โดยกลุ่ม ซึ่งจะชี้ให้เห็นว่าการเป็นสมาชิกของกลุ่มมีความสัมพันธ์กับฟังก์ชันที่หามาได้มากน้อยเพียงใด ถ้าค่าสหสัมพันธ์คาโนนิกัลสูงแสดงว่าฟังก์ชันนั้นสามารถใช้ในการจำแนกได้ดี ค่าสหสัมพันธ์คาโนนิกัลสามารถคำนวณได้จากสูตร

$$r^*_i = \sqrt{\frac{\lambda_i}{1 + \lambda_i}}$$

เมื่อ i = ฟังก์ชันจำแนกที่ i

3. ค่า Wilks' lambda เป็นสัดส่วนของความแปรปรวนภายในกลุ่มต่อความแปรปรวนทั้งหมด หรือ สัดส่วนความแปรปรวนทั้งหมดที่ไม่ได้อธิบายด้วยความแตกต่างระหว่างกลุ่ม ถ้าค่า Wilks' lambda มากเท่าใดตัวแปรที่เหลือจะอธิบายความแปรปรวนการเป็นสมาชิกของกลุ่มโดยฟังก์ชันใหม่จะน้อยลงเท่านั้น ค่า Wilks' lambda สามารถคำนวณได้จากสูตร

$$\Lambda = \prod_{i=k+1}^q \frac{1}{1 + \lambda_i}$$

เมื่อ k = จำนวนฟังก์ชันจำแนกที่ได้

นอกจากนี้การประเมินความใช้ได้ดีของฟังก์ชันจำแนก สามารถพิจารณาจากร้อยละจำนวนหน่วยวิเคราะห์ที่ฟังก์ชันสามารถจำแนกได้ถูกต้อง

2.4 การวิเคราะห์จำแนกด้วยโปรแกรมคอมพิวเตอร์

ในปัจจุบันการวิเคราะห์ข้อมูลด้วยการวิเคราะห์จำแนกสามารถทำได้โดยตรงด้วยการใช้โปรแกรมสำเร็จรูปในคอมพิวเตอร์ เช่น SPSS SAS เป็นต้น ซึ่งมีความสะดวกและสามารถคำนวณได้อย่างรวดเร็ว ในโปรแกรมคอมพิวเตอร์สามารถสร้างฟังก์ชันจำแนกได้ 2 วิธี คือ วิธีโดยตรง (direct method) และวิธีวิเคราะห์แบบมีลำดับขั้น (stepwise method) (สุชาติ ประสิทธิ์รัฐสินธุ์, 2537 ; Kiecka, 1987)

1. วิธีโดยตรง เป็นวิธีการวิเคราะห์ตัวแปรทุกตัว โดยไม่พิจารณาผลของตัวแปรจำแนกที่มีต่อฟังก์ชัน

2. วิธีวิเคราะห์แบบมีลำดับขั้น เป็นวิธีการพิจารณาเลือกตัวแปรเข้าสมการทีละตัว และพิจารณาผลของตัวแปรที่นำเข้าสมการแต่ละขั้นตอนด้วย ซึ่งมีวิธีการคัดเลือกหลายวิธี ได้แก่

2.1 พิจารณาค่า Wilks' lambda โดยเลือกตัวแปรที่ทำให้ค่า Wilks' lambda มีค่าน้อยที่สุด

2.2 พิจารณาค่า Rao' V หรือ Lowley - Hotelling Trace ซึ่งเป็นสถิติที่แสดงถึงความแตกต่างระหว่างกลุ่ม โดยถ้ากลุ่มมีความความแตกต่างกันมาก ค่า Rao' V จะสูง ในการเลือกตัวแปรด้วยวิธีนี้ต้องเลือกตัวแปรที่ทำให้ ค่า Rao' V สูงขึ้น

2.3 พิจารณาค่า Mahalanobis Distance (D^2) การเลือกตัวแปรโดยวิธีนี้ จะเลือกตัวแปรที่มีค่า D^2 มากที่สุดของคู่ที่ใกล้กันมากที่สุด หรือมีค่า D^2 น้อยที่สุดตั้งแต่แรก เข้าสมการก่อน

2.4 พิจารณาค่าความแปรปรวนเหลือที่น้อยที่สุด (minimizing residual variance) ตัวแปรที่ทำให้ความแปรปรวนที่เหลือไม่ได้ทั้งหมดเหลือที่น้อยที่สุด จะได้รับเลือกเข้าสมการ

2.5 พิจารณาค่า Between - Group F เป็นค่าสถิติที่แสดงความแตกต่างระหว่างกลุ่มในการเลือกตัวแปรด้วยวิธีนี้ จะเลือกตัวแปรที่มีค่า F ที่ใหญ่ที่สุด

แต่อย่างไรก็ตาม การใช้เทคนิควิธีวิเคราะห์จำแนกก็มีข้อที่ควรพิจารณาให้รอบคอบหลายประการ คือ ประการที่หนึ่ง ตัวแปรที่มีการแจกแจงแบบปกติจะช่วยให้การวิเคราะห์ดีขึ้นเพราะการอ้างอิงนัยสำคัญทางสถิติของตัวแปรอาศัยสมมติฐานการแจกแจงปกติของตัวแปร ถ้าเกิดปัญหาตัวแปรมีการแจกแจงที่ไม่เป็นปกติสามารถแก้ไขด้วยการเพิ่มจำนวนหน่วยวิเคราะห์ให้มากขึ้น ประการที่สอง การเก็บข้อมูลไม่ควรมีข้อมูลที่ขาดหายหรือใช้ไม่ได้ (missing data) เพราะจะเกิดความลำเอียงให้หน่วยวิเคราะห์ที่มีข้อมูลครบ ประการที่สาม การมีค่ารอบนอกหรือค่าปลายสุด (outliers) จะทำให้ค่าสถิติต่าง ๆ ผิดพลาดจากที่ควรจะเป็น ประการที่สี่ ตัวแปรต้นไม่ควรมีความสัมพันธ์กันสูง เพราะจะทำให้เกิดภาวะร่วมเส้นตรงพหุได้ ประการที่ห้า ตัวแปรควรมีความสัมพันธ์เชิงเส้น (linearity) เพราะการวิเคราะห์ค่าในนิคัลต้องใช้ค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สันซึ่งไวต่อความสัมพันธ์เชิงเส้นมากกว่าความสัมพันธ์แบบอื่น ผู้วิจัยสามารถตรวจสอบความสัมพันธ์เชิงเส้นได้โดยการลงจุดกระจาย (สุชาติ ประสิทธิ์รัฐสินธุ์, 2537)

ลักษณะโดยสรุปของการวิเคราะห์ถดถอยโลจิสติกและการวิเคราะห์จำแนก สามารถนำเสนอได้ดังตาราง 3

ตาราง 3 ลักษณะโดยสรุปของการวิเคราะห์ถดถอยโลจิสติกและการวิเคราะห์จำแนก

การวิเคราะห์ถดถอยโลจิสติก	การวิเคราะห์จำแนก
<p>หลักการวิเคราะห์</p> <p>สร้างสมการโลจิท เพื่อหารูปแบบความสัมพันธ์ระหว่างแอดัมต่อกับชุดตัวแปรต้น และจำแนกหน่วยวิเคราะห์เข้ากลุ่มตามค่าความน่าจะเป็นที่มีขนาดใหญ่ที่สุด</p> <p>รูปแบบสมการ</p> <p>สมการโลจิทได้จากการรวมเชิงเส้นของชุดตัวแปรต้น สามารถเขียนได้ดังนี้</p> $\ln \left \frac{\pi(x_1, x_2, \dots, x_p)}{1 - \pi(x_1, x_2, \dots, x_p)} \right = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p$ <p>สมการที่ได้มีลักษณะคล้ายกับสมการถดถอยพหุคูณ</p> <p>การประมาณค่าพารามิเตอร์</p> <p>ใช้วิธี ความเป็นไปได้สูงสุด (maximum likelihood = ML) ซึ่งเป็นการคำนวณทวนซ้ำ เพื่อให้ได้ค่าพารามิเตอร์ที่ใกล้เคียงกับข้อมูลเชิงประจักษ์มากที่สุด</p>	<p>หลักการวิเคราะห์</p> <p>สร้างฟังก์ชันจำแนกคาโนนิกัล เพื่อหารูปแบบความสัมพันธ์ระหว่างคะแนนจำแนกกับชุดตัวแปรต้น (ตัวแปรจำแนก) และจำแนกหน่วยวิเคราะห์เข้ากลุ่มได้ 3 วิธี ดังนี้</p> <ol style="list-style-type: none"> ใช้ฟังก์ชันจำแนก สร้างคะแนนจำแนกของแต่ละหน่วยวิเคราะห์ แล้วจัดหน่วยวิเคราะห์เข้ากลุ่มที่มีคะแนนสูงสุด ใช้ฟังก์ชันระยะห่างทั่วไป โดยจัดหน่วยวิเคราะห์เข้ากลุ่มที่มีระยะห่างระหว่างหน่วยวิเคราะห์กับศูนย์กลางกลุ่มน้อยที่สุด ใช้ความน่าจะเป็นสมาชิกภาพกลุ่มตามกฎของเบย์ โดยจัดหน่วยวิเคราะห์เข้ากลุ่มที่มีความน่าจะเป็นภายหลังใหญ่ที่สุด <p>รูปแบบสมการ</p> <p>ฟังก์ชันจำแนกคาโนนิกัล ได้จากการรวมเชิงเส้นของชุดตัวแปรจำแนก ที่มุ่งทำให้ความแตกต่างระหว่างกลุ่มมีค่าสูงสุด และความแตกต่างภายในกลุ่มมีค่าน้อยที่สุด รูปแบบฟังก์ชันเขียนได้ดังนี้</p> $f_{km} = u_0 + u_1 x_{1km} + u_2 x_{2km} + \dots + u_p x_{pkm}$ <p>ฟังก์ชันจำแนกนี้มีลักษณะคล้ายกับสมการถดถอยพหุคูณ</p> <p>การประมาณค่าพารามิเตอร์</p> <p>ใช้วิธีกำลังสองน้อยที่สุด (ordinary least squares = OLS) ที่ทำให้ค่าผลรวมกำลังสองของความคลาดเคลื่อนมีค่าน้อยที่สุด</p>

ตาราง 3 (ต่อ)

<p>ลักษณะข้อมูลที่ใช้ในการวิเคราะห์</p> <p>1. ตัวแปรตาม เป็นตัวแปรที่มีตั้งแต่ 2 ค่าขึ้นไป</p> <p>2. ตัวแปรต้น เป็นตัวแปรต่อเนื่อง หรือตัวแปรคัมมี</p> <p>ข้อตกลงเบื้องต้น</p> <p>1. ตัวแปรต้นไม่จำเป็นต้องมีการแจกแจงแบบปกติ</p> <p>2. ความแปรปรวนร่วมในแต่ละกลุ่มตัวอย่างไม่จำเป็นจะต้องเท่ากัน</p>	<p>ลักษณะข้อมูลที่ใช้ในการวิเคราะห์</p> <p>1. ตัวแปรตาม เป็นตัวแปรกลุ่มที่มีตั้งแต่ 2 กลุ่มขึ้นไป</p> <p>2. ตัวแปรต้น เป็นตัวแปรต่อเนื่อง หรือตัวแปรคัมมี</p> <p>ข้อตกลงเบื้องต้น</p> <p>1. ตัวแปรต้นต้องมีการแจกแจงแบบปกติ</p> <p>2. ความแปรปรวนร่วมในแต่ละกลุ่มตัวอย่างต้องเท่ากัน</p>
<p>ค่าสถิติทดสอบที่ใช้ในการวิเคราะห์</p> <p>1. การทดสอบนัยสำคัญค่าสัมประสิทธิ์ถดถอยโลจิสติกแต่ละตัว ใช้ค่า Wald statistics</p> <p>2. การตรวจสอบความสอดคล้องของโมเดลพิจารณาจาก ค่า $-2 LL$ likelihood ค่า Model Chi-Square ที่ทดสอบว่าสัมประสิทธิ์ของตัวแปรในโมเดลยกเว้นค่าคงที่มีค่าเป็นศูนย์ และค่า R^2_L ที่แสดงสัดส่วนความแปรปรวนที่อธิบายโดยโมเดล</p>	<p>ค่าสถิติทดสอบที่ใช้ในการวิเคราะห์</p> <p>1. การเปรียบเทียบค่าเฉลี่ยกลุ่มของแต่ละกลุ่มย่อยของตัวแปรจำแนก (หรือการทดสอบนัยสำคัญค่าสัมประสิทธิ์การจำแนก) ใช้ค่าสถิติทดสอบเอฟ</p> <p>2. การตรวจสอบความเหมาะสมของฟังก์ชันจำแนกที่ได้ พิจารณาจากค่าไอเกน (eigenvalue) ค่าสหสัมพันธ์คาโนนิกัล และ ค่า Wilks' lambda</p>

ตอนที่สี่ งานวิจัยที่เปรียบเทียบผลการวิเคราะห์ข้อมูลระหว่างการวิเคราะห์ถดถอยโลจิสติกกับการวิเคราะห์จำแนก

Efron (1975) ทำการเปรียบเทียบประสิทธิภาพการวิเคราะห์ข้อมูลระหว่างการวิเคราะห์ถดถอยโลจิสติกที่มีตัวแปรตามเป็นตัวแปรทวิภาคและการวิเคราะห์จำแนกโดยใช้การจำลองสถานการณ์ (simulation) ผลการวิจัยพบว่า เมื่อตัวแปรต้นมีการแจกแจงแบบปกติด้วย

เมทริกซ์ความแปรปรวนร่วมที่เท่ากันในกลุ่มตัวอย่าง การวิเคราะห์ถดถอยโลจิสติกจะมีประสิทธิภาพในการประมาณค่าน้อยกว่าการวิเคราะห์จำแนก

ต่อมา Press และ Wilson (1978) ได้สรุปประเด็นสำคัญเกี่ยวกับการวิเคราะห์ข้อมูล ระหว่างการวิเคราะห์ถดถอยโลจิสติกกับการวิเคราะห์จำแนกว่า การวิเคราะห์ถดถอยโลจิสติกมีประสิทธิภาพดีกว่าในสถานการณ์ต่าง ๆ ดังนี้

ก. เมื่อตัวแปรต้นมีการละเมิดข้อตกลงเบื้องต้นว่า ต้องมีแจกแจงแบบปกติพหุด้วย เมทริกซ์ความแปรปรวนร่วมที่เท่ากัน การประมาณค่าสัมประสิทธิ์ความชันด้วยฟังก์ชันจำแนก จะไม่คงที่ แม้ว่าจะมีกลุ่มตัวอย่างขนาดใหญ่ก็ตาม นั่นคือ ถ้าตัวแปรต้นเป็นตัวแปรทวิภาค ก็ไม่อาจคาดหวังได้ว่าฟังก์ชันจำแนกจะสามารถทำนายความน่าจะเป็นได้ถูกต้อง ในกรณีที่ตัวแปรต้นอย่างน้อยตัวหนึ่งเป็นตัวแปรทวิภาค ควรใช้วิธีการประมาณค่าที่มีความเที่ยงตรง เช่น วิธีความเป็นไปได้สูงสุด (maximum likelihood) ซึ่งเป็นวิธีที่ใช้ในการวิเคราะห์ถดถอยโลจิสติก

ข. เมื่อกลุ่มตัวอย่างมีขนาดใหญ่ แต่มีการละเมิดข้อตกลงเบื้องต้นที่ว่า ตัวแปรต้นต้องมีการแจกแจงแบบปกติพหุ ถ้าใช้วิธีความเป็นไปได้สูงสุดประมาณค่าสัมประสิทธิ์ความชันที่เป็นศูนย์จริง ๆ ค่าสัมประสิทธิ์ความชันที่ได้ก็มีแนวโน้มที่จะเป็นศูนย์ แต่ถ้าใช้ฟังก์ชันจำแนกประมาณค่า ค่าสัมประสิทธิ์ความชันที่ได้ก็มีแนวโน้มที่จะคลาดเคลื่อน

ค. ในโมเดลถดถอยโลจิสติกภายใต้เงื่อนไขที่ไม่ปกติ การประมาณค่าด้วยวิธีความเป็นไปได้สูงสุด จะให้โมเดลที่มีความเหมาะสมมากกว่าฟังก์ชันจำแนก

ง. การประมาณค่าด้วยวิธีความเป็นไปได้สูงสุดให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (mean squared error) น้อยกว่าฟังก์ชันจำแนก เนื่องจากวิธีความเป็นไปได้สูงสุดเป็นฟังก์ชันที่ให้ค่าสถิติเพียงพอ แต่ในฟังก์ชันจำแนกไม่มี

จ. วิธีความเป็นไปได้สูงสุดในการวิเคราะห์ถดถอยโลจิสติกจะบังคับให้ค่าที่คาดหวังเท่ากับค่าที่สังเกตได้ แต่ฟังก์ชันจำแนกจะไม่เป็นเช่นนั้น บางครั้งจะพบว่าค่าที่ได้จากการประมาณมากกว่าค่าที่เป็นจริง

นอกจากนี้ Press และ Wilson ได้เสนอผลการวิจัยที่ใช้กับข้อมูลเชิงประจักษ์ เมื่อมีตัวแปรตามเป็นตัวแปรทวิภาค 2 กรณี คือ

กรณีที่ 1 เป็นการศึกษาโดยใช้ฐานข้อมูลของ Wilson เมื่อ ค.ศ. 1977 ศึกษาผู้ป่วยเป็นโรคมะเร็งเต้านม ลักษณะข้อมูลที่ใช้ในส่วนของตัวแปรต้น มีทั้งตัวแปรต่อเนื่องและตัวแปรไม่ต่อเนื่อง แต่ส่วนใหญ่เป็นตัวแปรทวิภาค ผลการวิเคราะห์พบว่า การวิเคราะห์ถดถอย

โลจิสติกสามารถจำแนกผู้ป่วยเข้ากลุ่มได้ถูกต้องมากกว่าการวิเคราะห์จำแนก และใช้เวลาในการวิเคราะห์นานกว่า

กรณีที่ 2 Press และ Wilson (1978) ศึกษาการเปลี่ยนแปลงของประชากรใน 50 รัฐของสหรัฐอเมริกา ลักษณะข้อมูลที่ใช้ในการวิเคราะห์ ตัวแปรต้นมีทั้งตัวแปรต่อเนื่องและตัวแปรทวิภาค แต่ไม่มีตัวแปรต่อเนื่องตัวใดมีการแจกแจงแบบปกติ ผลการวิเคราะห์พบว่า การวิเคราะห์ถดถอยโลจิสติกสามารถจำแนกกลุ่มได้ถูกต้องมากกว่าการวิเคราะห์จำแนก แต่ใช้เวลาในการวิเคราะห์นานกว่า

ผลการวิจัยที่ได้จากข้อมูลเชิงประจักษ์ของ Press และ Wilson ที่กล่าวไว้ในตอนต้นสอดคล้องกับผลสรุปงานวิจัยของ Halparin, Blackwelder และ Verter (1971) ซึ่งถึงใน พัทรีเพ็งประโคน, 2539) ที่ทำการเปรียบเทียบประสิทธิภาพการวิเคราะห์ข้อมูลระหว่างการวิเคราะห์ถดถอยโลจิสติกที่มีตัวแปรตามเป็นตัวแปรทวิภาคและการวิเคราะห์จำแนกในการศึกษากับกลุ่มคนที่เป็นโรคหัวใจโคโรนารี พบว่า เมื่อตัวแปรต้นมีการละเมิดข้อตกลงเบื้องต้นเรื่องการแจกแจงปกติพหุและมีเมทริกซ์ความแปรปรวนร่วมเท่ากันในกลุ่มตัวอย่าง การประมาณค่าด้วยวิธีความเป็นไปได้สูงสุด ในการวิเคราะห์ถดถอยโลจิสติกจะให้โมเดลที่มีความเหมาะสมมากกว่าฟังก์ชันจำแนก แต่จะใช้เวลาในการวิเคราะห์นานกว่าด้วย

งานวิจัยของ ฮัมรินทร์ ทักษิณเสถียร (2536) ให้ผลการวิจัยที่คล้ายคลึงกับงานวิจัยของ Halparin, Blackwelder และ Verter (1971) และ Press และ Wilson (1978) โดยทำการศึกษาปัจจัยที่มีความสัมพันธ์กับภาวะความดันโลหิตสูงเนื่องจากการตั้งครรภ์ในโรงพยาบาลรามาธิบดี และเปรียบเทียบผลการวิเคราะห์จากการวิเคราะห์ถดถอยโลจิสติกกับการวิเคราะห์จำแนกเมื่อตัวแปรตามเป็นตัวแปรทวิภาค ผลการวิจัยพบว่า ตัวแปรต้นมีการละเมิดข้อตกลงเบื้องต้นในเรื่องเมทริกซ์ความแปรปรวนเท่ากันในกลุ่มตัวอย่าง สมการที่ได้จากการวิเคราะห์ถดถอยโลจิสติกมีความเหมาะสมกับข้อมูล สามารถอธิบายความแปรปรวนได้ดีกว่าการวิเคราะห์จำแนก แต่ทั้งสองวิธีสามารถจัดกลุ่มได้ถูกต้องเท่ากัน ในแง่ของการแปลผลวิธีการวิเคราะห์ถดถอยโลจิสติกจะดีกว่าในแง่ของความหมายและความเป็นจริงของตัวแปร และมีการผ่อนปรนในเรื่องระดับการวัดของตัวแปรมากกว่าการวิเคราะห์จำแนก

แต่การวิจัยของ พัทรี เพ็งประโคน (2539) ให้ผลการวิจัยต่างจากผลการวิจัยข้างต้น เมื่อศึกษาเปรียบเทียบผลการวิเคราะห์ข้อมูลระหว่างวิธีการวิเคราะห์ถดถอยโลจิสติกที่ตัวแปรตามเป็นตัวแปรทวิภาคกับวิธีการวิเคราะห์จำแนก ในการวิเคราะห์หาตัวแปรจำแนกกระหว่างชนิดที่มีวิถีชีวิตแบบประชาธิปไตยกับชนิดที่มีวิถีชีวิตแบบประชาธิปไตยสูง ผลการวิจัยพบ

ว่า สมการที่ได้ในการวิเคราะห์ถดถอยโลจิสติกอธิบายความแปรปรวนและจำแนกกลุ่มได้ถูกต้องน้อยกว่าการวิเคราะห์จำแนกเล็กน้อย ทั้งที่ตัวแปรต้นมีการละเมิดข้อตกลงเบื้องต้นเรื่องความแปรปรวนเท่ากันในกลุ่มตัวอย่าง

ส่วนการเปรียบเทียบผลการวิเคราะห์ข้อมูลระหว่างวิธีการวิเคราะห์ทั้งสองวิธีเมื่อตัวแปรตามเป็นตัวแปรพหุภาค ตามที่ Wilson และ Hardgrave (1995) ได้ศึกษาเปรียบเทียบสมรรถภาพในการวิเคราะห์ข้อมูลระหว่างเทคนิควิธี 4 วิธีด้วยกัน คือ การวิเคราะห์ถดถอยพหุคูณ การวิเคราะห์ถดถอยโลจิสติก การวิเคราะห์จำแนก และการวิเคราะห์เครือข่ายนิวรัล (neural network) โดยศึกษาระดับความสำเร็จของนักเรียนในโปรแกรม MBA ซึ่งแบ่งเป็น 3 กลุ่ม ผลการวิจัยพบว่า เมื่อตัวแปรต้นมีการละเมิดข้อตกลงเบื้องต้นเรื่องการแจกแจงแบบปกติพหุ การวิเคราะห์ถดถอยพหุคูณให้ผลการจำแนกได้ถูกต้องต่ำสุด การวิเคราะห์เครือข่ายนิวรัลให้ผลการวิเคราะห์ใกล้เคียงการวิเคราะห์จำแนกและการวิเคราะห์ถดถอยโลจิสติก แต่มีความคลุมเครือและแปลผลยาก ดังนั้นการวิเคราะห์จำแนกและการวิเคราะห์ถดถอยโลจิสติกจึงเป็นวิธีที่เหมาะสมกับการทำนายกลุ่ม แต่การวิเคราะห์จำแนกให้ผลการทำนายถูกต้องมากกว่าการวิเคราะห์ถดถอยโลจิสติกเล็กน้อย

สรุป จากงานวิจัยที่ผ่านมาพบว่า ถ้ามีการละเมิดข้อตกลงเบื้องต้นเรื่องตัวแปรต้นต้องมีการแจกแจงแบบปกติพหุและมีเมทริกซ์ความแปรปรวนเท่ากันแล้ว การวิเคราะห์ถดถอยโลจิสติกมักจะให้ผลการวิเคราะห์ดีกว่าการวิเคราะห์จำแนก แต่บางครั้งแม้ว่าจะมีการละเมิดข้อตกลงเบื้องต้นดังกล่าว การวิเคราะห์จำแนกกลับให้ผลการวิเคราะห์ดีกว่าการวิเคราะห์ถดถอยโลจิสติก เช่น งานวิจัยของพัชรี เฟิงประโคน (2539) และ Wilson และ Hardgrave (1995) จึงทำให้ไม่สามารถที่จะสรุปผลการเปรียบเทียบได้อย่างชัดเจน

สมมติฐานการวิจัยข้อที่ 2

จากการเสนอผลการศึกษาเปรียบเทียบประสิทธิภาพการวิเคราะห์ข้อมูลระหว่างการวิเคราะห์ถดถอยโลจิสติกและการวิเคราะห์จำแนกของ Press และ Wilson (1978) พบว่า การวิเคราะห์จำแนกใช้เวลาในการวิเคราะห์น้อยกว่าการวิเคราะห์ถดถอยโลจิสติก และผลการวิจัยของ พัทรี เฟิงประโคน (2539) ที่พบว่า การวิเคราะห์จำแนกสามารถจำแนกกลุ่มได้ถูกต้องมากกว่าการวิเคราะห์ถดถอยโลจิสติกทั้งที่ตัวแปรต้นมีการละเมิดข้อตกลงเบื้องต้นเรื่องความแปรปรวนเท่ากันในกลุ่มตัวอย่าง ซึ่งสอดคล้องกับผลการวิจัยของ Wilson และ Hardgrave

(1995) ที่ว่า เมื่อตัวแปรต้นมีการละเมิดข้อตกลงเบื้องต้นเรื่องการแจกแจงแบบปกติพอ การวิเคราะห์จำแนกให้ผลการทำนายถูกต้องมากกว่าการวิเคราะห์ถดถอยโลจิสติกเล็กน้อย จากผลการวิจัยดังกล่าวทำให้ผู้วิจัยตั้งสมมติฐานข้อที่สองว่า การวิเคราะห์จำแนกน่าจะให้ผลการวิเคราะห์ข้อมูลมีประสิทธิภาพสูงกว่าการวิเคราะห์ถดถอยโลจิสติกพหุวิภาคในด้านการทำนายกลุ่มและเวลาที่ใช้ในการวิเคราะห์



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย