

การพยากรณ์ความน่าจะเป็นการไม่สำเร็จการศึกษาภายใน 4 ปี ภายใต้ตัวแบบลอจิสต์ 2 กลุ่ม และตัวแบบลอจิสต์ออดส์เชิงสัดส่วน Probability Prediction of Failing to Graduate within 4 Years under Dichotomous and Proportional Odds Logit Models

ชญานูช อักษร, เพชร อ่อนลออ และวีรานันท์ พงศาภักดิ์*

ภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยศิลปากร

วิทยาเขตพระราชวังสนามจันทร์ ตำบลพระปฐมเจดีย์ อำเภอเมือง จังหวัดนครปฐม 73000

Chanyanooch Aksorn, Petcharee Onlaor and Veeranun Pongsapukdee*

Department of Statistics, Faculty of Science, Silpakorn University,

Sanam Chandra Palace Campus, Pra Pathom Chedi, Muang, Nakhon Pathom 73000

บทคัดย่อ

ปัจจุบันบัณฑิตปริญญาตรีในไทยมีจำนวนมากขึ้น แต่ยังมีปัญหาที่บัณฑิตใช้เวลาศึกษาเกินหลักสูตร 4 ปี ซึ่งอาจทำให้สูญเสียงบประมาณและโอกาสการก้าวหน้า งานวิจัยนี้ศึกษาอิทธิพลของปัจจัยเสี่ยงต่าง ๆ เพื่อนำมาพยากรณ์ความน่าจะเป็นการไม่สำเร็จการศึกษาภายใน 4 ปี โดยอาศัยข้อมูลการสำรวจประชากร 516 คนด้วยตัวอย่าง 258 คน จากนักศึกษาชั้นปีที่ 4 ปีการศึกษา 2556 คณะวิทยาศาสตร์ มหาวิทยาลัยศิลปากร ภายใต้แผนแบบการเลือกตัวอย่างแบบมีระบบ (systematic sampling design) และอาศัยตัวแบบเชิงสถิติในการวิเคราะห์ข้อมูล คือ ตัวแบบลอจิสต์ 2 กลุ่ม ที่ตัวแปรตอบสนองไม่มีอันดับ คือ ไม่สำเร็จการศึกษาและสำเร็จการศึกษา และตัวแบบลอจิสต์ออดส์เชิงสัดส่วน (proportional odds logit model) 3 กลุ่มที่ตัวแปรตอบสนองมี 3 อันดับ คือ ไม่สำเร็จการศึกษา สำเร็จการศึกษาตามเกณฑ์ และสำเร็จการศึกษาเกียรตินิยม ส่วนปัจจัยเสี่ยงศึกษาจากวรรณกรรมที่เกี่ยวข้องมี 30 ปัจจัย (x_1 - x_{30}) ผลของการวิจัยพบว่าตัวแบบมีภาวะสารูปดีและมี 4 ปัจจัยเสี่ยง ที่ส่งผลกระทบต่ออย่างมีนัยสำคัญเชิงสถิติที่ 0.05 คือ x_{18} , x_{21} , x_{27} และ x_{29} ผลการพยากรณ์ความน่าจะเป็นด้วยตัวแบบลอจิสต์ 2 กลุ่ม เช่น นักศึกษาคนหนึ่งมีพฤติกรรมเสียเวลาไปกับการเที่ยวเตร่และเรื่องไม่ไ้การเรียนในระดับมาก ($x_{18}(1)$) และเมื่ออ่านหนังสือเสร็จข้าพเจ้าจะไม่บันทึกสรุปสั้น ๆ (x_{27}) พบว่าความน่าจะเป็นไม่สำเร็จการศึกษาเท่ากับ 0.190 ถ้าพยากรณ์นักศึกษาคนหนึ่งที่มีพฤติกรรมเสียเวลาไปกับการเที่ยวเตร่และเรื่องไม่ไ้การเรียนในระดับปานกลาง ($x_{18}(2)$) และเมื่ออ่านหนังสือเสร็จข้าพเจ้าจะไม่บันทึกสรุปสั้น ๆ (x_{27}) ความน่าจะเป็นที่นักศึกษาไม่สำเร็จการศึกษาจะลดลงได้เป็น 0.129 ส่วนการพยากรณ์ความน่าจะเป็นด้วยตัวแบบลอจิสต์ออดส์เชิงสัดส่วน 3 กลุ่ม เช่น นักศึกษาคนหนึ่งเสียเวลาไปกับการเที่ยวเตร่และเรื่องไม่ไ้การเรียนในระดับมาก ($x_{18}(1)$) เมื่ออาจารย์มอบหมายให้ค้นคว้าหรืออ่านหนังสือเพิ่มเติมข้าพเจ้าจะรีบเข้าห้องสมุดทันทีที่มีเวลาว่างในระดับมาก ($x_{21}(1)$) เมื่ออ่านหนังสือ

เสร็จข้าพเจ้ามักจะไม่นับที่สรุปสั้น ๆ (x27) และการทบทวนบทเรียนข้าพเจ้าไม่ใช่วิธีตั้งคำถามให้ตนเองมากกว่า การอ่านหนังสือไปเรื่อย ๆ (x29) พบว่าความน่าจะเป็นไม่สำเร็จการศึกษาเท่ากับ 0.085 ความน่าจะเป็นสำเร็จ การศึกษาตามเกณฑ์เท่ากับ 0.860 และความน่าจะเป็นสำเร็จการศึกษาเกียรตินิยมเท่ากับ 0.055 ตามลำดับ และถ้า พยากรณ์นักศึกษาคนหนึ่งเสียเวลาไปกับการเที่ยวเตร่และเรื่องไม่ใช้การเรียนในระดับปานกลาง (x18(2)) เมื่อ อาจารย์มอบหมายให้ค้นคว้าหรืออ่านหนังสือเพิ่มเติมข้าพเจ้าจะรีบเข้าห้องสมุดทันทีที่มีเวลาว่างจากการเรียนใน ระดับปานกลาง (x21(2)) เมื่ออ่านหนังสือเสร็จข้าพเจ้ามักจะไม่นับที่สรุปสั้น ๆ (x27) และในการทบทวนบทเรียน ข้าพเจ้าไม่ใช่วิธีตั้งคำถามให้ตนเองมากกว่าการอ่านหนังสือไปเรื่อย ๆ (x29) ความน่าจะเป็นไม่สำเร็จการศึกษาจะ เท่ากับ 0.075 ความน่าจะเป็นสำเร็จการศึกษาตามเกณฑ์เท่ากับ 0.865 และความน่าจะเป็นสำเร็จการศึกษาเกียร ตินิยมเท่ากับ 0.06 ตามลำดับ สรุปได้ว่า ถ้านักศึกษามีปัจจัยเสี่ยงที่พบข้างต้นในระดับน้อยลง ความน่าจะเป็นไม่ สำเร็จภายใน 4 ปี จะมีแนวโน้มลดลงด้วย

คำสำคัญ: ปัจจัยเสี่ยงของการศึกษาในมหาวิทยาลัย; การสำเร็จการศึกษา; เกียรตินิยม; ตัวแบบเชิงสถิติ

Abstract

Thai bachelors are currently increasing; however, there are several graduates of failing to graduate within 4 years. This may potentially cause either over budgets or losing opportunity in life. This research investigates the effective risk factors and the probability prediction of level of failing to graduate in 2 cases. Case 1, nominal two grouped responses: fail to graduate and graduate. Case 2, ordinal three grouped responses: fail graduate, graduate, and honor graduate. Risk factors from relevant literature review consist of 30 factors (x1-x30). The sample of 258 fourth-year students were randomly selected using systematic sampling from 516 population units of the fourth-year science students, Silpakorn University in the academic year 2013. The results show that models' goodness-of-fit tests are significant at 0.05 with four risk factors, so that they are suited to use in models for prediction a probability. In forecasting the probability of failing to graduate using the dichotomous logit model, when a student who has level of his loaf time at high level (x18(1)), and when I completed reading book, I didn't do any short summary (x27), is found to be 0.190. Similarly when his loaf time is at moderate level (x18(2)), the probability is decreased to be 0.129. The forecasting probability of failing to graduate using the proportional odds logit model is equal to 0.08 when a student has level of his loaf time at high level (x18(1)), when the lecturer assigned me to research or read more books, I will rush to the library as soon as I have free time at high level (x21(1)), when I completed reading book, I didn't do any short summary (x27) and when reviewing the lessons I didn't have self-questioning rather than only do my reading (x29), which leading to the probability of graduate is 0.86, and the probability of honor graduate is 0.06, respectively. Alternatively, the forecasting probability when

a student has level of his loaf time at moderate level (x18(2)), when the lecturer assigned me to research or read more books, I will rush to the library as soon as I have free time at moderate level (x21(2)), when I completed reading book, I didn't do any short summary (x27) and when reviewing the lessons I didn't have self-questioning rather than only reading (x29), is found to be 0.075, and that the forecasting probability of graduate is 0.865, and the probability of honor graduate is 0.06, respectively. Hence, in conclusion, if a lower level of the risk factors the student can perform, a smaller probability of failing to graduate within 4 years is probably achieved.

Keywords: risk factors in education; university graduate; honor graduate; statistical models

1. บทนำ

ปัจจุบันมีปัญหาการศึกษาของนักศึกษาที่ไม่สำเร็จการศึกษาตามแผนการศึกษาของหลักสูตร ซึ่งจำนวนนักศึกษาที่ไม่สำเร็จการศึกษาตามเวลาในแผนแต่ละปีมีไม่น้อยเมื่อเทียบกับจำนวนนักศึกษาที่เข้าศึกษาในแต่ละปี ทำให้มีนักศึกษาตกค้างปีที่ 5-8 มากขึ้น การไม่สำเร็จการศึกษาอาจส่งผลกระทบต่อในหลาย ๆ ด้าน เช่น ด้านครอบครัว อาจทำให้ครอบครัวผิดหวัง และต้องปรับการวางแผนด้านการเงินตามที่ครอบครัวหวังไว้ ให้เพียงพอต่อรายจ่ายที่อาจเกิดจากการกู้ยืมเพื่อเลี้ยงครอบครัว ทำให้ลำบากในการใช้ชีวิตประจำวัน บางครั้งต้องเป็นภาระของครอบครัวอีกด้วย ด้านจิตใจ ทำให้เกิดอาการเครียด ซึ่งอาการเครียดเหล่านี้มาจากอาการท้อแท้ สิ้นหวังเนื่องจากความฝันของตนที่สร้างไว้ไม่เป็นจริง สามารถทำร้ายร่างกายหรืออาจเป็นอันตรายถึงชีวิตได้ ด้านสังคม บุคคลภายนอกอาจมองนักศึกษาในทางลบ สังคมไม่ยอมรับ ด้านการทำงาน ทำให้ไม่สามารถประกอบอาชีพที่ดีในเวลาอันควรได้ ต้องเป็นลูกจ้างมากกว่าเจ้านาย ใช้แรงงานในการทำงานมากกว่าใช้สติปัญญา อาจไม่สามารถวางแผนและไม่มีควมรับผิดชอบในการจัดการชีวิตตัวเองได้ จึงเป็นการเสียโอกาสทางการศึกษา เช่น ไม่สามารถใช้วิชาความรู้ให้เต็มความสามารถทางด้านการศึกษาอาจน้อยลง [1] ดังนั้น

การที่นักศึกษาไม่สำเร็จการศึกษามีจำนวนมาก อาจทำให้เกิดบุคคลไม่มีคุณภาพมากขึ้น ทำให้การบริหารประเทศไม่มีประสิทธิภาพอีกด้วย ปัจจัยสำคัญที่ส่งผลให้นักศึกษาไม่สำเร็จการศึกษาอาจเกิดจากปัจจัยด้านตัวนักศึกษาเอง เช่น นักศึกษามีผลการเรียนไม่ดี อาจเนื่องมาจากไม่ตั้งใจเรียน ไม่สร้างแรงจูงใจ ไม่อยากเรียน หรือเบื่อหน่ายการเรียน มีปัญหาด้านสุขภาพ หรือปัญหาด้านครอบครัว เช่น ผู้ปกครองมีฐานะยากจน หรือรายได้ต่ำ ผู้ปกครองต้องเดินทางไปประกอบอาชีพที่อื่น ผู้ปกครองต้องการให้นักเรียนประกอบอาชีพหรือหารายได้เอง ครอบครัวแตกแยก ผู้ปกครองมีการศึกษาน้อย บิดามารดาไม่เอาใจใส่ในการเรียนของบุตร และไม่เห็นความสำคัญของการศึกษา [2] ปัจจัยด้านมหาลัยและอาจารย์ เช่น จำนวนอาจารย์ไม่สมดุลกับนักศึกษา อาจารย์ขาดเจตคติที่ดีต่อการสอน มหาวิทยาลัยขาดอุปกรณ์การเรียนการสอน ขาดบริการด้านการแนะแนว และขาดบริการด้านการอำนวยความสะดวกในมหาวิทยาลัย การจัดโปรแกรมด้านการเรียนไม่น่าสนใจ หรือไม่สนองความต้องการของนักศึกษา อาจารย์ขาดการติดตามนักศึกษาที่ขาดเรียนและขาดความเป็นธรรมต่อนักศึกษา ปัจจัยด้านสิ่งแวดล้อม ชุมชน และวัฒนธรรม เช่น การอาศัยอยู่ในชุมชนหนาแน่นมีปัญหาเรื่องการสื่อสารภาษาที่ใช้ในมหาวิทยาลัย เนื่องจากนักศึกษานักเรียนภาษาถิ่นมากกว่า

ภาษาในชุมชน ไม่เห็นความสำคัญของการศึกษา [3] หากปัญหาการเรียนการศึกษาเกิดขึ้นจากหลายปัจจัยพร้อมกัน ไม่ว่าจะเป็นที่ตัวเอง หรือสภาพแวดล้อมการอบรมเลี้ยงดูทั้งที่บ้านและที่สถาบันการศึกษา จะยิ่งทำให้ปัญหาใหญ่ขึ้น และนำไปสู่การแก้ไขปัญหายากยิ่งขึ้น ดังนั้นปัญหาการไม่สำเร็จการศึกษาที่มากขึ้น อาจเป็นอุปสรรคต่อการพัฒนามนุษย์และทุนทางด้านสังคม ประเทศควรจัดการศึกษาให้ทั่วถึง อย่างมีประสิทธิภาพ/คุณภาพ ช่วยให้ประชากรส่วนใหญ่มีความรู้ความสามารถ มีทักษะและเป็นพลเมืองที่มีระเบียบวินัยที่ดี และช่วยกันรับผิดชอบต่อส่วนรวม งานวิจัยนี้ผู้วิจัยสนใจศึกษาปัจจัยเสี่ยงที่ส่งผลกระทบต่อความสำเร็จการศึกษาของนักศึกษา และพยากรณ์ความน่าจะเป็นการไม่สำเร็จการศึกษาตามหลักสูตรภายใน 4 ปี โดยใช้เทคนิคการวิเคราะห์ข้อมูลด้วยตัวแบบลอจิสต์ 2 กลุ่ม (dichotomous logit model) [4] และตัวแบบลอจิสต์ออดส์เชิงสัดส่วน (proportional

odds logit model) 3 กลุ่ม [5] ซึ่งตัวแบบที่สร้างมีการตรวจสอบภาวะสารูปดีด้วยตัวสถิติ (goodness-of-fit statistics) คือ ตัวสถิติ likelihood ratio, ตัวสถิติ score และตัวสถิติ Wald [6] โดยมีขอบเขตของตัวแปร ดังนี้

1.1 กรณีที่ 1 การวิเคราะห์ข้อมูลโดยใช้ตัวแบบลอจิสต์ เมื่อตัวแปรตอบสนอง (Y) คือ การสำเร็จการศึกษาแบบ 2 กลุ่ม ได้แก่ 1 แทนไม่สำเร็จการศึกษาตามหลักสูตร และ 2 แทนสำเร็จการศึกษาตามหลักสูตร

1.2 กรณีที่ 2 การวิเคราะห์ข้อมูลโดยใช้ตัวแบบลอจิสต์เชิงสัดส่วน เมื่อตัวแปรตอบสนอง (Y) คือ การสำเร็จการศึกษาแบบ 3 กลุ่ม ได้แก่ 1 แทนการไม่สำเร็จการศึกษาตามหลักสูตร 2 แทนการสำเร็จการศึกษาตามหลักสูตร และ 3 แทนการสำเร็จการศึกษาตามหลักสูตรเกียรตินิยม

ปัจจัย (ตัวแปรอธิบาย)	ปัจจัย (ตัวแปรอธิบาย)	ปัจจัย (ตัวแปรอธิบาย)
1. สาขาวิชาเอก(x1)	16. ข้าพเจ้ามีการวางแผน ทำตารางเวลาเพื่อทบทวนบทเรียนสม่ำเสมอและดูหนังสือในระดับใด (x16)	22.ข้าพเจ้าสนใจเรื่องคะแนนสอบในระดับใด (x22)
2. เพศ(x2)	17.ข้าพเจ้าเสียเวลาไปกับการเที่ยวเตร่ และเรื่องไม่ใช้การเรียนในระดับใด (x17)	23.ข้าพเจ้ามีความมุ่งมั่นที่จะเรียนให้ได้ผลดีที่สุดในระดับใด(x23)
3. ประเภทโรงเรียนของมัธยมศึกษาตอนปลาย(x3)	18.ข้าพเจ้าเสียเวลาไปกับการเที่ยวเตร่ และเรื่องไม่ใช้การเรียนในระดับใด (x18)	24.ข้าพเจ้ามีความมุ่งมั่นที่จะเรียนให้ได้ผลดีที่สุดในระดับใด(x24)
4. เกรตสะสมมัธยมศึกษาตอนปลาย(x4)	19.เมื่อข้าพเจ้าทำรายงานหรือทำการบ้านไม่เสร็จข้าพเจ้าจะลอกจากเพื่อนเพื่อให้มีงานส่งอาจารย์ในระดับใด(x19)	25.ในแต่ละปีการศึกษาท่านเคยเรียนพิเศษเกี่ยวกับบทเรียนในหลักสูตร(x25)
5. วิธีการรับเข้า(x5)	20.ข้าพเจ้าตั้งใจเรียนและทบทวนบทเรียนอย่างสม่ำเสมอในระดับใด (x20)	26.ในการทำข้อสอบข้าพเจ้าจะอ่านข้อความทั้งหมดอย่างใช้ความคิดที่จะลงมือเขียนตอบในระดับใด(x26)
6. สถานภาพสมรสของบิดามารดา(x6)	21.เมื่ออาจารย์มอบหมายให้ค้นคว้าหรืออ่านหนังสือเพิ่มเติมข้าพเจ้าจะรีบเข้าห้องสมุดทันทีที่มีเวลาว่างจากการเรียนในระดับใด(x21)	27.เมื่ออ่านหนังสือเสร็จข้าพเจ้ามักจะบันทึกสรุปสั้นๆ(x27)
7. อาชีพของบิดา(x7)		28.ข้าพเจ้าไม่ค่อยจดคำบรรยายของอาจารย์ (x28)
8. อาชีพของมารดา(x8)		29.ในการทบทวนบทเรียนข้าพเจ้าใช้วิธีตั้งคำถามให้ตนเองมากกว่าการอ่านหนังสือไปเรื่อยๆ(x29)
9. ท่านได้รับอุปการะด้านการเงินจากบิดา หรือมารดา(x9)		30.ในการสอบข้าพเจ้าจะแบ่งเวลาในตอนท้ายของการสอบไว้เพื่อตรวจสอบคำตอบก่อนที่จะส่งอาจารย์(x30)
10. ท่านได้รับอุปการะด้านการเงินจากบิดา และมารดา(x10)		
11. ท่านได้รับอุปการะด้านการเงินส่วนใหญ่จาก (x11)		
12. รายได้ของผู้ปกครอง/เดือน(x12)		
13. รายได้ของนักศึกษา/เดือน(x13)		
14. ข้าพเจ้ารับผิดชอบต่อการเรียนเช่น อ่านหนังสือค้นคว้า ทำการบ้าน ด้วยความเต็มใจโดยไม่มีใครบังคับ ในระดับใด(x14)		
15. ข้าพเจ้าขาดเรียนด้วยเหตุผลที่ไม่จำเป็นนัก เช่น นอนตื่นสาย ทำการบ้านหรือรายงานไม่เสร็จในระดับใด(x15)		

2. วิธีการทดลองและขอบเขตของการวิจัย

ข้อมูลที่น่ามาวิจัยเป็นข้อมูลจริงจากการสำรวจช่วงเดือนพฤศจิกายน - ธันวาคม 2556 ของนักศึกษาคณะวิทยาศาสตร์ ชั้นปีที่ 4 มหาวิทยาลัยศิลปากร ปีการศึกษา 2556 จากประชากร 516 คน โดยใช้ตัวอย่างทั้งหมด 258 คน ภายใต้แผนแบบการเลือกตัวอย่างแบบมีระบบ (systematic sampling design) วิธีการทดลองเพื่อให้มีการนำไปใช้และการตรวจสอบความแม่นยำของตัวแบบ ผู้วิจัยได้แบ่งข้อมูลแบบสุ่มออกเป็น 2 ส่วน ส่วนแรกประมาณร้อยละ 80 ขนาดตัวอย่าง 206 คน ใช้เป็นข้อมูลในการทดลอง (training set data) ส่วนหลังประมาณร้อยละ 20 ขนาดตัวอย่าง 52 คน ใช้เป็นข้อมูลในการตรวจสอบความถูกต้องของตัวแบบ (validation set data)

ขั้นตอนการวิเคราะห์ข้อมูล ตัวสถิติ และตัวแบบของการวิจัย ดังนี้

(1) สร้างตัวแบบของการวิจัย ด้วยวิธีการคัดเลือกตัวแปรแบบขั้นตอน (stepwise likelihoodratio, LR) ภายใต้การประมาณพารามิเตอร์วิธีภาวะน่าจะเป็นสูงสุดแบบวนซ้ำ (iterative maximum likelihood or ML) [7] และทดสอบสมมติฐานเชิงสถิติ ด้วยตัวสถิติทดสอบคือ likelihood ratio test, score test และตัวสถิติ Wald test

(2) วัดความเกี่ยวพันด้วยตัวสถิติเช่น Somers'D ตัวสถิติ gamma และตัวสถิติ c จากนั้นตรวจสอบความเหมาะสมของตัวแบบ ด้วยตัวสถิติอัตราส่วนภาวะน่าจะเป็น (G^2) หรือ deviance รวมถึงค่าสถิติอื่น ๆ ประกอบด้วย

(3) ตรวจสอบผลการพยากรณ์กลุ่มได้ถูกต้องด้วยตารางการจำแนกกลุ่ม (classification table) ของชุดข้อมูลในการทดลอง

(4) ตรวจสอบความถูกต้องของการพยากรณ์กลุ่มด้วยตารางการจำแนกกลุ่มของชุดข้อมูลในการตรวจสอบความถูกต้องของตัวแบบ

ตัวแบบในงานวิจัย การวิเคราะห์ข้อมูลเชิงสถิติอาศัยตัวแบบเชิงสถิติ 2 ตัวแบบ ดังนี้

(1) **ตัวแบบลอจิต 2 กลุ่ม** (dichotomous logit model) เมื่อตัวแปรตอบสนอง (Y) มี 2 กลุ่มและปัจจัยเสี่ยงหรือตัวแปรอธิบายหลายตัว (x 's) [4] ดังในสมการ (1)

$$Q(x) = \log\left[\frac{P(x)}{1-P(x)}\right] = \log\left(\frac{P(Y=1|x)}{1-P(Y=1|x)}\right) = \beta_0 + \beta'x \quad \dots (1)$$

เมื่อ β_0 , β แทนพารามิเตอร์จุดตัด (intercept) และเวกเตอร์ความชัน (slopes) ในตัวแบบตามลำดับ

(2) **ตัวแบบลอจิตออดส์เชิงสัดส่วน** (proportional odds logit model) เมื่อตัวแปรตอบสนองมี 3 กลุ่ม และตัวแปรอธิบายหลายตัว (x 's) ซึ่งอาจเรียกอีกชื่อหนึ่งว่า ตัวแบบลอจิต 3 กลุ่ม (trichotomous logit model) [8] ดังในสมการ (2)

$$\log\left(\frac{P(Y \leq j|x)}{1-P(Y \leq j|x)}\right) = \log\left(\frac{P_1(x) + \dots + P_j(x)}{P_{j+1} + \dots + P_J(x)}\right),$$

$$j = 1, \dots, J-1 = \text{logit}[P(Y \leq j|x)] = \beta_{j0} + \beta'x \quad \dots (2)$$

เมื่อ β_0 , β แทนพารามิเตอร์จุดตัด (intercept) และเวกเตอร์ความชัน (slope) ในตัวแบบตามลำดับ

ตัวแบบนี้ขึ้นอยู่กับข้อตกลงเบื้องต้นของเวกเตอร์ความชันในตัวแบบ (2) ที่ส่งผลต่อแต่ละระดับหรือกลุ่มของ Y_j , $j=1, \dots, J$ นั้นเท่ากัน กล่าวคือเวกเตอร์ $\beta_j = \beta$

การสร้างตัวแบบ (1) และ (2) ด้วยวิธีคัดเลือกแบบขั้นตอน โดยทำการวิเคราะห์ความเหมาะสมของตัวแบบ และการทดสอบเกี่ยวกับปัจจัยเสี่ยงต่าง ๆ รวมทั้งการตรวจสอบข้อตกลงเบื้องต้นของตัวแบบ (2) เกี่ยวกับ β ว่าส่งผลต่อทุกระดับของ Y เท่ากันหรือไม่ ถ้าใช่ ตัวแบบที่ใช้มีความเหมาะสม เพื่อการพยากรณ์ค่าความน่าจะเป็นต่อไป [9]

งานวิจัยนี้ประมวลผลข้อมูลและวิเคราะห์ข้อมูลด้วยโปรแกรมสำเร็จรูปชั้นสูง 2 โปรแกรม คือ SPSS version 18 ซึ่งใช้ในการเตรียมข้อมูลและตรวจสอบข้อมูล และ SAS Enterprise Guide 5.1 ซึ่งใช้ในการสร้างตัวแบบวิเคราะห์ตัวแบบและการทดสอบสมมติฐานต่าง ๆ ภายใต้ลิขสิทธิ์ของภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยศิลปากร วิทยาเขตพระราชวังสนามจันทร์

การศึกษาตามหลักสูตรภายใน 4 ปีอย่างมีนัยสำคัญที่ 0.05 (P-value = 0.0121, 0.0077) ตามลำดับ

ตารางที่ 1 ตัวแปรอธิบายที่มีผลกระทบต่อตัวแปรตอบสนอง 2 กลุ่ม

Effect	DF	Wald Chi-Square	P-value
x18	2	8.8323	0.0121
x27	1	7.0957	0.0077

3.ผลการวิจัย

3.1 กรณีที่ 1 การวิเคราะห์ข้อมูลโดยใช้ตัวแบบลอจิสต์เมื่อตัวแปรตอบสนองมี 2 กลุ่ม

จากตารางที่ 1 พบว่าปัจจัยเสี่ยงที่ถูกคัดเลือกเข้าตัวแบบลอจิสต์ 2 กลุ่ม ด้วยวิธีการเลือกตัวแปรแบบขั้นตอน ได้แก่ ข้าพเจ้าเสียเวลาไปกับการเที่ยวเตร่และเรื่องไม่เีการเรียนในระดับใด (x18) และเมื่ออ่านหนังสือเสร็จข้าพเจ้ามักจะบันทึกสรุปสั้น ๆ (x27) เป็นตัวแปรที่มีความเกี่ยวข้องกับการไม่สำเร็จ

จากตารางที่ 2 ผลการคัดเลือกตัวแปรแบบขั้นตอน พบว่าตัวแปรที่ถูกคัดเลือกคัดเลือกเข้าสมการคือ ข้าพเจ้าเสียเวลาไปกับการเที่ยวเตร่และเรื่องไม่เีการเรียนในระดับใด (x18) และเมื่ออ่านหนังสือเสร็จข้าพเจ้ามักจะบันทึกสรุปสั้น ๆ (x27) โดยแสดงกลุ่มหรือระดับของตัวแปรที่มีนัยสำคัญเชิงสถิติที่ 0.05 ได้แก่ x18(1) ในระดับมาก (P-value = 0.0181) และ x27(1) ในระดับไม่ได้บันทึกสรุปสั้น ๆ (P-value = 0.0077)

ตารางที่ 2 ค่าประมาณพารามิเตอร์ของตัวแบบลอจิสต์ 2 กลุ่ม เมื่อตัวแปรตอบสนองแบ่งเป็น 2 กลุ่มไม่มีอันดับ

พารามิเตอร์ (β)	องศาอิสระ	$\hat{\beta}$	S.E. β	ค่าสถิติ Wald	P-value
Intercept	1	-1.8108	0.2968	37.2297	<0.0001
x18 (1)	1	0.7475	0.3163	5.5854	0.0181
x18 (2)	1	0.5066	0.3337	2.3054	0.1289
x27 (1)	1	-0.6006	0.2255	7.0957	0.0077

3.1.1 การสร้างตัวแบบลอจิสต์ 2 กลุ่ม

จากตารางที่ 2 ได้ตัวแบบลอจิสต์ 2 กลุ่มดังในสมการ (3) คือ

$$\log \frac{P(x)}{1-P(x)} = -1.8108 + 0.8163 \times x18(1) + 0.5066 \times x18(2) - 0.6006 \times x27(1) \quad \dots(3)$$

$$\text{เมื่อ } P(Y = 1 | x) = P(x) = \frac{\exp(\hat{\phi})}{1 + \exp(\hat{\phi})}$$

$$\text{และ Odds} = \frac{P(x)}{1-P(x)} = \exp(\hat{\phi})$$

$$\text{เมื่อ } (\hat{\phi}) = -1.8108 + 0.8163 \times x18(1) + 0.5066 \times x18(2) - 0.6006 \times x27(1)$$

3.1.2 ตัวอย่างการพยากรณ์ความน่าจะเป็นในกรณี 1 ด้วยตัวแบบลอจิสต์ 2 กลุ่ม

ตัวอย่างการพยากรณ์ที่ 1.1

สมมติต้องการพยากรณ์นักเรียนคนหนึ่งที่มีพฤติกรรมเสียเวลาไปกับการเที่ยวเตร่และเรื่องไม่ไ้ การเรียนในระดับมาก หรือ $x_{18}(1)$ และเมื่ออ่านหนังสือเสร็จข้าพเจ้าจะไม่บันทึกสรุปสั้น ๆ หรือ $x_{27}(1)$ โดยกำหนด $x_{18}(1) = 1$ และ $x_{27}(1) = 1$ เมื่อตัวแปรอื่น ๆ คงตัวและเท่ากับ 0 จะคำนวณค่าความน่าจะเป็นได้ ดังนั้น

$$P(x) = \frac{\exp(-1.60)}{1 + \exp(-1.60)} = 0.17$$

หมายความว่าความน่าจะเป็นที่นักศึกษาไม่สำเร็จการศึกษาภายใน 4 ปี ซึ่งมีพฤติกรรมเสียเวลาไปกับการเที่ยวเตร่และเรื่องไม่ไ้ การเรียนในระดับมากหรือ $x_{18}(1)$ และเมื่ออ่านหนังสือเสร็จข้าพเจ้าจะไม่บันทึกสรุปสั้น ๆ หรือ $x_{27}(1)$ เท่ากับ 0.17 หรือประมาณร้อยละ 17 ซึ่งพบว่าค่อนข้างสูง เนื่องจากในทางกลับกันนักศึกษาคงกล่าวจะมีโอกาสสำเร็จการศึกษาภายใน 4 ปี เพียง 0.83 หรือประมาณร้อยละ 83 ซึ่งไม่ใช่ร้อยละ 100 ตามควรหรือตามหลักสูตร กล่าวอีกนัยหนึ่งนักศึกษาอาจต้องเรียนตกค้างอยู่ในชั้นปีที่ 5-8 ด้วยความน่าจะเป็นประมาณร้อยละ 17

ตัวอย่างการพยากรณ์ที่ 1.2

สมมติต้องการพยากรณ์นักเรียนคนหนึ่งที่มีพฤติกรรมเสียเวลาไปกับการเที่ยวเตร่และเรื่องไม่ไ้ การเรียนในระดับปานกลางหรือ $x_{18}(2)$ และเมื่ออ่านหนังสือเสร็จข้าพเจ้าจะไม่บันทึกสรุปสั้น ๆ หรือ $x_{27}(1)$

โดยกำหนด $x_{18}(2) = 1$ และ $x_{27}(1) = 1$ เมื่อตัวแปรอื่น ๆ คงตัวและเท่ากับ 0 จะคำนวณค่าความน่าจะเป็นได้ ดังนั้น

$$P(x) = \frac{\exp(-1.90)}{1 + \exp(-1.90)} = 0.13$$

หมายความว่า ความน่าจะเป็นที่นักศึกษาไม่สำเร็จการศึกษาภายใน 4 ปี ซึ่งมีพฤติกรรมเสียเวลาไปกับการเที่ยวเตร่และเรื่องไม่ไ้ การเรียนในระดับปานกลางหรือ $x_{18}(2)$ และเมื่ออ่านหนังสือเสร็จข้าพเจ้าจะไม่บันทึกสรุปสั้น ๆ หรือ $x_{27}(1)$ เท่ากับ 0.13 หรือ

ประมาณ ร้อยละ 13 ซึ่งพบว่าย่ำค่อนข้างสูง เนื่องจากในทางกลับกันนักศึกษาคงกล่าวจะมีโอกาสสำเร็จการศึกษาภายใน 4 ปี เพียง 0.87 หรือประมาณร้อยละ 87 ซึ่งไม่ใช่ ร้อยละ 100 ตามควรหรือตามหลักสูตร หรือนักศึกษาอาจต้องเรียนตกค้างอยู่ในชั้นปีที่ 5-8 ด้วยความน่าจะเป็นประมาณร้อยละ 13

สรุปในกรณีที่ 1 จะเห็นจากตัวอย่างการพยากรณ์ที่ 1.1-1.2 ว่า ถ้านักศึกษาคงกล่าวมีพฤติกรรมการเรียนดีขึ้นบ้าง ความน่าจะเป็นที่นักศึกษาจะไม่สำเร็จการศึกษาภายใน 4 ปี จะลดลงจากที่เคยมีร้อยละ 17 ลดลงเหลือร้อยละ 13 นอกจากนี้ประสิทธิภาพในการพยากรณ์ภาพรวมของตัวแบบลอจิสต์ 2 กลุ่มด้วยตารางการจำแนกกลุ่ม (classification table) แสดงไว้ในตารางที่ 3

ตารางที่ 3 ประสิทธิภาพในการพยากรณ์ภาพรวมของตัวแบบลอจิสต์ 2 กลุ่ม

ค่าสังเกต (Observed)		ค่าพยากรณ์ (Predicted)		
		Y		Percentage correct
		1	2	
Y	1	30	7	81.08
	2	31	190	86.36
ภาพรวม (Overall)				85.27

จากตารางที่ 3 แสดงประสิทธิภาพในการพยากรณ์ของตัวแบบลอจิสต์ พบว่าประสิทธิภาพในการพยากรณ์ที่เป็นส่วน ๆ หรือในแต่ละระดับของการไม่สำเร็จการศึกษาเช่น จะพยากรณ์การไม่สำเร็จการศึกษา (กลุ่ม 1) ได้ถูกต้องเท่ากับร้อยละ 81.08 และจะการพยากรณ์การสำเร็จการศึกษา (กลุ่ม 2) ได้ถูกต้องเท่ากับร้อยละ 86.36

ดังนั้นตัวแบบลอจิสต์ 2 กลุ่ม มีประสิทธิภาพในภาพรวมของการพยากรณ์กลุ่มหรือระดับการสำเร็จ

การศึกษา โดยสามารถพยากรณ์ได้ถูกต้องค่อนข้างสูง คือร้อยละ 85.27

3.2 กรณีที่ 2 การวิเคราะห์ข้อมูลโดยใช้ตัวแบบลอจิสติกส์เชิงสัดส่วน เมื่อตัวแปรตอบสนองมีอันดับ 3 กลุ่ม

ค่าประมาณต่าง ๆ และผลการทดสอบเกี่ยวกับพารามิเตอร์ของตัวแบบลอจิสติกส์เชิงสัดส่วน 3 กลุ่ม แสดงไว้ในตารางที่ 4

จากตารางที่ 4 พบว่าตัวแปรปัจจัยเสี่ยงที่เป็นตัวแปรอธิบายที่ถูกคัดเลือกเข้าตัวแบบลอจิสติกส์เชิงสัดส่วน ด้วยวิธีการเลือกตัวแปรแบบขั้นตอน ได้แก่ ข้าพเจ้าเสียเวลาไปกับการเที่ยวเตร่และเรื่องไม่ไ้การเรียนในระดับใด (x18) (P-value = 0.0201) เมื่ออาจารย์มอบหมายให้ค้นคว้าหรืออ่านหนังสือเพิ่มเติมข้าพเจ้าจะรีบเข้าห้องสมุดทันทีที่มีเวลาว่างจากการเรียนในระดับใด (x21) (P-value = 0.0459)

เมื่ออ่านหนังสือเสร็จข้าพเจ้ามักจะบันทึกสรุปสั้น ๆ (x27) (P-value = 0.0008) และในการทบทวนบทเรียนข้าพเจ้าใช้วิธีตั้งคำถามให้ตนเองมากกว่าการอ่านหนังสือไปเรื่อย ๆ (x29) (P-value = 0.0210) ซึ่งล้วนเป็นปัจจัยเสี่ยงหรือตัวแปรอธิบายที่ส่งผลกระทบต่อความสำเร็จการศึกษาภายใน 4 ปี อย่างมีนัยสำคัญทางสถิติที่ 0.05

ตารางที่ 4 ตัวแปรอธิบายที่มีความเกี่ยวข้องกับตัวแปรตอบสนองมีอันดับ 3 กลุ่ม

อิทธิพล	องศาอิสระ	ค่าสถิติ score	P-value
x18	2	7.8104	0.0201
x21	2	6.1637	0.0459
x27	1	11.136	0.0008
x29	1	5.3302	0.0210

ตารางที่ 5 ค่าประมาณพารามิเตอร์ของตัวแบบลอจิสติกส์เชิงสัดส่วน เมื่อตัวแปรตอบสนองมีอันดับ 3 กลุ่ม

พารามิเตอร์ (β)	องศาอิสระ	$\hat{\beta}$	S.E. ($\hat{\beta}$)	ค่าสถิติ Wald	P-value
Intercept	1	-1.5633	0.3346	21.8309	<0.0001
Intercept	1	3.5646	0.4427	64.8272	<0.0001
x18 (1)	1	0.9571	0.4445	4.6361	0.0313
x18 (2)	1	-0.0708	0.2795	0.0642	0.8
x21 (1)	1	0.3063	0.4929	0.3861	0.0344
x21 (2)	1	0.3606	0.3013	1.4324	0.2314
x27 (1)	1	-0.7209	0.216	11.136	0.0008
x29 (1)	1	-0.4384	0.1899	5.3302	0.021

จากตารางที่ 5 พบว่าค่าประมาณสัมประสิทธิ์ในตัวแบบลอจิสติกส์เชิงสัดส่วนของตัวแปรอธิบาย 4 ตัวที่มีส่งผลกระทบต่อความสำเร็จการศึกษาภายใน 4 ปี อย่างมีนัยสำคัญทางสถิติที่ 0.05 ประกอบด้วยระดับหรือกลุ่มที่ถูกคัดเลือกเข้าสมการ

อย่างมีนัยสำคัญทางสถิติที่ 0.05 ได้แก่ x18 (1) ในระดับมาก x21 (1) ในระดับปานกลาง x27 (1) ในระดับไม่ใช่ และ x29 (1) ในระดับไม่ใช่ ซึ่งแต่ละระดับให้ค่า P-value เท่ากับ 0.0313, 0.0344, 0.0008 และ 0.021 ตามลำดับ

นอกจากนี้ ก่อนนำตัวแบบลอจิสติกส์เชิงสัดส่วนไปใช้ ควรตรวจสอบข้อตกลงเบื้องต้นของตัวแบบเกี่ยวกับพารามิเตอร์ของตัวแบบว่าจะส่งผลต่อทุกระดับของ Y เท่ากันหรือไม่ ถ้าใช่หรือเท่ากัน ตัวแบบที่ใช้จึงมีความเหมาะสม และพบว่าตัวแบบอยู่ในเกณฑ์เหมาะสม (P-value = 0.0593)

3.2.1 การสร้างตัวแบบลอจิสติกส์เชิงสัดส่วน

จากตารางที่ 5 ได้ตัวแบบลอจิสติกส์เชิงสัดส่วนหรือตัวแบบ 3 กลุ่ม ดังในสมการ (4) - (5) ดังนี้

$$\log \left(\frac{P(Y \leq j = 1 | x)}{1 - [P(Y \leq j = 1 | x)]} \right) =$$

$$\text{logit} [P(Y \leq 1 | x)] = -1.5633 + 0.9571 \times 18(1) - 0.0708 \times 18(2) - 0.3063 \times 21(1) + 0.3606 \times 21(2) - 0.7209 \times 27(1) - 0.4384 \times 29(1) \quad \dots(4)$$

และ

$$\log \left(\frac{P(Y \leq j = 2 | x)}{1 - [P(Y \leq j = 2 | x)]} \right) =$$

$$\text{logit} [P(Y \leq 2 | x)] = 3.5646 + 0.9571 \times 18(1) - 0.0708 \times 18(2) - 0.3063 \times 21(1) + 0.3606 \times 21(2) - 0.7209 \times 27(1) - 0.4384 \times 29(1) \quad \dots(5)$$

3.2.2 ตัวอย่างการพยากรณ์ความน่าจะเป็นในกรณีที่ 2 ด้วยตัวแบบลอจิสติกส์เชิงสัดส่วน 3 กลุ่มมีอันดับ ประกอบด้วย

(1) ความน่าจะเป็นที่นักศึกษา ไม่สำเร็จ การศึกษาภายใน 4 ปี คือ

$$P(Y=1) = P(Y \leq 1) = \frac{\exp(\hat{\phi}_1)}{1 + \exp(\hat{\phi}_1)}$$

เมื่อ $\hat{\phi}_1 = -1.5633 + 0.9571 \times 18(1) - 0.0708 \times 18(2) - 0.3063 \times 21(1) + 0.3606 \times 21(2) - 0.7209 \times 27(1) - 0.4384 \times 29(1)$

(2) ความน่าจะเป็นที่นักศึกษา สำเร็จ การศึกษาภายใน 4 ปี คือ

$$P(Y=2) = P(Y \leq 2) - P(Y \leq 1)$$

$$= \frac{\exp(\hat{\phi}_2)}{1 + \exp(\hat{\phi}_2)} - \frac{\exp(\hat{\phi}_1)}{1 + \exp(\hat{\phi}_1)}$$

เมื่อ $\hat{\phi}_2 = 3.5646 + 0.9571 \times 18(1) - 0.0708 \times 18(2) - 0.3063 \times 21(1) + 0.3606 \times 21(2) - 0.7209 \times 27(1) - 0.4384 \times 29(1)$

(3) ความน่าจะเป็นที่นักศึกษาสำเร็จ การศึกษาเกียรตินิยม คือ

$$P(Y=3) = 1 - P(Y \leq 2)$$

3.2.3 ตัวอย่างการพยากรณ์ที่ 2.1

สมมติ ถ้าพยากรณ์นักศึกษาคนหนึ่งเสียเวลาไปกับการเที่ยวเตร่และเรื่องไม่ไ้การเรียนในระดับมาก (x18(1)) เมื่ออาจารย์มอบหมายให้ค้นคว้าหรืออ่านหนังสือเพิ่มเติมข้าพเจ้าจะรีบเข้าห้องสมุดทันทีที่มีเวลาว่างจากการเรียนในระดับมาก (x21(1)) เมื่ออ่านหนังสือเสร็จข้าพเจ้าจะไม่บันทึกสรุปสั้น ๆ (x27) และในการทบทวนบทเรียนข้าพเจ้าไม่ใช้วิธีตั้งคำถามให้ตนเองมากกว่าการอ่านหนังสือไปเรื่อยๆ (x29) เมื่อตัวแปรอื่น ๆ คงตัว โดยที่ $\hat{\phi}_1 = -1.46$ ดังนี้

$$P(Y \leq 1) = \frac{\exp(-1.46)}{1 + \exp(-1.46)} = 0.18$$

หมายความว่าความน่าจะเป็นที่นักศึกษาไม่สำเร็จ การศึกษาภายใน 4 ปี เมื่อนักศึกษาเสียเวลาไปกับการเที่ยวเตร่และเรื่องไม่ไ้การเรียนในระดับมาก (x18(1)) เมื่ออาจารย์มอบหมายให้ค้นคว้าหรืออ่านหนังสือเพิ่มเติมข้าพเจ้าจะรีบเข้าห้องสมุดทันทีที่มีเวลาว่างจากการเรียนในระดับมาก (x21(1)) เมื่ออ่านหนังสือเสร็จข้าพเจ้าจะไม่บันทึกสรุปสั้น ๆ (x27) และในการทบทวนบทเรียนข้าพเจ้าไม่ใช้วิธีตั้งคำถามให้ตนเองมากกว่าการอ่านหนังสือไปเรื่อยๆ (x29) เท่ากับ 0.18 หรือร้อยละ 18

$$P(Y=2) = \frac{\exp(3.67)}{1 + \exp(3.67)} - 0.18 = 0.80$$

เมื่อ $\hat{\phi}_2 = 3.67$ หมายความว่าความน่าจะเป็นที่นักศึกษาสำเร็จการศึกษาตามเกณฑ์ภายใน 4 ปี เมื่อ

นักศึกษาเสียเวลาไปกับการเที่ยวเตร่และเรื่องไม่ใช้การเรียนในระดับมาก(x18(1)) เมื่ออาจารย์มอบหมายให้ค้นคว้าหรืออ่านหนังสือเพิ่มเติมข้าพเจ้าจะรีบเข้าห้องสมุดทันทีที่มีเวลาวางจากการเรียนในระดับมาก(x21(1)) เมื่ออ่านหนังสือเสร็จข้าพเจ้ามักจะไม่นั่งทักสรุปล้น ๆ (x27) และในการทบทวนบทเรียนข้าพเจ้าไม่ใช่วิธีตั้งคำถามให้ตนเองมากกว่าการอ่านหนังสือไปเรื่อย ๆ (x29) เท่ากับ 0.80หรือร้อยละ 80

$$P(Y=3) = 1 - P(Y \leq 2) = 1 - 0.98 = 0.02$$

หมายความว่าความน่าจะเป็นที่นักศึกษาสำเร็จการศึกษาเกียรตินิยม เมื่อนักศึกษาเสียเวลาไปกับการเที่ยวเตร่และเรื่องไม่ใช้การเรียนในระดับมาก(x18(1)) เมื่ออาจารย์มอบหมายให้ค้นคว้าหรืออ่านหนังสือเพิ่มเติมข้าพเจ้าจะรีบเข้าห้องสมุดทันทีที่มีเวลาวางจากการเรียนในระดับมาก (x21(1)) เมื่ออ่านหนังสือเสร็จข้าพเจ้ามักจะไม่นั่งทักสรุปล้น ๆ (x27) และในการทบทวนบทเรียนข้าพเจ้าไม่ใช่วิธีตั้งคำถามให้ตนเองมากกว่าการอ่านหนังสือไปเรื่อย ๆ (x29) เท่ากับ 0.02 หรือร้อยละ 2

3.2.4 ตัวอย่างการพยากรณ์ที่ 2.2

สมมติ ถ้าพยากรณ์นักศึกษาคนหนึ่งเสียเวลาไปกับการเที่ยวเตร่และเรื่องไม่ใช้การเรียนในระดับปานกลาง (x18(2)) เมื่ออาจารย์มอบหมายให้ค้นคว้าหรืออ่านหนังสือเพิ่มเติมข้าพเจ้าจะรีบเข้าห้องสมุดทันทีที่มีเวลาวางจากการเรียนในระดับปานกลาง (x21(2)) เมื่ออ่านหนังสือเสร็จข้าพเจ้ามักจะไม่นั่งทักสรุปล้น ๆ (x27) และในการทบทวนบทเรียนข้าพเจ้าไม่ใช่วิธีตั้งคำถามให้ตนเองมากกว่าการอ่านหนังสือไปเรื่อย ๆ (x29) เมื่อตัวแปรอื่น ๆ คงตัว โดยที่ $\hat{\phi}_1 = -2.43$ ดังนั้น

$$P(Y \leq 1) = \frac{\exp(-2.43)}{1 + \exp(-2.43)} = 0.08$$

หมายความว่า ความน่าจะเป็นที่นักศึกษาไม่สำเร็จการศึกษายภายใน 4 ปี เมื่อนักศึกษาเสียเวลาไปกับการ

เที่ยวเตร่และเรื่องไม่ใช้การเรียนในระดับปานกลาง (x18(2)) เมื่ออาจารย์มอบหมายให้ค้นคว้าหรืออ่านหนังสือเพิ่มเติมข้าพเจ้าจะรีบเข้าห้องสมุดทันทีที่มีเวลาวางจากการเรียนในระดับปานกลาง (x21(2)) เมื่ออ่านหนังสือเสร็จข้าพเจ้ามักจะไม่นั่งทักสรุปล้น ๆ (x27) และในการทบทวนบทเรียนข้าพเจ้าไม่ใช่วิธีตั้งคำถามให้ตนเองมากกว่าการอ่านหนังสือไปเรื่อย ๆ (x29) เท่ากับ 0.08 หรือร้อยละ 8

$$P(Y=2) = \frac{\exp(2.695)}{1 + \exp(2.695)} - 0.08 = 0.86$$

โดยที่ $\hat{\phi}_2 = 2.695$ หมายความว่า ความน่าจะเป็นที่นักศึกษาสำเร็จการศึกษาตามเกณฑ์ภายใน 4 ปีเมื่อนักศึกษาเสียเวลาไปกับการเที่ยวเตร่และเรื่องไม่ใช้การเรียนในระดับปานกลาง(x18)เมื่ออาจารย์มอบหมายให้ค้นคว้าหรืออ่านหนังสือเพิ่มเติมข้าพเจ้าจะรีบเข้าห้องสมุดทันทีที่มีเวลาวางจากการเรียนในระดับปานกลาง(x21)เมื่ออ่านหนังสือเสร็จข้าพเจ้ามักจะไม่นั่งทักสรุปล้น ๆ (x27) และในการทบทวนบทเรียนข้าพเจ้าไม่ใช่วิธีตั้งคำถามให้ตนเองมากกว่าการอ่านหนังสือไปเรื่อย ๆ (x29) เท่ากับ 0.86 หรือร้อยละ 86 และ

$$P(Y=3) = 1 - P(Y \leq 2) = 1 - 0.94 = 0.06$$

หมายความว่า ความน่าจะเป็นที่นักศึกษาสำเร็จการศึกษาเกียรตินิยม เมื่อนักศึกษาเสียเวลาไปกับการเที่ยวเตร่และเรื่องไม่ใช้การเรียนในระดับปานกลาง (x18(2)) เมื่ออาจารย์มอบหมายให้ค้นคว้าหรืออ่านหนังสือเพิ่มเติมข้าพเจ้าจะรีบเข้าห้องสมุดทันทีที่มีเวลาวางจากการเรียนในระดับปานกลาง (x21(2)) เมื่ออ่านหนังสือเสร็จข้าพเจ้ามักจะไม่นั่งทักสรุปล้น ๆ (x27) และในการทบทวนบทเรียนข้าพเจ้าไม่ใช่วิธีตั้งคำถามให้ตนเองมากกว่าการอ่านหนังสือไปเรื่อย ๆ (x29) เท่ากับ 0.06 หรือร้อยละ 6

สรุปในกรณีนี้ 2 จะเห็นจากตัวอย่างของการพยากรณ์ที่ 2.1-2.2 ว่า ถ้านักศึกษาค้นดังกล่าวมีพฤติกรรมเรียนขยันขันดีขึ้นบ้าง ความ

น่าจะเป็นที่นักศึกษาจะไม่สำเร็จการศึกษาภายใน 4 ปี จะลดลง ส่วนความน่าจะเป็นที่นักศึกษาจะสำเร็จการศึกษาภายใน 4 ปีและความน่าจะเป็นที่นักศึกษาสำเร็จการศึกษาเกียรตินิยมจะเพิ่มขึ้น

นอกจากนี้ประสิทธิภาพในการพยากรณ์ภาพรวมของตัวแบบลอจิสติกส์เชิงสัดส่วน 3 กลุ่ม ด้วยตารางจำแนกกลุ่ม (classification table) แสดงไว้ในตารางที่ 6

ตารางที่ 6 ประสิทธิภาพในการพยากรณ์ของตัวแบบลอจิสติกส์เชิงสัดส่วน 3 กลุ่ม

ค่าสังเกต (Observed)	ค่าพยากรณ์ (Predicted)			
	1	2	3	Percent correct
1	27	6	4	72.97
2	5	155	7	74.88
3	2	4	8	57.14
รวม (Overall percentage)				73.64

จากตารางที่ 6 แสดงประสิทธิภาพในการพยากรณ์ของตัวแบบลอจิสติกส์เชิงสัดส่วน 3 กลุ่ม พบว่าประสิทธิภาพในการพยากรณ์กลุ่มของการไม่สำเร็จการศึกษา (กลุ่ม 1) พยากรณ์ได้ถูกต้องร้อยละ 72.97 กลุ่มของการสำเร็จการศึกษาตามเกณฑ์ (กลุ่ม 2) พยากรณ์ได้ถูกต้องร้อยละ 74.88 และกลุ่มของการสำเร็จการศึกษาเกียรตินิยม (กลุ่ม 3) พยากรณ์ได้ถูกต้องร้อยละ 57.14 ดังนั้นประสิทธิภาพในการพยากรณ์กลุ่มในภาพรวมของตัวแบบลอจิสติกส์เชิงสัดส่วน 3 กลุ่ม เมื่อตัวแปรตอบสนองจำแนกประเภทแบบมีอันดับสามารถพยากรณ์ได้ถูกต้องร้อยละ 73.64 ซึ่งผลการตรวจสอบความถูกต้องของตัวแบบของข้อมูลในการทดลองดังกล่าว และข้อมูลในการตรวจสอบความถูกต้องของตัวแบบให้ผลลัพธ์ที่ใกล้เคียงกันในแต่ละกรณีศึกษา

4. สรุปผลการวิจัย

4.1 กรณีตัวแปรตอบสนองแบบจำแนกประเภท 2 กลุ่ม แบบไม่มีอันดับ

ตัวแบบลอจิสติก 2 กลุ่ม ดังในสมการ (1) มีภาวะสารูปดี และมีความเหมาะสมอย่างมีนัยสำคัญทางสถิติที่ 0.05 ปัจจัยเสี่ยงหรือตัวแปรที่ถูกคัดเลือกเข้าสมการคือ ข้าพเจ้าเสียเวลาไปกับการเที่ยวเตร่และเรื่องไม่เิกการเรียนในระดับใด(x18) และเมื่ออ่านหนังสือเสร็จข้าพเจ้ามักจะบันทึกสรุปสั้นๆ (x27) จากค่าสถิติ likely-hood ratio เท่ากับ 22.7954 ค่าสถิติ score เท่ากับ 22.0319 และค่าสถิติ Wald เท่ากับ 17.2781 โดยให้ค่า P-value ที่สอดคล้องกันมีค่าน้อยกว่า 0.05 หมายความว่าพารามิเตอร์ของตัวแบบลอจิสติกมีค่าแตกต่างจากศูนย์ ภายใต้ระดับนัยสำคัญทางสถิติ 0.05 ดังนั้นตัวแปรอธิบายดังกล่าวจึงมีความเหมาะสมที่จะใช้พยากรณ์ และสำหรับตัวสถิติ deviance เท่ากับ 165.6783 โดยให้ค่า P-value ที่สอดคล้องกันมีค่า มากกว่า 0.05 ซึ่งให้ผลยืนยันว่าตัวแบบมีความเหมาะสมที่จะใช้ในการพยากรณ์ต่อไป ภายใต้ระดับนัยสำคัญ 0.05 นอกจากนี้ พบว่าร้อยละของจำนวนคู่ของค่าสังเกตที่มีความสอดคล้องกัน (percent concordant) ซึ่งเท่ากับร้อยละ 66.5 ซึ่งมากกว่าร้อยละของจำนวนคู่ของค่าสังเกตที่ไม่สอดคล้องกัน (percent discordant) ซึ่งเท่ากับร้อยละ 13.7 และพบว่าค่าสถิติ Somers' D เท่ากับ 0.528 ให้ความเกี่ยวพันเชิงบวก หมายความว่าตัวแปรตอบสนองและตัวแปรอธิบายมีความเกี่ยวพันค่อนข้างมากด้วย ดังตารางที่ 7

ดังนั้นถ้า นักศึกษาคนดังกล่าวมีพฤติกรรมการเรียนดีขึ้นบ้าง ความน่าจะเป็นที่นักศึกษาจะไม่สำเร็จการศึกษาภายใน 4 ปี จะลดลงจากเดิมร้อยละ 17 เป็นร้อยละ 13 นอกจากนี้ประสิทธิภาพในการพยากรณ์ภาพรวมของตัวแบบลอจิสติก 2 กลุ่ม ดังตาราง

จำแนกกลุ่ม พบว่าประสิทธิภาพในการพยากรณ์ในระดับของการไม่สำเร็จการศึกษา เช่น พยากรณ์การไม่สำเร็จการศึกษา (กลุ่ม 1) ได้ถูกต้องเท่ากับร้อยละ 81.08 และการพยากรณ์การสำเร็จการศึกษา (กลุ่ม 2) ได้ถูกต้องเท่ากับร้อยละ 86.36 (ตารางที่ 3) และตัว

แบบลอจิสต์ 2 กลุ่ม มีประสิทธิภาพในภาพรวมของการพยากรณ์กลุ่มหรือระดับการไม่สำเร็จการศึกษาและระดับการสำเร็จการศึกษา โดยสามารถพยากรณ์ได้ถูกต้องค่อนข้างสูงคือร้อยละ 85.27 ดังตารางที่ 7

ตารางที่ 7 สรุปตัวอย่างการพยากรณ์ค่าความน่าจะเป็นของการไม่สำเร็จและสำเร็จการศึกษาด้วยตัวแบบลอจิสต์ 2 กลุ่ม

ตัวแปรอธิบาย (x)	ความน่าจะเป็นการไม่สำเร็จการศึกษา	ความน่าจะเป็นของการสำเร็จการศึกษา
x18(1) ระดับมาก และ x27 (1) ไม่ใช่	0.17	0.83
x18(2) ระดับปานกลาง และ x27(1) ไม่ใช่	0.13	0.87

4.2 กรณีตัวแปรตอบสนองเป็นแบบจำแนกประเภท 3 กลุ่ม แบบมีอันดับ

จากตัวแบบลอจิสต์อันดับสามส่วน 3 กลุ่ม ในสมการ (4)-(5) ซึ่งมีภาวะสารูปดีและมีความเหมาะสมอย่างมีนัยสำคัญทางสถิติที่ 0.05 พบว่าปัจจัยที่ส่งผลต่อการไม่สำเร็จการศึกษา และถูกเลือกเข้าตัวแบบคือ ข้าพเจ้าเสียเวลาไปกับการเที่ยวเตร่และเรื่องไม่ใช่การเรียนในระดับใด (x18) เมื่ออาจารย์มอบหมายให้ค้นคว้าหรืออ่านหนังสือเพิ่มเติมข้าพเจ้าจะรีบเข้าห้องสมุดทันทีที่มีเวลาว่างจากการเรียนในระดับใด (x21) เมื่ออ่านหนังสือเสร็จข้าพเจ้ามักจะบันทึกสรุปสั้น ๆ (x27) และในการทบทวนบทเรียนข้าพเจ้าใช้วิธีตั้งคำถามให้ตนเองมากกว่าการอ่านหนังสือไปเรื่อย ๆ (x29) โดยให้ค่าสถิติ likelihood ratio เท่ากับ 30.0359 ค่าสถิติ score เท่ากับ 28.7477 ค่าสถิติ Wald เท่ากับ 26.1792 ซึ่งค่า P-Value ของแต่ละตัวสถิติมีค่าน้อยกว่า 0.05 หมายความว่าพารามิเตอร์ของตัวแบบที่สอดคล้องกันมีค่าไม่เท่ากับศูนย์ ภายใต้ระดับนัยสำคัญเชิงสถิติ 0.05 ดังนั้นตัวแปรอธิบายจึงมีความเหมาะสม และจากการวิเคราะห์ข้อมูลพบว่าร้อยละของจำนวนคู่ของค่าสังเกตที่มีความสอดคล้องกัน

(percent concordant) ซึ่งเท่ากับร้อยละ 71.3 ดีกว่าหรือมากกว่าร้อยละของจำนวนคู่ของค่าสังเกตที่ไม่สอดคล้องกัน (percent discordant) ซึ่งเท่ากับร้อยละ 21.3 และพบว่าค่าสถิติ Somers' D เท่ากับ 0.500 ค่าสถิติ Gamma เท่ากับ 0.539 Tau-a เท่ากับ 0.173 และค่าสถิติ c เท่ากับ 0.750 ทุกค่าให้ความเกี่ยวพันเชิงบวกระหว่างตัวแปรค่อนข้างสูง ดังตารางที่ 8

เช่น สมมติ ถ้าพยากรณ์นักศึกษาคนหนึ่งเสียเวลาไปกับการเที่ยวเตร่และเรื่องไม่ใช่การเรียนในระดับมาก (x18(1)) เมื่ออาจารย์มอบหมายให้ค้นคว้าหรืออ่านหนังสือเพิ่มเติมข้าพเจ้าจะรีบเข้าห้องสมุดทันทีที่มีเวลาว่างจากการเรียนในระดับมาก (x21(1)) เมื่ออ่านหนังสือเสร็จข้าพเจ้าจะไม่บันทึกสรุปสั้น ๆ (x27) และในการทบทวนบทเรียนข้าพเจ้าไม่ใช้วิธีตั้งคำถามให้ตนเองมากกว่าการอ่านหนังสือไปเรื่อย ๆ (x29) พบว่าความน่าจะเป็นที่นักศึกษาไม่สำเร็จการศึกษาเท่ากับ 0.085 ความน่าจะเป็นที่นักศึกษาสำเร็จการศึกษาตามเกณฑ์เท่ากับ 0.860 และความน่าจะเป็นที่สำเร็จการศึกษาเกียรตินิยมเท่ากับ 0.055

แต่ถ้าพยากรณ์นักศึกษาคนหนึ่งเสียเวลาไปกับการเที่ยวเตร่และเรื่องไม่ใช่การเรียนในระดับปาน

กลาง (x18(2)) เมื่ออาจารย์มอบหมายให้ค้นคว้าหรืออ่านหนังสือเพิ่มเติมข้าพเจ้าจะรีบเข้าห้องสมุดทันทีที่มีเวลาว่างจากการเรียนในระดับปานกลาง (x21(2)) เมื่ออ่านหนังสือเสร็จข้าพเจ้ามักจะไม่นั่งทบทวนสรุปสั้น ๆ (x27) และในการทบทวนบทเรียนข้าพเจ้าไม่ใช้วิธีตั้งคำถามให้ตนเองมากกว่าการอ่านหนังสือไปเรื่อย ๆ

(x29) พบว่าความน่าจะเป็นที่นักศึกษาไม่สำเร็จการศึกษาเท่ากับ 0.075 ความน่าจะเป็นที่นักศึกษาสำเร็จการศึกษาตามเกณฑ์เท่ากับ 0.865 และความน่าจะเป็นที่นักศึกษาสำเร็จการศึกษาเกียรตินิยมเท่ากับ 0.060

ตารางที่ 8 สรุปตัวอย่างการพยากรณ์ค่าความน่าจะเป็นของการไม่สำเร็จและสำเร็จการศึกษาด้วยตัวแบบลอจิสติก 3 กลุ่ม

ตัวแปรอธิบาย	ความน่าจะเป็นของการไม่สำเร็จการศึกษา	ความน่าจะเป็นของการสำเร็จการศึกษาตามเกณฑ์	ความน่าจะเป็นของการสำเร็จการศึกษาตามเกณฑ์
x18 (1) ระดับมาก, x21 (1) ระดับมาก, x27 (ไม่ใช่), x29 (ไม่ใช่)	0.085	0.860	0.055
x18 (2) ระดับปานกลาง, x21 (2) ระดับปานกลาง, x27 (ไม่ใช่), x29 (ไม่ใช่)	0.075	0.865	0.06

ดังนั้น ถ้านักศึกษาค้นคว้ามีพฤติกรรมการเรียนที่ดีขึ้นบ้าง จะพบว่าความน่าจะเป็นที่นักศึกษาไม่สำเร็จการศึกษาภายใน 4 ปีจะลดลงได้ และความน่าจะเป็นที่นักศึกษาจะสำเร็จการศึกษาภายใน 4 ปี จะเพิ่มขึ้นตามลำดับ

นอกจากนี้ประสิทธิภาพในการพยากรณ์ภาพรวมของตัวแบบลอจิสติกออดส์เชิงสัดส่วน 3 กลุ่ม ดังตารางจำแนกกลุ่ม ในตารางที่ 6 สามารถพยากรณ์ภาพรวมทั้ง 3 กลุ่มได้ถูกต้องร้อยละ 73.64 ซึ่งค่อนข้างสูงสำหรับกรณีที่มีจำนวนกลุ่ม 3 กลุ่ม

5.รายการอ้างอิง

[1] อัญชิสา อยู่ยืนยง, 2550, การศึกษาถึงปัจจัยที่มีอิทธิพลส่งผลให้เยาวชนตัดสินใจออกจากโรงเรียนกลางคันในระดับมัธยมศึกษาการศึกษา

ตำบลบ้านบัว อำเภอเมือง จังหวัดบุรีรัมย์, มหาวิทยาลัยธรรมศาสตร์, กรุงเทพฯ.

[2] บังอร ชูพงศ์, 2523, การศึกษาสภาพปัญหาที่ส่งผลต่อการปรับตัวของนักเรียนอาชีวศึกษา, วิทยานิพนธ์ปริญญาโท, มหาวิทยาลัยศรีนครินทรวิโรฒประสานมิตร, กรุงเทพฯ.

[3] แบบสอบถามเกี่ยวกับนิสัยในการเรียน, แหล่งที่มา : <http://www.elecptn.rmutl.ac.th>, 16 กันยายน 2556.

[4] Agresti, A., 2007, Introduction to Categorical Data Analysis, John Wiley & Sons, Inc., New York, 372p.

[5] McCullagh, P., 1980, Regression models for ordinal data, J. Royal Statist. B42: 109-142.

- [6] Agresti, A., 2002, Categorical Data Analysis, John Wiley & Sons, Inc., New York, 710 p.
- [7] วีรานันท์ พงศาภักดี, 2555, การวิเคราะห์ข้อมูลจำแนกประเภท: ทฤษฎีและการประยุกต์ด้วย GLIM, SPSS, SAS, และ MTB, พิมพ์ครั้งที่ 3, โรงพิมพ์มหาวิทยาลัยศิลปากร, นครปฐม, 495น.
- [8] Pongsapukdee, V. and Sukgumphaphan, S., 2008, Three cumulative link models for ordinal responses: A review of methods and power comparisons, Songklanakarin J. Sci. Technol. 30: 805-811.
- [9] Walker, S.H. and Duncan, D.B., 1967, Estimation of the probability of an event as a function of several independent variables, Biometrika 54: 167-17