

## บทที่ 3

### วิธีดำเนินการวิจัย

การศึกษาวิจัยครั้งนี้เป็นการศึกษาถึงทัศนคติของครูเกี่ยวกับการเข้ารับบริการหรือไม่เข้ารับบริการในคลินิกสตรีวัยหมดประจำเดือนและศึกษาถึงปัจจัยที่มีผลต่อการเข้ารับบริการหรือไม่เข้ารับบริการในคลินิกสตรีวัยหมดประจำเดือน ในเขตอำเภอเมือง จังหวัดพิษณุโลก โดยมีระเบียบของการดำเนินการวิจัยดังนี้

3.1 ประชากรและกลุ่มตัวอย่าง

3.2 เครื่องมือที่ใช้ในการวิจัย

3.3 การเก็บรวบรวมข้อมูล

3.4 สถิติที่ใช้ในการวิเคราะห์

#### 3.1 ประชากรและกลุ่มตัวอย่าง

การวิจัยครั้งนี้เป็นการศึกษาถึงทัศนคติของครูเกี่ยวกับการเข้ารับบริการหรือไม่เข้ารับบริการในคลินิกสตรีวัยหมดประจำเดือน และปัจจัยที่ส่งผลต่อการเข้ารับบริการในคลินิกสตรีวัยหมดประจำเดือน ในเขตอำเภอเมือง จังหวัดพิษณุโลก มีขอบเขตการศึกษา ดังนี้

##### ขอบเขตประชากร

ศึกษาคูสตรีที่อายุตั้งแต่ 40-59 ปีในสังกัดสำนักงานการประถมศึกษาแห่งชาติ กรมสามัญศึกษา และกรมอาชีวศึกษาในเขตอำเภอเมือง จังหวัดพิษณุโลก มีจำนวน 1,453 คน

ตาราง 3.1 แสดงจำนวนครูสตรีวัยหมดประจำเดือนแยกตามสังกัด

สังกัด	จำนวนครู
โรงเรียนสังกัดการประถมศึกษา	784
โรงเรียนในสังกัดกรมสามัญศึกษา	537
โรงเรียนสังกัดกรมอาชีวศึกษา	132
รวม	1,453

ในการวิจัยครั้งนี้เป็นการศึกษาโดยขยายผลจากกระบวนการวนวิชา การสัมมนา ที่ผู้วิจัยได้นำเสนอเรื่อง สถิติวิเคราะห์ปัจจัยที่มีผลต่อการเข้ารับบริการในคลินิกสตรีวัยหมดประจำเดือน ของข้าราชการครูสังกัดกรมสามัญศึกษา กระทรวงศึกษาธิการ ในเขตเทศบาลเมือง จังหวัดพิษณุโลก โดยที่กำหนดขอบเขตประชากร คือ ครูสตรีวัยหมดประจำเดือนสังกัดกรมสามัญศึกษา กระทรวงศึกษาธิการ ในเขต เทศบาลเมือง จังหวัดพิษณุโลก โดยที่ ขนาดตัวอย่างที่ใช้ในการศึกษาทั้งหมด 197 คน ซึ่งทำการสุ่มตัวอย่าง อย่างมีระบบจากบัญชีรายชื่อข้าราชการ ได้ข้อมูลลักษณะทั่วไปของครูสตรีวัยหมดประจำเดือนที่สำคัญ ดังนี้

ตาราง 3.2 แสดงข้อมูลลักษณะทั่วไปของครูสตรีวัยหมดประจำเดือน

	ข้อมูล	จำนวน (คน)	ร้อยละ
การเข้ารับบริการในคลินิก สตรีวัยหมดประจำเดือน	เคยเข้ารับบริการ	65	33.0
	ไม่เคยเข้ารับบริการ	132	67.0
ลักษณะการมีประจำเดือน	ไม่ปกติ ขาดหายน้อยกว่า 1 ปี	18	9.1
	ขาดหายนานเกินกว่า 1 ปี	55	27.9
	มาสม่ำเสมอ	124	62.9

#### ตัวอย่าง

ในการศึกษาวิจัยครั้งนี้ทำการสุ่มตัวอย่างโดยวิธี Stratified-Systematic Sampling โดยมีขั้นตอนดังนี้

1. กำหนดขนาดตัวอย่างที่จะใช้ในการศึกษานี้ โดยข้อมูลสัดส่วนของครูสตรีวัยหมดประจำเดือนที่เข้ารับบริการในคลินิกซึ่งได้จากตาราง 3.2 เท่ากับ 0.33 และต้องการให้ความคลาดเคลื่อนในการประมาณไม่เกิน 5 % ด้วยระดับความเชื่อมั่น 95 % สามารถคำนวณขนาดตัวอย่างได้ดังนี้

$$n = \frac{Z_{\alpha/2}^2 pq}{d^2} = \frac{(1.96)^2 (0.33)(0.67)}{0.05^2} = 339.7 \approx 340 \text{ คน}$$

ได้ขนาดตัวอย่างอย่างน้อย 340 คน

2. ทำการสุ่มตัวอย่างแบบชั้นภูมิโดยกำหนดให้สังกัดเป็นชั้นภูมิและจากแต่ละชั้นภูมิสุ่มตัวอย่างตามสัดส่วนของประชากรดังนี้

$$n_h = \frac{N_h \times n}{N}$$

กำหนด  $N$  คือ ขนาดประชากร,  $n$  คือ ขนาดตัวอย่างทั้งหมด

$N_h$  คือ ขนาดประชากรในแต่ละชั้นภูมิ และ  $n_h$  คือ ขนาดตัวอย่างแต่ละชั้นภูมิ

ตาราง 3.3 แสดงขนาดตัวอย่างครูสตรีวัยหมดประจำเดือนในเขตอำเภอเมืองจังหวัดพิษณุโลก แยกตามสังกัดที่ได้จากการคำนวณ

สังกัด	จำนวนครู
โรงเรียนสังกัดการประถมศึกษา	183
โรงเรียนในสังกัดกรมสามัญศึกษา	126
โรงเรียนสังกัดกรมอาชีวศึกษา	31
รวม	340

3. สุ่มตัวอย่างครูในแต่ละชั้นภูมิด้วยวิธีการสุ่มอย่างมีระบบ โดยอาศัยบัญชีรายชื่อข้าราชการครู ซึ่งได้จัดเรียงลำดับตามอายุ

แต่การศึกษาครั้งนี้ได้ทำการเก็บแบบสอบถามจากครูสตรี จำนวน 400 คน ทั้งนี้เนื่องจากในทางปฏิบัติได้ส่งแบบสอบถามเกินกว่าจำนวนที่ระบุไว้ เพื่อป้องกันการได้รับแบบสอบถามคืน ไม่ครบตามจำนวนที่ระบุ โดยสามารถแบ่งออกตามสังกัด ดังนี้

ตาราง 3.4 แสดงขนาดตัวอย่างครูสตรีวัยหมดประจำเดือนในเขตอำเภอเมืองจังหวัดพิษณุโลก แยกตามสังกัด

สังกัด	จำนวนครู
โรงเรียนสังกัดการประถมศึกษา	214
โรงเรียนในสังกัดกรมสามัญศึกษา	150
โรงเรียนสังกัดกรมอาชีวศึกษา	36
รวม	400

### 3.2 เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยครั้งนี้ ได้ใช้แบบสอบถามจำนวน 5 ตอน คือ  
ตอนที่ 1 ข้อมูลส่วนบุคคล เป็นข้อมูลทั่วไป แบ่งออกเป็น

ปัจจัยด้านประชากร สังคมและเศรษฐกิจ ประกอบด้วย อายุตนเอง น้ำหนัก ส่วนสูง ศาสนา การปฏิบัติงาน จำนวนปีที่ศึกษา สถานภาพสมรส อายุสามี รายได้ครัวเรือน รายจ่ายครัวเรือน ความเพียงพอของรายได้ ลักษณะครอบครัว และโรคประจำตัว

ปัจจัยด้านการเจริญพันธุ์ ประกอบด้วย จำนวนบุตร จำนวนครั้งการคลอดบุตร จำนวนครั้งการแท้งบุตร การคลอดบุตรครั้งสุดท้าย การคุมกำเนิด ระยะการคุมกำเนิด วิธีการคุมกำเนิด ลักษณะประจำเดือนหลังการคุมกำเนิด อายุการมีประจำเดือนครั้งแรก ลักษณะการมีประจำเดือนของสตรีวัยหมดประจำเดือน และการเข้ารับบริการในคลินิกสตรีวัยหมดประจำเดือน

ตอนที่ 2 อาการที่เกี่ยวกับความไม่สุขสบายของภาวะการหมดประจำเดือน มีทั้งหมด 20 ข้อ เป็นการถามเกี่ยวกับความรุนแรงของกลุ่มอาการความไม่สุขสบายของภาวะการหมดประจำเดือน โดยที่คำถามแต่ละข้อจะมีคำตอบให้เลือก 4 คำตอบ ดังนี้ ไม่มี เป็นบางครั้งหรือนาน ๆ ครั้ง บ่อยครั้ง และเป็นประจำ

ตอนที่ 3 ความรู้เกี่ยวกับภาวะการหมดประจำเดือน มีทั้งหมด 10 ข้อ เป็นข้อคำถามเกี่ยวกับ ความรู้ของสตรีวัยหมดประจำเดือนเกี่ยวกับการเปลี่ยนแปลงทางสรีรวิทยา มีคำตอบให้เลือก 2 ข้อ คือ ใช่กับไม่ใช่

ตอนที่ 4 ทักษะคิดเกี่ยวกับภาวะการหมดประจำเดือน มีทั้งหมด 10 ข้อ เป็นข้อคำถามเกี่ยวกับ ความรู้สึก ความเชื่อของสตรีวัยหมดประจำเดือนเกี่ยวกับการหมดประจำเดือน การรับประทานอาหาร การออกกำลังกาย การมีเพศสัมพันธ์ และการเข้ารับการรักษาในคลินิกสตรีวัยหมดประจำเดือนคำถามแต่ละข้อจะมีคำตอบให้เลือก 3 คำตอบ โดยมีความหมาย ดังนี้

ไม่เห็นด้วย หมายถึง ข้อความนั้นไม่ตรงกับความเข้าใจ ความรู้สึกนึกคิดหรือความเชื่อของนักเรียนเป็นส่วนมาก

ไม่แน่ใจ หมายถึง นักเรียนไม่แน่ใจว่ามีความเข้าใจ ความรู้สึกนึกคิดหรือความเชื่ออย่างไรกับข้อความนั้น และไม่สามารถตัดสินใจได้ว่าเห็นด้วยหรือไม่เห็นด้วย

เห็นด้วย หมายถึง ข้อความนั้นตรงกับความเข้าใจ ความรู้สึกนึกคิดหรือความเชื่อของนักเรียนเป็นส่วนมาก

ตอนที่ 5 การปฏิบัติตัวเพื่อดูแลสุขภาพของสตรีวัยหมดประจำเดือน มีทั้งหมด 10 ข้อ เป็นข้อคำถาม เกี่ยวกับ การปฏิบัติตัวให้เหมาะสมกับการเปลี่ยนแปลง ที่เกิดขึ้นเนื่องจากภาวะการหมดประจำเดือน มีคำตอบให้เลือก 2 ข้อ คือ ปฏิบัติหรือไม่ปฏิบัติ

### เกณฑ์การให้คะแนน

1. ตอนที่ 2 อาการที่เกี่ยวกับความไม่สุขสบายของภาวะการหมดประจำเดือน มีหลักเกณฑ์การให้คะแนนดังนี้

ไม่มี	ให้	0	คะแนน
เป็นบางครั้งหรือนานๆครั้ง	ให้	1	คะแนน
บ่อยครั้ง	ให้	2	คะแนน
เป็นประจำ	ให้	3	คะแนน

2. ตอนที่ 3 ความรู้เกี่ยวกับภาวะการหมดประจำเดือน มีหลักเกณฑ์การให้คะแนนคือ ตอบถูกได้ 1 คะแนน ตอบผิดหรือไม่ตอบได้ 0 คะแนน

3. ตอนที่ 4 ทักษะคิดเกี่ยวกับภาวะการหมดประจำเดือน มีหลักเกณฑ์การให้คะแนนตามวิธีของไลเคิร์ต (Likert) ดังนี้

3.1 ข้อความด้านบวก การให้คะแนนเป็นดังนี้

เห็นด้วย	ให้	3	คะแนน
ไม่แน่ใจ	ให้	2	คะแนน
ไม่เห็นด้วย	ให้	1	คะแนน

3.2 ข้อความด้านลบ การให้คะแนนเป็นดังนี้

เห็นด้วย	ให้	1	คะแนน
ไม่แน่ใจ	ให้	2	คะแนน
ไม่เห็นด้วย	ให้	3	คะแนน

4. ตอนที่ 5 การปฏิบัติตัวเพื่อดูแลสุขภาพของสตรีวัยหมดประจำเดือน มีหลักเกณฑ์การให้คะแนน ดังนี้ ปฏิบัติตัวถูกต้องได้ 1 คะแนน ตอบผิดหรือไม่ตอบได้ 0 คะแนน

### การแปลความหมาย

1. ตอนที่ 2 อาการที่เกี่ยวกับความไม่สุขสบายของภาวะการหมดประจำเดือน มีการแปลความหมาย โดยใช้เกณฑ์ประเมิน ดังนี้ (คะแนนเต็ม 60 คะแนน)(กัลยา วานิชชัยบัญชา, 2544)

คะแนน 41-60 อยู่ในกลุ่มที่มีอาการรุนแรงมาก

คะแนน 21-40 อยู่ในกลุ่มที่มีอาการรุนแรง

คะแนน 1-20 อยู่ในกลุ่มที่มีอาการไม่รุนแรง

คะแนน 0 อยู่ในกลุ่มที่ไม่มีอาการ

2. ตอนที่ 3 ความรู้เกี่ยวกับภาวะการหมดประจำเดือน มีการแปลความหมาย ดังนี้ คือ นำคะแนนที่ได้จากการทดสอบโดยคะแนนเต็ม 10 คะแนน นำมาหาค่าร้อยละ แล้วนำไปเปรียบเทียบกับเกณฑ์ประเมินของกระทรวงศึกษาธิการ (ทัสสนี นุชประยูรและเติมศรี ชำนิจารกิจ,2537)

คะแนนร้อยละ 80-100 อยู่ในเกณฑ์ดีมาก

คะแนนร้อยละ 60-79 อยู่ในเกณฑ์ดี

คะแนนร้อยละ 50-59 อยู่ในเกณฑ์พอใช้

คะแนนร้อยละ 0-49 อยู่ในเกณฑ์ที่ต้องแก้ไข

3. ตอนที่ 4 ทักษะคิดเกี่ยวกับภาวะการหมดประจำเดือน มีการแปลความหมาย ดังนี้ คือ ผู้วิจัยได้จำแนกระดับทักษะคิดออกเป็น 3 ระดับ ได้แก่ ทักษะคิดอยู่ในระดับต่ำ ทักษะคิดอยู่ในระดับปานกลางและทักษะคิดอยู่ในระดับสูง โดยใช้เกณฑ์ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานในการแบ่งระดับคะแนน (Munro, 2001) ได้ผลดังนี้

ทักษะคิดอยู่ในระดับต่ำ หมายถึง คะแนนรวมที่ได้อยู่ในระดับต่ำกว่าค่า  $X - SD$

ทักษะคิดอยู่ในระดับปานกลาง หมายถึง คะแนนรวมที่ได้อยู่ในระดับค่า  $X \pm SD$

ทักษะคิดอยู่ในระดับสูง หมายถึง คะแนนรวมที่ได้อยู่ในระดับสูงกว่าค่า  $X + SD$

จากข้อมูลที่รวบรวมมาทำการวิจัยในครั้งนี้ได้คะแนนทักษะคิดเกี่ยวกับภาวะการหมดประจำเดือน ซึ่งมีคะแนนเฉลี่ยเท่ากับ 25.78 และส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 2.98 (จากคะแนนเต็ม 30 คะแนน) ได้ผลดังนี้

ทักษะคิดอยู่ในระดับต่ำ หมายถึง คะแนนรวมที่มีต่ำกว่า 21.51

ทักษะคิดอยู่ในระดับปานกลาง หมายถึง คะแนนรวมที่มีค่าอยู่ระหว่าง 21.51-28.35

ทักษะคิดอยู่ในระดับสูง หมายถึง คะแนนรวมที่มีค่าสูงกว่า 28.35

4. ตอนที่ 5 การปฏิบัติตัวเพื่อดูแลสุขภาพของสตรีวัยหมดประจำเดือน มีการแปลความหมาย ดังนี้ คือ นำคะแนนที่ได้จากการทดสอบโดยมีคะแนนเต็ม 10 คะแนน แล้วนำมาหาค่าร้อยละ แล้วเปรียบเทียบกับเกณฑ์ ดังนี้ (สมหวัง พิธิยานุวัฒน์ และอุษาวดี จันทรสุนธิ,2537)

ปฏิบัติถูกมากกว่าร้อยละ 79 อยู่ในเกณฑ์ดี

ปฏิบัติถูก ร้อยละ 50-79 อยู่ในเกณฑ์พอใช้

ปฏิบัติถูกน้อยกว่าร้อยละ 50 อยู่ในเกณฑ์ที่ต้องแก้ไข

## การหาคุณภาพของเครื่องมือ

### การตรวจสอบความตรงของเครื่องมือ (Validity)

โดยผู้วิจัยนำแบบสอบถามไปหาความความเที่ยงตรงด้านเนื้อหาและความถูกต้องของภาษาโดยผู้ทรงคุณวุฒิ จำนวน 5 ท่าน คือ อาจารย์แพทย์ผู้เชี่ยวชาญด้านการดูแลสุขภาพสตรีวัยหมดประจำเดือน 2 ท่าน อาจารย์พยาบาลผู้เชี่ยวชาญด้านการดูแลสุขภาพสตรีวัยหมดประจำเดือน 3 ท่าน แล้วจึงนำมาปรับปรุงแก้ไขตามความคิดเห็นและข้อเสนอแนะของผู้ทรงคุณวุฒิอีกครั้ง ก่อนนำไปทดลองใช้

### การตรวจสอบความเชื่อมั่นของเครื่องมือ (Reliability)

ผู้วิจัยได้นำแบบสอบถามไปทดลองใช้กับกลุ่มครูสตรีวัยหมดประจำเดือนที่มีลักษณะเดียวกับกลุ่มตัวอย่างที่จะศึกษาจำนวน 30 ราย แล้วตรวจสอบวัดความเชื่อถือได้ของแบบสอบถาม ด้วยวิธีการวิเคราะห์โดยวิธี ครอนบาค แอลฟา (Cronbach's  $\alpha$ ) ได้ค่าความเชื่อมั่นของค่าวัดของข้อคำถามเกี่ยวกับเรื่อง อาการที่เกี่ยวกับความไม่สุขสบาย ความรู้และทัศนคติเกี่ยวกับภาวะการหมดประจำเดือน และการปฏิบัติตัวเพื่อดูแลสุขภาพเท่ากับ 84.45 %

### 3.3 การเก็บรวบรวมข้อมูล

1. ติดต่อสำนักงานการประถมศึกษา สำนักงานสามัญศึกษาและสถาบันการศึกษาในสังกัดกรมอาชีวศึกษาในเขตอำเภอเมืองจังหวัดพิษณุโลก เพื่อทราบจำนวนครูสตรีที่อายุตั้งแต่ 40 ปีขึ้นไป เพื่อกำหนดกรอบตัวอย่าง
2. ติดต่อภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยเชียงใหม่ เพื่อทำหนังสือขอความร่วมมือไปยังโรงเรียนในสังกัดสำนักงานการประถมศึกษา สำนักงานสามัญศึกษาและสถาบันการศึกษาในสังกัดกรมอาชีวศึกษาในเขตอำเภอเมืองจังหวัดพิษณุโลก
3. ขอความร่วมมือจากทางโรงเรียน ในสังกัดสำนักงานการประถมศึกษา สำนักงานสามัญศึกษาและสถาบันการศึกษาในสังกัดกรมอาชีวศึกษา ในเขตอำเภอเมือง จังหวัดพิษณุโลก โดยในบางโรงเรียนผู้วิจัยได้ดำเนินการเก็บแบบสอบถามเอง โดยนำแบบสอบถามไปให้กลุ่มตัวอย่างตอบและรอนำกลับ ในบางโรงเรียนได้ขอความร่วมมือกับเจ้าหน้าที่ ให้ช่วยดำเนินการเก็บแบบสอบถาม โดยที่ผู้วิจัยได้อธิบายให้เจ้าหน้าที่เข้าใจถึงวัตถุประสงค์ของการวิจัย และลักษณะของกลุ่มตัวอย่างที่จะศึกษา รวมทั้งชี้แจงให้เข้าใจเนื้อหาของแบบสอบถาม

### 3.4 ตัวแปรที่ใช้ในการวิเคราะห์

#### ตัวแปรอิสระที่นำมาวิเคราะห์มีดังนี้

##### ปัจจัยด้านประชากร

- อายุ (AGE)
- ศาสนา(REG1)
- ค่าดัชนีมวลกาย (BMI)
- การศึกษา(EDU1)
- การปฏิบัติงานสอน(WORK1)
- การปฏิบัติงานเพิ่ม(WORK2)
- สถานภาพสมรส(STATUS)
- อายุสามี (AGEHBCA)
- การมีโรคประจำตัว(DISEASE)

##### ปัจจัยด้านสังคมและเศรษฐกิจ

- รายได้ครัวเรือน(INCOME)
- รายจ่ายครัวเรือน(EXPH)
- ความเพียงพอของรายได้(EN)
- ลักษณะครอบครัว(FAMILY)
- จำนวนสมาชิกในครัวเรือน(NFAMILY)

##### ปัจจัยด้านการเจริญพันธุ์

- จำนวนครั้งของการผ่านการคลอดบุตร(PARITY)
- จำนวนครั้งการแท้งบุตร(ABORTION)
- จำนวนบุตรที่มีชีวิตอยู่(CHILD)
- วิธีการคุมกำเนิด(TYPECTCP)
- ลักษณะประจำเดือนหลังการคุมกำเนิด(MENA)
- อายุเมื่อมีประจำเดือนครั้งแรก(AGESTMEN)
- ลักษณะการมีประจำเดือน(MENS)

##### ปัจจัยด้านความไม่สุขสบาย

- คะแนนรวมความรุนแรงของอาการไม่สุขสบาย(TOTALSMP)



### ปัจจัยด้านความรู้ ทักษะ และการปฏิบัติตัวเพื่อดูแลสุขภาพตนเอง

- คะแนนรวมความรู้เกี่ยวกับภาวะการหมดประจำเดือน(TOTALKN)
- คะแนนรวมทักษะ (TOTALATI)
- คะแนนรวมการปฏิบัติตัวที่ถูกต้อง(TOTALPAC)

### ตัวแปรตาม คือ

- การเข้ารับบริการในคลินิกสตรีวัยหมดประจำเดือน (CLINIC)

### 3.5 สถิติที่ใช้ในการวิเคราะห์

1. วิเคราะห์หาค่าความเชื่อถือได้ของแบบสอบถาม โดยใช้สูตรสัมประสิทธิ์แอลฟาของครอนบาค (Chronbach's Alpha Coefficient)
2. นำเสนอข้อมูลทั่วไป โดยใช้สถิติพรรณนา แสดงความถี่ ร้อยละ ค่าเฉลี่ย ค่าต่ำสุด ค่าสูงสุดและส่วนเบี่ยงเบนมาตรฐาน
3. ศึกษาความสัมพันธ์ของทักษะเกี่ยวกับภาวะการหมดประจำเดือนรายข้อ กับการเข้ารับบริการหรือไม่เข้ารับบริการในคลินิกวัยหมดประจำเดือน โดยทดสอบความสัมพันธ์และวัดระดับความสัมพันธ์โดยใช้ Gamma และศึกษาความสัมพันธ์ของคะแนนรวมของทักษะเกี่ยวกับภาวะการหมดประจำเดือนกับการเข้ารับบริการหรือไม่เข้ารับบริการในคลินิกสตรีวัยหมดประจำเดือน โดยทดสอบความสัมพันธ์ด้วย Eta
4. ศึกษาปัจจัยที่มีผลต่อ ทักษะเกี่ยวกับภาวะการหมดประจำเดือนของครูที่เข้ารับบริการและไม่เข้ารับบริการในคลินิกสตรีวัยหมดประจำเดือน โดยใช้เทคนิคการวิเคราะห์ถดถอยเชิงกลุ่ม (Category Regression)
5. ศึกษาปัจจัยที่มีผลต่อ การเข้ารับบริการหรือไม่เข้ารับบริการในคลินิกสตรีวัยหมดประจำเดือน โดยใช้เทคนิคการวิเคราะห์ถดถอยเชิงกลุ่ม (Category Regression)
6. วิเคราะห์หาสมการความสัมพันธ์ ของปัจจัยที่มีผลต่อการเข้ารับบริการหรือไม่เข้ารับบริการในคลินิกสตรีวัยหมดประจำเดือน โดยใช้เทคนิคการวิเคราะห์ถดถอยโลจิสติก (Logistic Regression Analysis)

### ความหมายที่ควรทราบ

#### Significance (Sig) ของค่าสถิติใด ๆ

หมายถึง ค่าที่ใช้ในการตัดสินใจว่าจะยอมรับหรือปฏิเสธสมมติฐาน ( $H_0$ ) จากการใช้ค่าสถิติวิเคราะห์ โดยจะนำไปเปรียบเทียบกับค่าระดับนัยสำคัญ ( $\alpha$ )

การตัดสินใจเกี่ยวกับสมมติฐานที่ระดับนัยสำคัญ  $\alpha$

จะยอมรับ  $H_0$  ถ้า Sig (ค่าสถิติ)  $> \alpha$

จะปฏิเสธ  $H_0$  ถ้า Sig (ค่าสถิติ)  $\leq \alpha$

#### สัมประสิทธิ์แอลฟาของครอนบาค (Chronbach's Alpha Coefficient : Alpha)

หมายถึง ค่าความเชื่อถือได้ หรือค่าความสัมพันธ์ระหว่างตัวแปร มีสูตรดังนี้

$$\alpha = \frac{K}{K-1} \left[ 1 - \frac{\sum_{i=1}^K S_i^2}{S^2} \right]$$

โดยที่  $K$  = จำนวนข้อของคำถามเกี่ยวกับเรื่อง อาการที่เกี่ยวกับความไม่สุขสบาย ความรู้และทัศนคติเกี่ยวกับภาวะการหมดประจำเดือน และการปฏิบัติตัวเพื่อดูแลสุขภาพ

$S_i^2$  = ความแปรปรวนของค่าวัดแบบสอบถามเกี่ยวกับเรื่อง อาการที่เกี่ยวกับความไม่สุขสบาย ความรู้และทัศนคติเกี่ยวกับภาวะการหมดประจำเดือน และการปฏิบัติตัวเพื่อดูแลสุขภาพ แต่ละข้อ

$S^2$  = ความแปรปรวนของค่าวัดแบบสอบถามเกี่ยวกับเรื่อง อาการที่เกี่ยวกับความไม่สุขสบาย ความรู้และทัศนคติเกี่ยวกับภาวะการหมดประจำเดือน และการปฏิบัติตัวเพื่อดูแลสุขภาพ ทั้งฉบับ

#### Gamma

เพื่อหาค่าความสัมพันธ์ระหว่างตัวแปรเชิงอันดับ กับตัวแปรเชิงอันดับ โดยสามารถบอกได้ทั้งขนาดและทิศทาง มีสูตรดังนี้

$$\gamma = \frac{P-Q}{P+Q}$$

โดยที่

$$-1 < \gamma < 1$$

โดยที่  $P$  คือ จำนวนคู่ที่สอดคล้องกัน (concordant pairs)  
 $Q$  คือ จำนวนคู่ที่ไม่สอดคล้องกัน (discordant pairs)

การคำนวณหาค่า  $P$  และ  $Q$  จะพิจารณาจากตาราง

ตัวแปรตาม ตัวแปรอิสระ	ระดับที่ 1	ระดับที่ 2	ระดับที่ 3
ระดับที่ 1	a	b	c
ระดับที่ 2	d	e	f

โดยที่  $P = a(e + f) + bf$

$$Q = c(d + e) + bd$$

Eta

เพื่อหาค่าความสัมพันธ์พร้อมทดสอบความสัมพันธ์ระหว่างตัวแปรเชิงปริมาณ กับตัวแปรกลุ่ม มีสูตรดังนี้

$$Eta^2 = \frac{\text{Sum of Squares Between Group}}{\text{Total Sum of Squares}} = \frac{\sum \frac{T_i^2}{n} - \frac{(\sum T_i)^2}{n}}{\sum \sum X_{ij}^2 - \frac{(\sum T_i)^2}{n}}$$

โดยที่  $F = \frac{MSR}{MSE}$

และ  $MSR = \frac{SSR}{1} = \sum \frac{T_i^2}{n} - \frac{(\sum T_i)^2}{n}$

$$MSE = \frac{SSE}{n-2} = \frac{\sum \sum X_{ij}^2 - \sum \frac{T_i^2}{n_i}}{n-2}$$

เมื่อ  $X_{ij}$  : ข้อมูลของหน่วยตัวอย่างในแถวที่  $i$  คอลัมน์ที่  $j$

$T_i$  : ผลรวมของข้อมูลตัวอย่างชุดที่  $i$

$n$  : จำนวนข้อมูลทั้งหมด

มีสมมติฐานในการทดสอบ คือ

$H_0$  : ตัวแปรทั้ง 2 ไม่มีความสัมพันธ์กัน

$H_1$  : ตัวแปรทั้ง 2 มีความสัมพันธ์กัน

จะยอมรับ  $H_0$  ถ้า  $\text{Sig}(F) > \alpha$

จะปฏิเสธ  $H_0$  ถ้า  $\text{Sig}(F) \leq \alpha$

### การวิเคราะห์ถดถอยเชิงกลุ่ม (Category Regression)

เทคนิคการวิเคราะห์ถดถอยเชิงกลุ่มเป็นวิธีการวิเคราะห์ที่มีวัตถุประสงค์เพื่อเปรียบเทียบความสำคัญของตัวแปรอิสระทั้งหลายที่มีต่อตัวแปรตาม

เมื่อ Y คือ ตัวแปรตาม

$X_i$  คือ ตัวแปรอิสระตัวที่ i

โดยตัวแปรตามและตัวแปรอิสระอยู่ในระดับ Optimal scaling ซึ่งระดับ Optimal scaling มี 3 ระดับ ดังนี้

1. Nominal level ใช้กับตัวแปรที่มีมาตรวัดนามบัญญัติ กล่าวคือ ค่าของตัวแปรไม่สามารถเปรียบเทียบได้ว่ากลุ่มใดมากกว่ากัน เช่น เพศ อาชีพ

2. Ordinal level ใช้กับตัวแปรที่มีมาตรวัดเรียงอันดับ กล่าวคือ ค่าของตัวแปรสามารถเปรียบเทียบได้ว่ากลุ่มใดมากกว่ากัน แต่ไม่สามารถบอกปริมาณความมากกว่าหรือน้อยกว่าว่าเป็นเท่าใด เช่น ระดับความพึงพอใจ ระดับทัศนคติ

3. Numerical level ใช้กับตัวแปรที่มีมาตรวัดอันตรภาค หรือมาตรวัดอัตราส่วน กล่าวคือค่าของตัวแปรสามารถบอกปริมาณความแตกต่างได้ว่าค่าใดมากกว่ากันและบอกได้ว่าแต่ละค่าห่างกันเท่าใด เช่น รายได้ น้ำหนัก

### การหาค่าความสัมพันธ์

1. สัมประสิทธิ์การถดถอยมาตรฐาน ( $Beta_i$ ) ของตัวแปรอิสระใด หมายถึง น้ำหนักความสำคัญของตัวแปรอิสระนั้นที่มีต่อตัวแปรตาม ค่านี้เป็นค่ามาตรฐานของสัมประสิทธิ์การถดถอย โดยที่

$$Beta_i = \hat{\beta}_i \frac{S_i}{S_y}$$

เมื่อ  $\hat{\beta}_i$  : ค่าสัมประสิทธิ์การถดถอยของตัวแปรอิสระ  $X_i$  ของตัวอย่าง

$S_i, S_y$  : ค่าเบี่ยงเบนมาตรฐานของตัวแปรอิสระตัวที่ i และ

ตัวแปรตาม y ตามลำดับ ของตัวอย่าง

มีสมมติฐานสำหรับการทดสอบค่าสัมประสิทธิ์การถดถอยของประชากร ( $\beta_i$ ) คือ

$$H_0: \beta_i = 0$$

$H_1$ : มี  $\beta_i$  อย่างน้อย 1 ตัวที่ไม่เป็น 0

โดยใช้สถิติทดสอบ

$$t = \frac{\hat{\beta}_i - \beta_i}{s_{\hat{\beta}_i}}$$

และ  $F = t^2$

เมื่อ  $S_{\hat{\beta}_i}$ : ค่าเบี่ยงเบนมาตรฐานของสัมประสิทธิ์การถดถอยของตัวอย่าง

จะยอมรับ  $H_0$  ถ้า  $F < 3.84$  หรือ  $\text{Sig } \alpha \geq 0.05$

จะปฏิเสธ  $H_0$  ถ้า  $F \geq 3.84$  หรือ  $\text{Sig } \alpha < 0.05$

2. Zero Order ( $r_i$ ): ค่าสัมประสิทธิ์สหสัมพันธ์ของตัวแปรอิสระนั้นกับตัวแปรตาม

3.  $R^2$ : ค่าที่แสดงให้ทราบว่า ตัวแปรอิสระทั้งหลายในสมการสามารถอธิบายความผันแปรของตัวแปรตามได้มากน้อยเพียงใด

4. Part Correlation: ค่ารากที่สองของ  $R^2$  ที่เปลี่ยนไป เมื่อเพิ่มตัวแปรอิสระตัวใหม่ ( $X_i$ ) เข้าไป โดยที่

$$\text{Part Correlation} = \sqrt{R_y^2 - R_i^2} = \sqrt{R_{\text{change}}^2}$$

เมื่อ  $R_i^2$ : ค่า  $R^2$  ใหม่ เมื่อเพิ่มตัวแปรอิสระ ( $X_i$ ) เข้ามา

$R_y^2$ : ค่า  $R^2$  เดิม

5. Partial Correlation ( $Pr_i$ ): ค่าสัมประสิทธิ์สหสัมพันธ์ของตัวแปรอิสระนั้นกับตัวแปรตาม เมื่อมีการขจัดอิทธิพลของตัวแปรอิสระอื่นๆ ในสมการแล้ว

$$Pr_i = \sqrt{\frac{R_y^2 - R_i^2}{1 - R_i^2}}$$

6. Importance: ค่าน้ำหนักความสำคัญของตัวแปรอิสระนั้น ที่มีต่อตัวแปรตาม จากน้ำหนักความสำคัญของตัวแปรอิสระทั้งหมด ซึ่งมีผลรวมเป็น 1.00

$$\text{Importance} = \frac{\text{Beta}_i(r_i)}{R^2}$$

### การวิเคราะห์ถดถอยโลจิสติก (Logistic Regression)(กัลยา วาณิชย์บัญชา,2544)

เป็นวิธีการวิเคราะห์ที่มีวัตถุประสงค์และแนวคิดเหมือนกับการวิเคราะห์การถดถอยทั่วไป คือ เพื่อศึกษาความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระและนำเสนอผลการถดถอยที่ได้ไปประมาณหรือพยากรณ์ค่าตัวแปรตาม เมื่อกำหนดค่าตัวแปรอิสระ ซึ่งการวิเคราะห์การถดถอยทั่วไปนั้น ตัวแปรตามจะมีลักษณะเป็นข้อมูลแบบต่อเนื่อง แต่การวิเคราะห์การถดถอยโลจิสติกนั้นตัวแปรตามเป็นข้อมูลแบบไม่ต่อเนื่อง นั่นคือเป็นตัวแปรเชิงกลุ่ม

#### การวิเคราะห์การถดถอยโลจิสติก แบ่งออกเป็น 2 ประเภท

1. การวิเคราะห์การถดถอยโลจิสติกแบบทวิ (Binary Logistic Regression) ตัวแปรตามเป็นตัวแปรเชิงกลุ่มที่มีค่าได้เพียง 2 ค่า (Dichotomous variable) เช่น เป็นหรือไม่เป็น หายหรือไม่หาย ซื้อหรือไม่ซื้อ เป็นต้น

2. การวิเคราะห์การถดถอยโลจิสติกแบบพหุ (Multinomial Logistic Regression) ตัวแปรตามเป็นตัวแปรเชิงกลุ่มที่มีค่ามากกว่า 2 ค่า เช่นในการวิเคราะห์ตลาดภาพยนตร์ บริษัทผู้ผลิตภาพยนตร์ต้องการพยากรณ์ว่า ภาพยนตร์ประเภทใดเป็นที่นิยม ซึ่งแบ่งออกเป็น 3 ประเภท คือ 1. ภาพยนตร์ประเภทบู๊ 2. ภาพยนตร์ประเภทตลก 3. ภาพยนตร์ประเภทชีวิต 4. ภาพยนตร์ประเภทรักโรแมนติก เป็นต้น

ในการศึกษาครั้งนี้กล่าวเพียง Binary Logistic Regression เท่านั้น

#### การวิเคราะห์การถดถอยโลจิสติกแบบทวิ (Binary Logistic Regression)

เมื่อตัวแปรตามมีค่าได้เพียง 2 ค่า คือ 0 (ไม่เกิดเหตุการณ์ที่สนใจ) กับ 1 (เกิดเหตุการณ์ที่สนใจ) และ ตัวแปรอิสระเป็นตัวแปรเชิงปริมาณหรือตัวแปรทวิ จะพบว่าความสัมพันธ์ระหว่างตัวแปรอิสระ และตัวแปรตาม ไม่ได้อยู่ในรูปเชิงเส้น แต่อยู่ในรูปแบบของสมการการถดถอยโลจิสติก ดังนี้

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p \quad (\text{ประชากร})$$

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \dots + \hat{\beta}_p X_p \quad (\text{ตัวอย่าง}) \quad (1.1)$$

$$P(\text{เกิดเหตุการณ์}) = \frac{e^{\hat{Y}}}{1 + e^{\hat{Y}}} = \frac{1}{1 + e^{-\hat{Y}}} \quad (1.2)$$

เมื่อ  $X_i$  : ตัวแปรอิสระซึ่งอาจเป็นตัวแปรเชิงปริมาณหรือตัวแปรทวิ :  $i = 1, 2, \dots, p$

$\beta_i$  : เป็นสัมประสิทธิ์การถดถอยโลจิสติกของ  $X_i$

เรียกสมการที่ (1.2) ว่า Logistic Response Function ดังนั้นการพยากรณ์ค่าความน่าจะเป็นของเหตุการณ์ คือ

$$P(\text{เกิดเหตุการณ์}) = \frac{e^{\hat{y}}}{1+e^{\hat{y}}} = \frac{1}{1+e^{-\hat{y}}}$$

$$P(\text{ไม่เกิดเหตุการณ์}) = 1 - P(\text{เกิดเหตุการณ์}) = \frac{1}{1+e^{\hat{y}}}$$

เมื่อ  $P(\text{เกิดเหตุการณ์})$  หมายถึง ความน่าจะเป็นที่จะเกิดเหตุการณ์ที่สนใจ และ  $P(\text{ไม่เกิดเหตุการณ์})$  หมายถึง ความน่าจะเป็นที่จะไม่เกิดเหตุการณ์ที่สนใจ สำหรับผู้ที่มีลักษณะของปัจจัยเสี่ยง หากความน่าจะเป็นที่จะเกิดเหตุการณ์มีค่าน้อยกว่า 0.5 สามารถตีความได้ว่าเหตุการณ์นั้นจะมีโอกาสเกิดขึ้นได้น้อยกว่าเหตุการณ์ที่ไม่สนใจ ถ้าความน่าจะเป็นที่จะเกิดเหตุการณ์มีค่ามากกว่า 0.5 แสดงว่าเหตุการณ์ที่สนใจมีโอกาที่จะเกิดขึ้นมากกว่าอีกเหตุการณ์หนึ่ง แต่ถ้าความน่าจะเป็นมีค่าเท่ากับ 0.5 อาจตัดสินใจว่า โอกาสที่จะเกิดหรือไม่เกิดเหตุการณ์ที่สนใจกับอีก เหตุการณ์หนึ่งเท่ากัน

สามารถปรับสมการความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระที่ไม่ได้อยู่ในรูปแบบเชิงเส้น ให้อยู่ในรูปแบบเชิงเส้นได้โดยทำให้เป็นอัตราส่วน odds (odds ratio) โดยที่

Odds ratio หมายถึง อัตราส่วนระหว่างโอกาสที่เกิดเหตุการณ์ใดเหตุการณ์หนึ่ง กับโอกาสที่จะไม่เกิดเหตุการณ์นั้น สำหรับผู้ที่มีลักษณะของปัจจัยเสี่ยงอย่างใดอย่างหนึ่ง

ดังนั้น

$$\text{odds} = \frac{P(\text{เกิดเหตุการณ์})}{P(\text{ไม่เกิดเหตุการณ์})} = \frac{e^{\hat{y}}/1+e^{\hat{y}}}{1/1+e^{\hat{y}}} = e^{\hat{y}} \quad (1.3)$$

$$\log(\text{odds}) = \log\left[ \frac{P(\text{เกิดเหตุการณ์})}{P(\text{ไม่เกิดเหตุการณ์})} \right] \quad (1.4)$$

$$\text{หรือ } \log(\text{odds}) = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \dots + \hat{\beta}_p X_p \quad (1.5)$$

สำหรับการประมาณค่า  $\hat{y}$  จะประมาณจาก  $P(\text{เกิดเหตุการณ์})$  โดยใช้สมการ (1.2) และสำหรับการประมาณค่าสัมประสิทธิ์การถดถอย  $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_p$  จะประมาณโดยใช้วิธีภาวะน่าจะเป็นสูงสุด (Maximum Likelihood)

การตีความหมายค่าต่าง ๆ ที่ได้จากการสร้างแบบจำลองการถดถอยโลจิสติก

1. ค่า Initial -2 times log-likelihood (-2LL<sub>0</sub>) เป็นค่าที่แบบจำลองสร้างขึ้น เมื่อเริ่มสร้างสมการ โดยไม่มีตัวแปรอิสระใด ๆ เข้ามาเกี่ยวข้อง เป็นค่าที่บอกถึงการวัดค่าความเหมาะสมของ logistic model โดยพิจารณาจากค่า -2 log likelihood (-2LL)

ถ้าค่า -2LL มีค่าน้อยหรือเข้าใกล้ศูนย์ แสดงว่าแบบจำลองมีความเหมาะสม

2. ค่า  $\beta_0$  เป็นค่าคงที่ของสมการ และสมมติฐานของการทดสอบค่าคงที่ คือ

$$H_0 : \text{ค่าคงที่ของแบบจำลองมีค่าเท่ากับ } 0 (\beta_0 = 0)$$

$$H_1 : \text{ค่าคงที่ของแบบจำลองมีค่าไม่เท่ากับ } 0 (\beta_0 \neq 0)$$

3. ค่าสัมประสิทธิ์การถดถอย (Regression Coefficientor :  $\hat{\beta}_i$ ) เป็นค่าความชันของเส้นตรง คือค่าที่แสดงการเปลี่ยนแปลงของ  $\ln[p(1-p)]$  นี้จะแสดงให้เห็นว่าเมื่อ X เปลี่ยนไป 1 หน่วย Y จะเปลี่ยนไปโดยเฉลี่ยเท่าไร

4. Wald-Statistic เป็นตัวสถิติที่นำมาทดสอบว่า ตัวแปรอิสระใดควรเข้าในสมการกรณีกลุ่มตัวอย่างข้อมูลมีขนาดใหญ่ โดยมีการแจกแจงแบบไคสแควร์ และมีค่าความเป็นอิสระ (degree of freedom) เป็น 1 ค่า Wald มีสูตรการคำนวณ ดังนี้

$$\text{Wald} = \left( \frac{\hat{\beta}_i}{\text{S.E.}_i} \right)^2$$

โดยที่  $\hat{\beta}_i$  คือ ค่าสัมประสิทธิ์การถดถอยของตัวแปรอิสระตัวที่ i

S.E.<sub>i</sub> คือ ค่าความคลาดเคลื่อนมาตรฐาน (Standard error) ของค่าสัมประสิทธิ์การถดถอยตัวแปรอิสระตัวที่ i ( $\hat{\beta}_i$ )

สมมติฐานสำหรับ Wald-Statistic คือ

$$H_0 : \text{สัมประสิทธิ์การถดถอยของตัวแปรอิสระใด ๆ เท่ากับ } 0 (\beta_i = 0)$$

$$H_1 : \text{สัมประสิทธิ์การถดถอยของตัวแปรอิสระใด ๆ ไม่เท่ากับ } 0 (\beta_i \neq 0)$$



5. **Pesudo  $R^2$**  เป็นค่าที่บอกสัดส่วน หรือเปอร์เซ็นต์ที่สามารถอธิบายความผันแปร ใน logistic regression model ซึ่งจะคล้ายกับค่า  $R^2$  ในการวิเคราะห์ความถดถอยเชิงเส้น

$$\diamond \text{ Cox \& Snell } R^2 = 1 - \left[ \frac{L(0)}{L(B)} \right]^{\frac{2}{W}}$$

โดยที่  $L(0)$  คือค่า Likelihood สำหรับ Model ที่มีเพียงค่าคงที่

$L(B)$  คือค่า Likelihood สำหรับ Model ที่มีตัวแปรอิสระตามที่กำหนด

$$W = \sum_i^n w_i \text{ เมื่อ } w_i \text{ คือ ค่าถ่วงน้ำหนักของตัวอย่างตัวที่ } i$$

$$\diamond \text{ Nagelkerke } R^2 = \frac{\text{Cox \& Snell } R^2}{\text{Cox \& Snell } R^2_{Max}}$$

6. **ค่าแบบจำลองไคสแควร์ (Model Chi-square)** เป็นค่าความแตกต่างระหว่าง  $-2LL_0$  กับ  $-2LL_1$  ซึ่งบอกให้ทราบว่า สมการในแบบจำลองนั้นเหมาะสมหรือไม่ ซึ่งเปรียบเสมือนค่า  $F$  ในการวิเคราะห์ถดถอยพหุ (Multiple regression) มีสมมติฐานของค่าแบบจำลองไคสแควร์ คือ

$H_0$ : สัมประสิทธิ์การถดถอยของตัวแปรอิสระทุกตัวในแบบจำลองเท่ากับ 0

$$\text{หรือ } \beta_1 = \beta_2 = \dots = \beta_p = 0$$

$H_1$ : มีสัมประสิทธิ์การถดถอยของตัวแปรอิสระในแบบจำลองอย่างน้อยหนึ่งตัวที่ไม่เท่ากับ 0

7. **ค่าปรับปรุง (Improvement)** เป็นค่าการเปลี่ยนแปลงของ  $-2$  times Log-likelihood เมื่อเสร็จขั้นตอนการสร้างแบบจำลองแล้ว ซึ่งบอกให้ทราบว่าสัมประสิทธิ์ของตัวแปรอิสระตัวสุดท้ายนั้นควรเข้าในแบบจำลองหรือไม่ ค่าปรับปรุงนี้เปรียบเสมือน  $F$ -Change ในการวิเคราะห์ถดถอยพหุ (Multiple Regression) สมมติฐานของค่าปรับปรุง คือ

$H_0$ : สัมประสิทธิ์การถดถอยของตัวแปรอิสระตัวสุดท้ายที่เข้าในแบบจำลอง เท่ากับ 0

$H_1$ : สัมประสิทธิ์การถดถอยของตัวแปรอิสระตัวสุดท้ายที่เข้าในแบบจำลองไม่เท่ากับ 0

การพิจารณาความเหมาะสมของแบบจำลอง

1. **ค่า Current - 2 Times Log-Likelihood ( $-2LL_1$ )** เป็นค่าที่แบบจำลองสร้างขึ้น โดยมี ตัวแปรอิสระเข้าสมการ ซึ่งจะเป็นค่าที่บอกให้ทราบว่า ตัวแปรอิสระที่อยู่ในแบบจำลอง ขณะนั้น เป็นตัวแปรที่ทำให้แบบจำลองที่ได้เหมาะสมหรือไม่ โดยสมมติฐานของค่า  $-2LL_1$  คือ

$H_0$ : แบบจำลองนี้มีความเหมาะสมกับข้อมูล

$H_1$ : แบบจำลองนี้มีความไม่เหมาะสมกับข้อมูล

ถ้าค่า  $-2LL_1$  มีค่าน้อยหรือเข้าใกล้ศูนย์ และมีค่าน้อยกว่าค่า  $-2LL_0$  แสดงว่าแบบจำลองมีความเหมาะสม

2. Hosmer-lemeshow goodness-of fit ใช้ในการตรวจสอบความเหมาะสมของแบบจำลองโดยจะแบ่งจำนวนข้อมูล (case) โดยสุ่ม ออกเป็น 10 กลุ่ม ๆ ละพอ ๆ กัน แบ่งพิจารณาจากค่าประมาณของโอกาสที่เหตุการณ์จะเกิด โดยใช้สถิติทดสอบไคสแควร์ ขนาดตัวอย่างต้องมากพอที่ทำให้จำนวนข้อมูลที่คาดไว้ (Expected value:  $E_i$  เมื่อ  $i = 1, \dots, 10$ ) จะต้องมากกว่า 5 และไม่มีกลุ่มใดที่มีค่า  $E_i < 1$  มีกรคำนวณดังนี้

$$\chi^2 = \sum_{i=1}^{10} \frac{(O_i - E_i)^2}{E_i}$$

เมื่อ  $O_i$  = ความถี่จากการสังเกต

$E_i$  = ความถี่คาดหวัง

มีสมมติฐานของค่าที่คำนวณได้จากการทดสอบภาวะรูปดี คือ

$H_0$  : แบบจำลองนี้มีความเหมาะสมกับข้อมูล

$H_1$  : แบบจำลองนี้มีความไม่เหมาะสมกับข้อมูล

จะยอมรับ  $H_0$  ถ้า  $\text{Sig}(\chi^2) > \alpha$

จะปฏิเสธ  $H_0$  ถ้า  $\text{Sig}(\chi^2) \leq \alpha$

3. ตารางการจำแนกกลุ่ม (Classification Table) ทำให้ทราบว่าตัวแปรอิสระที่ได้ในแบบจำลองมีความสามารถในการจำแนกกลุ่มจริงและกลุ่มที่ได้จากการทำนายมากหรือน้อยเพียงไร สำหรับแบบจำลองที่ได้นั้น ร้อยละในการจำแนกกลุ่มได้อย่างถูกต้องควรจะมีค่าสูง อย่างน้อยร้อยละ 80 คำนวณได้จาก

$$\text{ร้อยละการจำแนกถูกต้อง} = \frac{\text{จำนวนที่จำแนกถูกต้อง} \times 100}{\text{จำนวนที่จำแนกถูกต้อง} + \text{จำนวนที่จำแนกไม่ถูกต้อง}}$$

การคัดเลือกตัวแปรอิสระเข้าสมการ

1. Enter เป็นวิธีการเลือกตัวแปรอิสระเข้าสมการในขั้นตอนเดียว โดยให้ตัวแปรอิสระทุกตัวเข้าไปในสมการ แล้วเลือกตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม โดยพิจารณาจากค่านัยสำคัญของการทดสอบ

**2. Forward : Conditional** เป็นวิธีการเลือกตัวแปรอิสระเข้าสมการทีละตัว โดยจะพิจารณาตัวแปรที่มีค่านัยสำคัญของการทดสอบต่ำเข้าก่อนทีละตัว (เปรียบเทียบค่านัยสำคัญของการทดสอบกับค่าความน่าจะเป็นของ likelihood ratio statistic ที่ขึ้นกับค่าประมาณโดยวิธี conditional )

**3. Forward : LR** เป็นวิธีการเลือกตัวแปรอิสระเข้าสมการทีละตัว โดยจะพิจารณาตัวแปรที่มีค่านัยสำคัญของการทดสอบต่ำเข้าก่อนทีละตัว (เปรียบเทียบค่านัยสำคัญของการทดสอบกับค่าความน่าจะเป็นของ likelihood ratio statistic ที่ขึ้นกับค่าประมาณโดยวิธี maximum likelihood)

**4. Forward : Wald** เป็นวิธีการเลือกตัวแปรอิสระเข้าสมการทีละตัว โดยจะพิจารณาตัวแปรที่มีค่านัยสำคัญของการทดสอบต่ำเข้าก่อนทีละตัว (เปรียบเทียบค่านัยสำคัญของการทดสอบกับค่าความน่าจะเป็นของสถิติ Wald )

**5. Backward : Conditional** เป็นวิธีการเลือกตัวแปรอิสระทุกตัวเข้าสมการก่อน แล้วตัดตัวแปรอิสระที่มีค่านัยสำคัญของการทดสอบสูงออกก่อนทีละตัว (เปรียบเทียบค่านัยสำคัญของการทดสอบกับค่าความน่าจะเป็นของ likelihood ratio statistic ที่ขึ้นกับค่าประมาณโดยวิธี conditional )

**6. Backward : LR** เป็นวิธีการเลือกตัวแปรอิสระทุกตัวเข้าสมการก่อน แล้วตัดตัวแปรอิสระที่มีค่านัยสำคัญของการทดสอบสูงออกก่อนทีละตัว (เปรียบเทียบค่านัยสำคัญของการทดสอบกับค่าความน่าจะเป็นของ likelihood ratio statistic ที่ขึ้นกับค่าประมาณโดยวิธี maximum likelihood)

**7. Backward : Wald** เป็นวิธีการเลือกตัวแปรอิสระทุกตัวเข้าสมการก่อน แล้วตัดตัวแปรอิสระที่มีค่านัยสำคัญของการทดสอบสูงออกก่อนทีละตัว (เปรียบเทียบค่านัยสำคัญของการทดสอบกับค่าความน่าจะเป็นของสถิติ Wald)

**การตรวจสอบเพื่อให้ได้แบบจำลองที่ดีที่สุดโดยวิธีพิจารณาจากค่าความคลาดเคลื่อน**

หลังจากที่ได้แบบจำลองที่เหมาะสมแล้ว เมื่อแทนค่าตัวแปรต่าง ๆ ในแบบจำลองเพื่อพยากรณ์ความน่าจะเป็น พบว่า มีข้อมูลบางตัวที่ผิดปกติ คือ ค่าความน่าจะเป็นของกลุ่มจริงไม่สอดคล้องกับกลุ่มที่ได้จากการพยากรณ์ ซึ่งเป็นความคลาดเคลื่อนที่ไม่อาจหลีกเลี่ยงได้ ดังนั้นจึงมีความจำเป็นต้องตรวจสอบและแก้ไขในแบบจำลองที่ได้ เพื่อพยากรณ์ความน่าจะเป็นให้ใกล้เคียงกับข้อมูลจริงมากที่สุด โดยพิจารณาจากค่าความคลาดเคลื่อน ดังนี้

ค่าความคลาดเคลื่อนมาตรฐาน (Standardized Residual) เป็นค่าความคลาดเคลื่อนที่  
ปรับอยู่ในรูปแบบคะแนนมาตรฐานแล้ว โดยที่

$$Z_{\text{residual}_i} = \frac{\text{Residual}_i}{[(\text{Prob}(\text{พยากรณ์}_i)) (1 - \text{Prob}(\text{พยากรณ์}_i))]^{0.5}}$$

ถ้า  $Z_{\text{residual}_i}$  มีค่ามากกว่าที่กำหนด ( $Z_{\text{residual}_i} > 3$ ) แสดงว่าหน่วยตัวอย่าง(case)  
นั้น ควรตัดทิ้ง เพราะมีค่าความคลาดเคลื่อนมากกว่าที่กำหนด