การประมาณขนาดประชากรและสร้างตัวแบบทำนายเพศนกประจำถิ่น ในเขตห้ามล่าสัตว์ป่าบึงบอระเพ็ด จังหวัดนครสวรรค์

Population Size Estimation and Sexual Predictive Model Construction of the Resident Birds in Bueng Boraphet Non-Hunting Area, Nakhon Sawan Province

อุทัยทิพย์ ทนเถื่อน (Uthaithip Thonthuen)[*]
จุฑาภรณ์ สินสมบูรณ์ทอง (Juthaphorn Sinsomboonthong)[**]
มีนา ปทุมสูตร (Mena Patummasut)[***]
ไกรรัตน์ เอี่ยมอำไพ (Krairat Eiamampai)[****]
จิระเดช บุญมาก (Jiradaj Boonmak)[*****]

บทคัดย่อ

งานวิจัยครั้งนี้มีวัตถุประสงค์เพื่อประมาณค่าแบบช่วงของขนาดประชากรนกประจำถิ่นในเขตห้ามล่าสัตว์ป่าบึงบอระเพ็ด จังหวัดนครสวรรค์ ที่ระดับความเชื่อมั่น 95% ด้วยวิธีจับซ้ำโดยใช้ตัวประมาณ 3 ตัว คือ ตัวประมาณของ Zelterman, ตัวประมาณของ Chao และตัวประมาณแบบถ่วงน้ำหนัก นอกจากนี้ในงานวิจัยยังมีการสร้างตัวแบบการทำนายเพศนกกระติ๊ดแดง ซึ่งเป็นนกประจำถิ่นชนิดหนึ่ง ผลการศึกษาพบว่า การประมาณขนาดประชากรนกประจำถิ่นแบบจุดโดยใช้ตัวประมาณของ Chao ให้ค่าประมาณความคลาดเคลื่อนมาตรฐานต่ำที่สุด และมีช่วงความเชื่อมั่น 95% ของขนาดประชากรนกประจำถิ่นอยู่ระหว่าง 2,234.005 ถึง 6,575.783 ส่วนตัวประมาณของ Zelterman และตัวประมาณแบบถ่วงน้ำหนักให้ค่าประมาณแบบจุดที่ใกล้เคียงกัน และให้ค่าประมาณความคลาดเคลื่อนมาตรฐานสูงกว่าตัวประมาณของ Chao แต่ตัวประมาณของ Chao มีความกว้าง

[*] นิสิตปริญญาโท ภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ บางเขน
(Student, Department of Statistics, Faculty of Science, Kasetsart University, Bangkhen. Email; authaithip.sow@gmail.com)

[**] อาจารย์ที่ปรึกษาหลัก, ผู้ช่วยศาสตราจารย์ ดร.; ประจำภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ บางเขน
(Major Advisor, Assistance Professor Dr., Department of Statistics, Faculty of Science, Kasetsart University, Bangkhen. Email; fscijps@ku.ac.th)

[***] อาจารย์ที่ปรึกษาร่วม, อาจารย์; ประจำภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ บางเขน
(Co-Advisor, Lecturer, Department of Statistics, Faculty of Science, Kasetsart University, Bangkhen. Email; fscimnp@ku.ac.th)

[****] นักวิทยาศาสตร์ชำนาญการพิเศษ สถานีวิจัยสัตว์ป่าบึงบอระเพ็ด กรมอุทยานแห่งชาติ สัตว์ป่า และพันธุ์พืช
(Scientist, Senior Professional Level, Bueng Boraphet Wildlife Research Division, Department of National Parks, Wildlife and Plant Conservation, Thailand. Email; jacanaeiam@gmail.com)

[*****] เจ้าพนักงานป่าไม้ชำนาญงาน เขตห้ามล่าสัตว์ป่าบึงบอระเพ็ด กรมอุทยานแห่งชาติ สัตว์ป่า และพันธุ์พืช
(Forestry Officer, Professional Level, Bueng Boraphet Non-Hunting Area, Department of National Parks, Wildlife and Plant Conservation, Thailand. Email; Jiradej.31@hotmail.co.th)

Veridian E-Journal, Science and Technology Silpakorn University
Volume 5 Number 3 May – June 2018 ISSN 2408 - 1248

ของช่วงความเชื่อมั่น 95% ของขนาดประชากรนกประจำถิ่นแคบกว่าตัวประมาณของ Zelterman และตัวประมาณแบบถ่วงน้ำหนัก สำหรับการสร้างตัวแบบทำนายเพศนกกระติ๊ดแดง พิจารณาจากตัวแปรอิสระ 6 ตัว ได้แก่ ความยาวปีก (x_1), ความยาวท้ายทอยถึงปลายปาก (x_2), ความยาวปาก (x_3), ความยาวแข้ง (x_4), ความยาวหาง (x_5) และน้ำหนัก (x_6) ซึ่งใช้เทคนิคการวิเคราะห์การถดถอยโลจิสติคและการวิเคราะห์จำแนกกลุ่ม สรุปได้ว่าการวิเคราะห์การถดถอยโลจิสติคมีอำนาจจำแนกได้ถูกต้องร้อยละ 91.30 สูงกว่าการวิเคราะห์จำแนกกลุ่มที่มีอำนาจจำแนกได้ถูกต้องร้อยละ 86.96 และตัวแปรอิสระที่ถูกคัดเลือกเข้าสมการถดถอยโลจิสติคโดยวิธี Forward Stepwise มี 5 ตัว ได้แก่ ความยาวปีก (x_1), ความยาวท้ายทอยถึงปลายปาก (x_2), ความยาวแข้ง (x_4), ความยาวหาง (x_5) และน้ำหนัก (x_6) สำหรับการวิเคราะห์จำแนกกลุ่ม ตัวแปรอิสระที่ถูกคัดเลือกเข้าสมการจำแนกกลุ่มโดยวิธี Stepwise มี 4 ตัว ได้แก่ ความยาวปีก (x_1), ความยาวแข้ง (x_4), ความยาวหาง (x_5) และน้ำหนัก (x_6)

คำสำคัญ : วิธีจับซ้ำ, การประมาณขนาดประชากร, ตัวแบบทำนาย, การวิเคราะห์การถดถอยโลจิสติค, การวิเคราะห์จำแนกกลุ่ม

Abstract

The purposes of this research were to estimate the population size of resident birds in Bueng boraphet non-hunting area, Nakhon Sawan province by using the 95% confidence interval of three estimators—Zelterman's estimator, Chao's estimator and Weighted estimator—for capture-recapture method. In addition, the sexual predictive models of red avadavat, one species of resident birds, were constructed. The results showed that the population size estimation of the resident birds by using Chao's estimator has the smallest estimated standard error for point estimation, and the 95% confidence interval of this is between 2,234.005 to 6,575.783. The point estimates of Zelterman's and Weighted estimators are close together and their estimated standard errors are higher than that of Chao's estimator. However, the width of the 95% confidence interval of Chao's estimator for population size of the resident birds is narrower than those of Zelterman's and Weighted estimators. For the sexual predictive model construction of the red avadavat, six independent variables—wing length (x_1), occipital to beak length (x_2), beak length (x_3), shin length (x_4), tail length (x_5), weight (x_6)—are considered by using logistic regression and discriminant analysis. It is found that the percentage of correct prediction of logistic regression analysis is 91.30 that is higher than that of discriminant analysis which is equals 86.96. Further, the independent variables selection for fitting logistic regression model by forward stepwise method consists of five independent variables include, wing length (x_1), occipital to beak length (x_2), shin length (x_4), tail length (x_5) and weight (x_6). For discriminant analysis, the independent variables

selection for fitting discriminant function by stepwise method consists of four independent variables including, wing length (x_1), shin length (x_4), tail length (x_5) and weight (x_6).

Keywords: Capture-recapture Method, Population Size Estimation, Predictive Model, Logistic Regression Analysis, Discriminant Analysis

บทนำ

วิธีจับซ้ำหรือการทำเครื่องหมายแล้วจับซ้ำ (Capture-recapture Method) เป็นวิธีการหาขนาดประชากรโดยประมาณของสัตว์ที่เคลื่อนที่ได้และมีจำนวนมาก (ไพรัช, 2527) ซึ่งเหมาะที่จะใช้กับสัตว์ที่หลบซ่อนตัวเก่งหรือเคลื่อนที่เร็ว แต่สามารถดักจับและทำเครื่องหมายที่ตัวสัตว์ได้ การทำเครื่องหมายมีหลายวิธีซึ่งแตกต่างกันตามความเหมาะสมและชนิดของสัตว์ อย่างไรก็ตามการทำเครื่องหมายบนตัวสัตว์ต้องคำนึงถึงจริยธรรมและสวัสดิภาพความปลอดภัยของสัตว์เป็นหลักด้วย ซึ่งวิธีการดังกล่าวสามารถประมาณขนาดประชากรนกได้เช่นกัน โดยสามารถใช้ตาข่ายดักจับและติดเครื่องหมายให้กับนก เนื่องจากนกเป็นสัตว์ที่มีกระดูกสันหลังที่มีวิวัฒนาการเป็นสัตว์ปีกที่บินได้แท้ มีพัฒนาการโครงสร้างร่างกายให้เหมาะสมกับการบิน สภาพภูมิประเทศ และระยะทาง เป็นผลให้นกมีชนิดพันธุ์ที่หลากหลาย มีขนาดประชากร สภาพถิ่นอาศัย บทบาทหน้าที่ทางนิเวศ และพฤติกรรมการดำรงชีวิตที่แตกต่างกัน ปัจจุบันนกจัดเป็นพวกสัตว์บกมีกระดูกสันหลังที่ประสบความสำเร็จในการดำรงชีวิตมีถิ่นกระจายพันธุ์พบได้ทุกภูมิภาคของโลก และมีความหลากหลายชนิดพันธุ์มากที่สุดถึง 9,672 ชนิด รวมทั้งนกยังเป็นตัวชี้วัดสภาพ แวดล้อมและการเปลี่ยนแปลงสภาพแวดล้อมได้เหมาะสมดีกว่าสัตว์จำพวกอื่นๆ โดยเฉพาะชนิดพันธุ์เด่นในท้องถิ่น

การประมาณขนาดประชากรด้วยข้อมูลจากวิธี Capture-recapture มีผู้เสนอและพัฒนาตัวประมาณขนาดประชากรไว้มากมายหลายตัวประมาณ เช่น ในปี ค.ศ. 1951 Chapman (Chapman, 1951) ได้เสนอตัวประมาณขนาดประชากรที่ไม่เอนเอียงซึ่งเป็นตัวประมาณที่ปรับมาจากตัวประมาณของ Lincoln-Petersen ซึ่งเป็นตัวประมาณที่เอนเอียง (Petersen, 1896; Lincoln, 1930) โดยตัวประมาณดังกล่าวเป็นตัวประมาณที่เหมาะสำหรับการจับซ้ำเพียงสองครั้ง ซึ่งต่อมาได้มีผู้เสนอและพัฒนาตัวประมาณที่เหมาะสำหรับการจับซ้ำตั้งแต่สองครั้งขึ้นไป โดยในปี ค.ศ. 1988 Zelterman (Zelterman, 1988) ได้เสนอตัวประมาณภายใต้การแจกแจงปัวซงที่ถูกตัดทอนศูนย์ (Zero-truncated Poisson Distribution) เพื่อใช้ในการประมาณขนาดประชากร ซึ่งเป็นตัวประมาณที่ให้ค่าความเอนเอียงต่ำแต่ความแปรปรวนสูง ถัดมาในปี ค.ศ. 1989 Chao (Chao, 1989) ได้เสนอตัวประมาณขนาดประชากรสำหรับข้อมูลที่บางเบา (Sparse Data) ซึ่งพบว่าเป็นตัวประมาณที่ให้ค่าความเอนเอียงน้อยกว่าตัวประมาณของ Darroch (Darroch, 1958) นอกจากนี้ ในปี ค.ศ. 2008 Böhning (Böhning, 2008) ได้เสนอตัวประมาณแบบถ่วงน้ำหนัก โดยเป็นการถ่วงน้ำหนักจากตัวประมาณของ Zelterman และตัวประมาณของ Chao ซึ่งเป็นการรวมด้านดีของสองตัวประมาณ กล่าวคือ ตัวประมาณของ Zelterman ให้ค่าความเอนเอียงต่ำแต่ความแปรปรวนสูง ในขณะที่ตัวประมาณของ Chao ให้ค่าความเอนเอียงติดลบแต่ความแปรปรวนต่ำ เป็นต้น

Veridian E-Journal, Science and Technology Silpakorn University
Volume 5 Number 3 May – June 2018 ISSN 2408 - 1248

จากการศึกษางานวิจัยที่เกี่ยวข้องกับการประมาณขนาดประชากรด้วยข้อมูลจากวิธี Capture-recapture ดังกล่าวข้างต้น ทำให้ผู้วิจัยสนใจที่จะประมาณขนาดประชากรนกประจำถิ่นในพื้นที่เขตห้ามล่าสัตว์ป่าบึงบอระเพ็ด จังหวัดนครสวรรค์ โดยใช้ข้อมูลจากวิธี Capture-recapture ซึ่งในงานวิจัยนี้ศึกษาตัวประมาณขนาดประชากร 3 ตัว คือ ตัวประมาณของ Zelterman ตัวประมาณของ Chao และตัวประมาณแบบถ่วงน้ำหนัก โดยทำการศึกษาในพื้นที่บึงบอระเพ็ด เนื่องจากบึงบอระเพ็ดเป็นทะเลสาบน้ำจืดขนาดใหญ่เป็นระบบนิเวศน้ำที่มีคุณค่าและมีเอกลักษณ์เหมาะแก่การเป็นที่อยู่อาศัยของนกนานาชนิด จึงเหมาะแก่การเป็นพื้นที่ศึกษาระบบนิเวศและการสำรวจประชากรนก (สำนักงานนโยบายและแผนทรัพยากรธรรมชาติและสิ่งแวดล้อม, 2546) ซึ่งผลการศึกษาของงานวิจัยนี้ สามารถนำไปเป็นแนวทางในการอนุรักษ์และฟื้นฟูประชากรนกประจำถิ่น และเพื่อเป็นข้อมูลพื้นฐานในการติดตามการเปลี่ยนแปลงของประชากรนกประจำถิ่นด้วย

นอกจากผู้วิจัยจะสนใจศึกษาการประมาณขนาดประชากรนกประจำถิ่นแล้ว ผู้วิจัยยังสนใจศึกษาการสร้างตัวแบบในการทำนายเพศนกเพื่อเป็นข้อมูลในการอนุรักษ์พันธุ์นกประจำถิ่นในเขตห้ามล่าสัตว์ป่าบึงบอระเพ็ด จังหวัดนครสวรรค์ โดยทั่วไปแล้วการจำแนกเพศนกสามารถทำได้หลากหลายวิธี ไม่ว่าจะเป็นการพิจารณาความแตกต่างที่ปรากฏภายนอก เช่น สีขน สีปาก รูปร่าง ขนาด ฯลฯ หรือการใช้ขนาดความแตกต่างของช่องเชิงกรานเป็นตัวจำแนกเพศนก หรือการตรวจดูอัณฑะและรังไข่โดยใช้กล้องส่องตรวจภายในช่องท้อง หรือการตรวจ DNA จากเลือดหรือโคนขนของนก รวมทั้งการดูพฤติกรรมจำเพาะของแต่ละเพศ (ไชยรัตน์, ม.ป.ป.) จะเห็นว่าการจำแนกเพศนกที่ได้กล่าวมานั้นบางวิธีอาจไม่สามารถจำแนกเพศนกได้อย่างถูกต้อง หรือบางวิธีอาจเกิดอันตรายกับตัวนกได้ รวมทั้งบางวิธีอาจมีค่าใช้จ่ายสูงในการจำแนกเพศนก ดังนั้นผู้วิจัยจึงมีความสนใจใช้เทคนิคการวิเคราะห์ทางสถิติเพื่อทำนายเพศนก ได้แก่ การวิเคราะห์การถดถอยโลจิสติค (Logistic Regression Analysis) และการวิเคราะห์จำแนกกลุ่ม (Discriminant Analysis) เนื่องจากการใช้เทคนิคการวิเคราะห์ทางสถิติดังกล่าวนี้สามารถช่วยจำแนกเพศนกได้โดย ไม่ต้องเสียค่าใช้จ่ายสูงหรือส่งผลอันตรายต่อตัวนก รวมทั้งการใช้เทคนิคการวิเคราะห์ทางสถิติสามารถบอกได้ว่าปัจจัยใดบ้างที่มีผลต่อการจำแนกเพศนกประจำถิ่น

วัตถุประสงค์การวิจัย

1) เพื่อประมาณขนาดประชากรนกประจำถิ่นในเขตห้ามล่าสัตว์ป่าบึงบอระเพ็ด จังหวัดนครสวรรค์ ที่ระดับความเชื่อมั่น 95% เมื่อสุ่มตัวอย่างด้วยวิธี Capture-recapture

2) เพื่อสร้างตัวแบบการทำนายเพศนกประจำถิ่นในเขตห้ามล่าสัตว์ป่าบึงบอระเพ็ด จังหวัดนครสวรรค์

วิธีการดำเนินการวิจัย

1) กำหนดพื้นที่ที่ทำการศึกษา คือ บึงบอระเพ็ด ซึ่งเป็นบึงน้ำจืดที่ใหญ่ที่สุดของภาคเหนือตอนล่าง โดยตั้งอยู่ในเขตอำเภอเมือง อำเภอชุมแสง และอำเภอท่าตะโก จังหวัดนครสวรรค์ มีพื้นที่ทั้งหมดประมาณ 212.38 ตารางกิโลเมตร หรือประมาณ 132,737.14 ไร่ มีระดับความลึกเฉลี่ยของน้ำในบึงประมาณ 1.6 เมตร และเป็นพื้นที่ชุ่มน้ำที่มีความสำคัญระดับนานาชาติของประเทศไทย (สำนักงานนโยบายและแผนทรัพยากรธรรมชาติและสิ่งแวดล้อม, 2546)

2) การเก็บข้อมูลปฐมภูมิของนกประจำถิ่นในเขตห้ามล่าสัตว์ป่าบึงบอระเพ็ด มีขั้นตอนดังนี้

2.1) กำหนดจุด/สถานีเพื่อจับนกประจำถิ่นใส่ห่วงขาให้กระจายทั่วพื้นที่เขตห้ามล่าสัตว์ป่าบึงบอระเพ็ด โดยแบ่งออกเป็น 6 สถานี ประกอบด้วย สถานีอุทยานนกน้ำ (สวนบัว), สถานีแหลมตาเส็ง, สถานีทางเข้าประมงเชิงพาณิชย์, สถานีแหลมนา, สถานีเนินระฆัง และสถานีฝายเก่า

2.2) ในแต่ละสถานีทำการวางตาข่าย เพื่อใช้ในการจับนกประจำถิ่นจำนวน 25 ผืนต่อสถานี ซึ่งขนาดตาข่ายแต่ละผืนมีความกว้าง 2 เมตร และความยาว 9 เมตร

2.3) ทำการวางตาข่ายและจับนกประจำถิ่นครั้งแรกในเดือนตุลาคม พ.ศ. 2560 โดยในแต่ละสถานีจะใช้เวลา 1 วัน โดยทำการวางตาข่ายและเริ่มจับนกประจำถิ่นตั้งแต่เวลา 06.00 น. – 12.00 น. เมื่อดักจับได้นกประจำถิ่นแล้ว ทำการใส่ห่วงขา บันทึกข้อมูลรายละเอียดเกี่ยวกับตัวนกประจำถิ่น ชนิดของนกประจำถิ่น จำนวนนกประจำถิ่น พร้อมถ่ายรูปประกอบแล้วปล่อยนกประจำถิ่นคืนสู่ธรรมชาติ

2.4) ทำการจับนกประจำถิ่นซ้ำเป็นครั้งที่ 2 ในเดือนพฤศจิกายน พ.ศ. 2560 โดยวางตาข่ายที่สถานีเดิมทั้ง 6 สถานี บันทึกข้อมูลแล้วปล่อยนกประจำถิ่นคืนสู่ธรรมชาติ

2.5) ทำการจับนกประจำถิ่นซ้ำเป็นครั้งที่ 3 ในเดือนธันวาคม พ.ศ. 2560 โดยวางตาข่ายที่สถานีเดิมทั้ง 6 สถานี บันทึกข้อมูลแล้วปล่อยนกประจำถิ่นคืนสู่ธรรมชาติ

3) การเก็บรวบรวมข้อมูลทุติยภูมิของนกประจำถิ่นในเขตห้ามล่าสัตว์ป่าบึงบอระเพ็ด ซึ่งเก็บรวบรวมในปี พ.ศ. 2559 เพื่อนำข้อมูลมาสร้างตัวแบบการทำนายเพศนกประจำถิ่น ประกอบด้วยข้อมูลชนิดนก เพศนก ความยาวปีก ความยาวท้ายทอยถึงปลายปาก ความยาวปาก ความยาวแข้ง ความยาวหาง และน้ำหนัก สำหรับงานวิจัยนี้จะสร้างตัวแบบทำนายเพศนกกระติ๊ดแดง ซึ่งเป็นนกประจำถิ่นในเขตห้ามล่าสัตว์ป่าบึงบอระเพ็ด

4) การสุ่มตัวอย่างด้วยวิธี Capture-recapture เป็นวิธีการหาจำนวนโดยประมาณของสัตว์ที่เคลื่อนที่ได้และมีจำนวนมาก ซึ่งเหมาะที่จะใช้กับสัตว์ที่หลบซ่อนตัวเก่งหรือเคลื่อนที่เร็ว แต่สามารถดักจับและทำเครื่องหมายที่ตัวสัตว์ได้ การทำเครื่องหมายมีหลายวิธีซึ่งแตกต่างกันตามความเหมาะสมและชนิดของสัตว์ อย่างไรก็ตามการทำเครื่องหมายบนตัวสัตว์ต้องคำนึงถึงจริยธรรมและสวัสดิภาพความปลอดภัยของสัตว์เป็นหลัก ซึ่งการทำเครื่องหมายและจับซ้ำอาจทำซ้ำๆ ได้หลายรอบก็ได้ซึ่งอาจมีปัจจัยอื่นมาทำให้การวิเคราะห์ยุ่งยากขึ้น (ประชุม, 2552) หรือทำการจับซ้ำมาเพียงแค่ครั้งเดียวก็ได้ (ธงชัย, ม.ป.ป.) โดยการสุ่มตัวอย่างด้วยวิธี Capture-recapture มีข้อสมมติเบื้องต้น และตัวประมาณขนาดประชากรด้วยวิธี Capture-recapture ที่เกี่ยวข้องในงานวิจัยนี้ มีดังนี้

4.1) ข้อสมมติเบื้องต้นของการประมาณขนาดประชากรด้วยวิธี Capture-recapture (ไพรัช, 2527) มีดังนี้

4.1.1) จำนวนหน่วยชีวิตของประชากรต้องคงที่ หมายความว่าไม่มีการเพิ่มหรือลดจำนวนอันเนื่องมาจากการเกิด การตาย การอพยพเข้า และการอพยพออก

4.1.2) ทุกๆ หน่วยชีวิตจะต้องมีโอกาสถูกจับขึ้นมาเท่ากัน และการจับแต่ละครั้งจะต้องเป็นแบบสุ่ม ในการจับขึ้นมาครั้งที่ 2 จะต้องนับสัตว์ที่มีเครื่องหมายติดทุกตัว

4.1.3) การสุ่มตัวอย่างแต่ละครั้งจะต้องเป็นอิสระกัน

Veridian E-Journal, Science and Technology Silpakorn University
Volume 5 Number 3 May – June 2018 ISSN 2408 - 1248

4.1.4) ต้องมีระยะเวลาช่วงหนึ่ง หลังจากจับขึ้นมาติดเครื่องหมายครั้งแรกแล้วปล่อยลงไป เพื่อให้มีการกระจายกันภายในกลุ่มระหว่างพวกที่ติดและไม่ติดเครื่องหมาย ซึ่งระยะเวลาที่ใช้นั้นขึ้นอยู่กับชนิดของสัตว์

4.1.5) เครื่องหมายที่ติดจะไม่หลุดหายและไม่เป็นอันตรายหรือกระทบกระเทือนต่อการดำรงชีวิต

4.2) ตัวประมาณขนาดประชากรด้วยวิธี Capture-recapture ที่เกี่ยวข้องในงานวิจัยมี 3 ตัว ได้แก่

4.2.1) ตัวประมาณขนาดประชากรของ Zelterman ในปี ค.ศ.1988 Zelterman ได้แนะนำตัวประมาณขนาดประชากรด้วยวิธี Capture-recapture ภายใต้การแจกแจงปัวซงที่ถูกตัดทอนศูนย์ (Zero-truncated Poisson Distribution) ซึ่งเป็นตัวประมาณที่มีค่าความเอนเอียงต่ำแต่มีค่าความแปรปรวนสูง คือ

$$\hat{N}_{Zel} = \frac{n}{1- \exp\left(\frac{-2f_2}{f_1}\right)}$$

ตัวประมาณความแปรปรวนของ $\hat{N}_{Zel}$ คือ

$$\hat{V}\left(\hat{N}_{Zel}\right) = nG\left(\hat{\lambda}\right)\left[1+nG\left(\hat{\lambda}\right)\hat{\lambda}^2\left(\frac{1}{f_1}+\frac{1}{f_2}\right)\right] \text{ เมื่อ } G\left(\hat{\lambda}\right) = \frac{\exp\left(-\hat{\lambda}\right)}{\left(1-\exp\left(-\hat{\lambda}\right)\right)^2} \text{ และ } \hat{\lambda} = \frac{2f_2}{f_1}$$

และตัวประมาณความคลาดเคลื่อนมาตรฐานของ $\hat{N}_{Zel}$ คือ $SE\left(\hat{N}_{Zel}\right) = \sqrt{\hat{V}\left(\hat{N}_{Zel}\right)}$

โดยที่ f_1 คือ จำนวนสัตว์ที่จับซ้ำมา 1 ครั้ง
f_2 คือ จำนวนสัตว์ที่จับซ้ำมา 2 ครั้ง
f_3 คือ จำนวนสัตว์ที่จับซ้ำมา 3 ครั้ง
n คือ จำนวนสัตว์ที่ถูกจับมาทั้งหมด ซึ่ง $n = f_1 + f_2 + f_3$

ดังนั้น การประมาณช่วงความเชื่อมั่น 95% สำหรับขนาดประชากรโดยใช้ตัวประมาณของ Zelterman $\left(\hat{N}_{Zel}\right)$ คือ $\hat{N}_{Zel} \pm 1.96\left[SE\left(\hat{N}_{Zel}\right)\right]$

4.2.2) ตัวประมาณขนาดประชากรของ Chao ในปี ค.ศ.1989 Chao ได้เสนอตัวประมาณขนาดประชากรด้วยวิธี Capture-recapture สำหรับข้อมูลที่บางเบา (Sparse Data) หรือข้อมูลที่มีการจับซ้ำเพียงหนึ่งครั้งหรือสองครั้งเท่านั้น คือ

$$\hat{N}_{Chao} = n + \frac{f_1^2}{2f_2}$$

ตัวประมาณความแปรปรวนของ $\hat{N}_{Chao}$ คือ

$$\hat{V}\left(\hat{N}_{Chao}\right) = 0.25\left(\frac{f_1^4}{f_2^3}\right) + \frac{f_1^3}{f_2^2} + 0.5\left(\frac{f_1^2}{f_2}\right) - 0.25\left(\frac{f_1^4}{\left(f_2^2 n\right)}\right) - 0.5\left(\frac{f_1^4}{f_2\left(2f_2 n + f_1^2\right)}\right)$$

และตัวประมาณความคลาดเคลื่อนมาตรฐานของ $\hat{N}_{Chao}$ คือ $SE\left(\hat{N}_{Chao}\right) = \sqrt{\hat{V}\left(\hat{N}_{Chao}\right)}$

โดยที่ f_1 คือ จำนวนสัตว์ที่จับซ้ำมา 1 ครั้ง

f_2 คือ จำนวนสัตว์ที่จับซ้ำมา 2 ครั้ง

f_3 คือ จำนวนสัตว์ที่จับซ้ำมา 3 ครั้ง

n คือ จำนวนสัตว์ที่ถูกจับมาทั้งหมด ซึ่ง $n = f_1 + f_2 + f_3$

ดังนั้น การประมาณช่วงความเชื่อมั่น 95% สำหรับขนาดประชากรโดยใช้ตัวประมาณของ Chao $\left(\hat{N}_{Chao}\right)$ ที่ Burnham (Burnham, 1978) แนะนำให้ใช้การแปลง log เพื่อปรับปรุงช่วงความเชื่อมั่นแบบทั่วไป คือ

$$\left[n + \frac{\left(\hat{N}_{Chao} - n\right)}{C}, n + \left(\hat{N}_{Chao} - n\right)C\right] \text{ โดยที่ } C = \exp\left\{1.96\left[\log\left(1 + \frac{\hat{V}\left(\hat{N}_{Chao}\right)}{\left(\hat{N}_{Chao} - n\right)^2}\right)\right]^{1/2}\right\}$$

4.2.3) ตัวประมาณขนาดประชากรแบบถ่วงน้ำหนัก ในปี ค.ศ.2008 Böhning ได้เสนอตัวประมาณขนาดประชากรด้วยวิธี Capture-recapture แบบถ่วงน้ำหนัก โดยเป็นการถ่วงน้ำหนักจากตัวประมาณของ Zelterman และตัวประมาณของ Chao ซึ่ง Böhning ได้มีการรวมด้านดีของทั้งสอง ตัวประมาณเพื่อสร้างตัวประมาณใหม่แบบถ่วงน้ำหนัก ดังนั้น ตัวประมาณขนาดประชากรแบบถ่วงน้ำหนัก ที่ Böhning ได้เสนอไว้ คือ

$$\hat{N}_w = \frac{1}{2}\left(\hat{N}_{Zel} + \hat{N}_{Chao}\right) = n + \frac{f_1^2}{2f_2}\left(\frac{1}{2} + \frac{n}{2(f_1 + f_2)}\right)$$

ตัวประมาณความแปรปรวนของ $\hat{N}_w$ คือ

$$\hat{V}\left(\hat{N}_w\right) = \frac{n}{\hat{N}_w}\left(\hat{N}_w - n\right) + \left(\frac{1}{2} + \frac{n}{2(f_1 + f_2)}\right)^2 \frac{f_1^3}{f_2^2}\left(1 + \frac{f_1}{4f_2}\left(1 - \frac{f_2}{n}\right)\right)$$

และตัวประมาณความคลาดเคลื่อนมาตรฐานของ $\hat{N}_w$ คือ $SE\left(\hat{N}_w\right) = \sqrt{\hat{V}\left(\hat{N}_w\right)}$

โดยที่ f_1 คือ จำนวนสัตว์ที่จับซ้ำมา 1 ครั้ง

f_2 คือ จำนวนสัตว์ที่จับซ้ำมา 2 ครั้ง

f_3 คือ จำนวนสัตว์ที่จับซ้ำมา 3 ครั้ง

n คือ จำนวนสัตว์ที่ถูกจับมาทั้งหมด ซึ่ง $n = f_1 + f_2 + f_3$

ดังนั้น การประมาณช่วงความเชื่อมั่น 95% สำหรับขนาดประชากรโดยใช้ตัวประมาณแบบถ่วงน้ำหนัก $\left(\hat{N}_w\right)$ คือ $\hat{N}_w \pm 1.96\left[SE\left(\hat{N}_w\right)\right]$

Veridian E-Journal, Science and Technology Silpakorn University
Volume 5 Number 3 May – June 2018 ISSN 2408 - 1248

5) การวิเคราะห์การถดถอยโลจิสติค (Logistic Regression Analysis)

5.1) วัตถุประสงค์ของการวิเคราะห์การถดถอยโลจิสติค (ยุทธ, 2555) เพื่อศึกษาว่าตัวแปรอิสระใดบ้างที่สามารถใช้อธิบายโอกาสการเกิดเหตุการณ์หรือการไม่เกิดเหตุการณ์ที่สนใจของตัวแปรตาม พร้อมทั้งศึกษาระดับความสัมพันธ์ของตัวแปรอิสระแต่ละตัว และเพื่อทำนายโอกาสที่จะเกิดเหตุการณ์ที่สนใจจากสมการถดถอยโลจิสติคที่เหมาะสม โดยเลือกตัวแปรอิสระที่เหมาะสมเพื่อทำให้เปอร์เซ็นต์ของความถูกต้องในการทำนายมีค่าสูงสุด

5.2) ข้อสมมติเบื้องต้นของการวิเคราะห์การถดถอยโลจิสติคมีดังนี้

5.2.1) ตัวแปรอิสระหรือตัวแปรทำนาย (x) เป็นตัวแปรที่มีระดับข้อมูลอยู่ในระดับช่วงเป็นอย่างต่ำ กรณีที่เป็นข้อมูลเชิงกลุ่มให้แปลงเป็นตัวแปรหุ่นที่มีค่าเป็น 0 กับ 1 เท่านั้น ส่วนตัวแปรตามกรณีที่เป็นการวิเคราะห์การถดถอยโลจิสติคแบบทวิ (Binary Logistic Regression Analysis) จะกำหนด 2 ค่า คือ 0 กับ 1 (ยุทธ, 2555)

5.2.2) ค่าเฉลี่ยของความคลาดเคลื่อนสุ่มเป็นศูนย์หรือ $\mathrm{E}(\varepsilon) = 0$ (กัลยา, 2551ก)

5.2.3) ตัวแปรอิสระไม่มีความสัมพันธ์หรือไม่เกิดปัญหา Multicollinearity (กัลยา, 2551ข) โดยพิจารณาจากค่า Tolerance และค่า Variance Inflation Factor (VIF) ซึ่งหากค่า Tolerance มีค่าเข้าใกล้ 1 และค่า Variance Inflation Factor (VIF) ไม่ได้มีค่าเข้าใกล้ 10 แสดงว่าตัวแปรอิสระไม่มีความสัมพันธ์กันหรือไม่เกิดปัญหา Multicollinearity (ปภากร, 2556)

5.2.4) การวิเคราะห์การถดถอยโลจิสติคจะต้องใช้ขนาดตัวอย่าง (n) มากกว่าการวิเคราะห์การถดถอยแบบปกติ โดยจะใช้ขนาดตัวอย่าง (n) ไม่น้อยกว่า 30p โดยที่ p คือ จำนวนตัวแปรอิสระ (กัลยา, 2551ข)

5.3) วิธีเลือกตัวแปรอิสระเข้าสมการถดถอยโลจิสติค มี 3 วิธี (ยุทธ, 2555) ดังนี้

5.3.1) Enter Method เป็นวิธีที่เลือกตัวแปรอิสระทั้งหมดเข้าสมการถดถอยโลจิสติคพร้อมกันในขั้นตอนเดียว ซึ่งในการพิจารณาตัวแปรอิสระที่เหมาะสมที่จะนำมาใช้ทำนายในสมการถดถอยโลจิสติค ผู้วิจัยจะต้องเป็นผู้ตัดสินใจเองว่าตัวแปรอิสระตัวใดบ้างที่มีความสัมพันธ์กับตัวแปรตามหรือตัวแปรอิสระตัวใดบ้างควรจะอยู่ในสมการถดถอยโลจิสติค โดยพิจารณาจากค่าสถิติทดสอบ ถ้ามีนัยสำคัญทางสถิติถือว่าตัวแปรอิสระนั้นควรอยู่ในสมการถดถอยโลจิสติค

5.3.2) Forward Method วิธีการนี้จะคัดเลือกตัวแปรอิสระที่อธิบายความผันแปรของ ตัวแปรตามได้สูงสุดและมีนัยสำคัญทางสถิติเข้าสมการก่อน จากนั้นจึงเลือกตัวแปรอิสระที่อธิบายความผันแปรของตัวแปรตามได้อันดับรองลงมาและมีนัยสำคัญทางสถิติเข้าสมการตามลำดับ การนำตัวแปรอิสระเข้าสมการจะทำเช่นนี้ไปเรื่อยๆ จนกระทั่งไม่มีตัวแปรอิสระตัวใดที่อธิบายความผันแปรของตัวแปรตามได้อย่างมีนัยสำคัญทางสถิติอีกแล้ว โดยที่การนำตัวแปรอิสระเข้าสมการแบบ Forward Method มีวิธีย่อยๆ อีก 3 วิธี ได้แก่

ก) วิธี Forward Stepwise: Likelihood Ratio วิธีนี้บางทีเรียกว่า Forward LR ซึ่งวิธีนี้จะเริ่มจากการนำตัวแปรอิสระเข้าสมการทีละ 1 ตัว โดยที่ตัวแปรอิสระที่เลือกเข้าสมการทำให้ค่าทำนายโอกาสที่จะเกิดเหตุการณ์ที่สนใจถูกต้องมากขึ้น เกณฑ์ในการพิจารณาเลือกตัวแปรอิสระเข้าสมการ คือ เลือกตัวแปรอิสระที่ให้ค่าแสดงความสัมพันธ์มากที่สุดก่อนและมีนัยสำคัญทางสถิติ เมื่อนำตัวแปรอิสระเข้าสมการแล้วจะมีการตรวจสอบตัวแปรอิสระนั้นอีกว่าควรจะถูกตัดออกหรือควรจะคงอยู่ในสมการโดยพิจารณาจากอัตราส่วนความ

เป็นไปได้หรือจากการเปลี่ยนแปลงของ -2LL ถ้าค่า -2LL ลดลงแสดงว่าตัวแปรอิสระนั้นควรจะคงอยู่ในสมการถดถอยโลจิสติค

ข) วิธี Forward Stepwise: Wald วิธีนี้เหมือนกับวิธี Forward LR ทุกประการเพียงแต่จะพิจารณาจากค่าสถิติทดสอบของวาลด์ (Wald Statistic) เท่านั้น

ค) วิธี Forward Stepwise: Condition วิธีนี้จะเหมือนกับวิธี Forward LR แต่แตกต่างกันตรงที่วิธี Forward LR เป็นวิธีที่ไม่มีเงื่อนไข (Unconditional) ส่วนวิธีนี้จะมีเงื่อนไข (Condition)

5.3.3) Backward Method เป็นวิธีที่นำตัวแปรอิสระทั้ง p ตัว ($x_1, x_2, ...,x_p$) เข้าสมการพร้อมกัน จากนั้นพิจารณาตัวแปรอิสระที่อธิบายความผันแปรของตัวแปรตามได้น้อยที่สุดออกจากสมการก่อน ทำเช่นนี้ไปเรื่อยๆ จนกระทั่งเหลือตัวแปรอิสระที่สามารถอธิบายความผันแปรของตัวแปรตามได้อย่างมีนัยสำคัญทางสถิติ การนำตัวแปรอิสระเข้าสมการแบบ Backword Method มีวิธีย่อยๆ อีก 3 วิธี ได้แก่

ก) วิธี Backward Stepwise: Likelihood Ratio วิธีนี้บางทีเรียกว่า Backward LR ซึ่งเป็นวิธีที่ตรงกันข้ามกับวิธี Forward Stepwise: Likelihood Ratio เนื่องจากเป็นวิธีที่นำตัวแปรอิสระทั้งหมด p ตัว ($x_1, x_2, ...,x_p$) เข้าสมการ แล้วพิจารณาว่าจะนำตัวแปรอิสระตัวใดออกจากสมการ โดยพิจารณานำออกทีละ 1 ตัว โดยพิจารณาจากเกณฑ์การนำตัวแปรอิสระออกจากสมการคือ จะนำตัวแปรอิสระที่ไม่มีผลต่อการทำนายโอกาสที่จะเกิดเหตุการณ์ที่สนใจออกจากสมการ โดยที่ตัวแปรอิสระตัวแรกที่จะนำออกจากสมการจะเป็นตัวแปรที่มีผลต่อการทำนายโอกาสที่จะเกิดเหตุการณ์ที่สนใจน้อยที่สุด สำหรับเกณฑ์การนำตัวแปรอิสระออกจากสมการจะพิจารณาจากอัตราส่วนความเป็นไปได้หรือจากการเปลี่ยนแปลงของ -2LL เหมือนวิธี Forward Stepwise: Likelihood Ratio

ข) วิธี Backward Stepwise: Wald วิธีนี้จะเหมือนกับวิธี Backward LR ทุกประการเพียงแต่จะพิจารณาจากค่าสถิติทดสอบของวาลด์ (Wald Statistic) เท่านั้น

ค) วิธี Backward Stepwise: Condition วิธีนี้จะเหมือนกับวิธี Backward LR ทุกประการแต่แตกต่างกันตรงที่วิธี Backward LR เป็นวิธีที่ไม่มีเงื่อนไข ส่วนวิธีนี้จะมีเงื่อนไข

5.4) ตัวแบบการวิเคราะห์การถดถอยโลจิสติคทวิ ตัวแปรตาม (y) มี 2 ค่า คือ ไม่เกิดเหตุการณ์ (y=0) หรือเกิดเหตุการณ์ (y=1) มีความสัมพันธ์กับตัวแปรอิสระ (x) ซึ่งไม่อยู่ในรูปเชิงเส้น ทั้งนี้เพราะว่าตัวแปรตามมี 2 ค่า คือ 0 กับ 1 จึงเป็นไปไม่ได้ที่ความสัมพันธ์จะอยู่ในรูปเชิงเส้น จึงต้องมีการปรับให้ความสัมพันธ์อยู่ในรูปเชิง กรณีตัวแปรอิสระมากกว่า 1 ตัว การวิเคราะห์การถดถอยโลจิสติคที่เป็นแบบทวิจะได้ฟังก์ชันความน่าจะเป็นของการเกิดเหตุการณ์ที่สนใจดังนี้ (ยุทธ, 2555)

$$P_y = \frac{1}{1+e^{-f(x)}} = \frac{1}{1+e^{-(b_0 + b_1x_1 +...+ b_px_p)}} = \frac{e^{b_0 + b_1x_1 +...+ b_px_p}}{1+e^{b_0 + b_1x_1 +...+ b_px_p}}$$

เมื่อ P_y คือ ความน่าจะเป็นของการเกิดเหตุการณ์ที่สนใจ (y=1)

e คือ ฟังก์ชันเลขชี้กำลัง (Exponential Function) โดยที่ e = 2.71828...

f(x) คือ ฟังก์ชันของตัวแปรทำนาย โดยที่ $f(x) = b_0 + b_1x_1 +...+ b_px_p$

b_i คือ สัมประสิทธิ์การถดถอยโลจิสติค โดยที่ i = 0, 1, ..., p

Veridian E-Journal, Science and Technology Silpakorn University
Volume 5 Number 3 May – June 2018 ISSN 2408 - 1248

สมมติให้ P_y คือ ความน่าจะเป็นของการเกิดเหตุการณ์ที่สนใจ (y=1)

Q_y คือ ความน่าจะเป็นของการไม่เกิดเหตุการณ์ที่สนใจ (y=0)

ดังนั้น $Q_y = 1-P_y$ หรือ $Q_y = 1-\left(\frac{e^{b_0 + b_1x_1 +...+ b_px_p}}{1 + e^{b_0 + b_1x_1 +...+ b_px_p}}\right) = \frac{1}{1+e^{b_0 + b_1x_1 +...+ b_px_p}}$

จากความสัมพันธ์ระหว่างตัวแปรอิสระกับตัวแปรตามของการวิเคราะห์การถดถอยโลจิสติคที่ไม่เป็นรูปเชิงเส้น จึงต้องมีการปรับให้ความสัมพันธ์อยู่ในรูปเชิงเส้น โดยอยู่ในรูปแบบของ odds หรือ odd ratio โดยที่ odds หมายถึง อัตราส่วนระหว่างโอกาสที่จะเกิดเหตุการณ์ที่สนใจกับโอกาสที่จะไม่เกิดเหตุการณ์ที่สนใจซึ่งจะได้

$$\text{odds} = \frac{P_y}{Q_y}$$

ค่าของ odds แสดงถึงโอกาสที่จะเกิดเหตุการณ์ที่สนใจเป็นกี่เท่าของโอกาสที่จะไม่เกิดเหตุการณ์ที่สนใจ เช่น odds ของการโยนเหรียญ 1 ครั้ง เท่ากับ 1 แสดงว่า โอกาสที่จะเกิดเหตุการณ์ที่สนใจกับโอกาสที่จะไม่เกิดเหตุการณ์ที่สนใจเท่ากัน หรือถ้า odds มีค่าเท่ากับ 2.5 แสดงว่า โอกาสที่จะเกิดเหตุการณ์ที่สนใจเป็น 2.5 เท่าของโอกาสที่จะไม่เกิดเหตุการณ์ที่สนใจ นั่นคือ ถ้า odds มีค่ามากกว่า 1 แสดงว่าโอกาสที่จะเกิดเหตุการณ์ที่สนใจนั้นมากกว่าโอกาสที่จะไม่เกิดเหตุการณ์ที่สนใจ การเขียนตัวแบบโลจิสติคจะอยู่ในรูป log ของ odds เรียกว่า logit หรือ Logistic Response Function ซึ่ง logit เขียนในรูปสมการดังนี้

$$\text{logit} = \log\left(\frac{P_y}{Q_y}\right) = \log\left(\frac{P_y}{1-P_y}\right) = b_0 + b_1x_1 +...+ b_px_p$$

เมื่อได้ log ของ odds หรือ logit แล้ว รูปแบบของตัวแปรตามจึงสามารถทำนายได้ด้วยชุดของตัวแปรอิสระเชิงเส้นตรง สำหรับการทำนายค่า y ที่เป็น P_y ในการวิเคราะห์การถดถอยโลจิสติคจะใช้สมการดังนี้

$$P_y = \frac{e^{b_0 + b_1x_1 +...+ b_px_p}}{1 + e^{b_0 + b_1x_1 +...+ b_px_p}}$$

6) การวิเคราะห์จำแนกกลุ่ม (Discriminant Analysis)

การวิเคราะห์จำแนกกลุ่มเป็นเทคนิคทางสถิติที่ใช้ในการจำแนกกลุ่ม (Classification) ตั้งแต่ 2 กลุ่มขึ้นไป โดยใช้ตัวแปรอิสระตั้งแต่ 1 ตัวขึ้นไป ในการพยากรณ์ค่าของตัวแปรตามจำนวน 1 ตัว ซึ่งตัวแปรตามมีได้หลายค่า แต่ละค่าจะแสดงกลุ่มที่สังกัด โดยตัวแปรอิสระหรือตัวแปรที่ทำให้กลุ่มแตกต่างกันจะเรียกว่า ตัวแปรจำแนกกลุ่ม (Discriminator Variable) ซึ่งเป็นตัวแปรที่นำมาใช้ในการแบ่งกลุ่มพยากรณ์ และควรเป็นตัวแปรเชิงปริมาณ ตัวแปรอิสระที่ทำให้กลุ่มพยากรณ์แตกต่างกันอาจมีเพียง 1 ตัว หรือตั้งแต่ 2 ตัวขึ้นไป ซึ่งในการวิเคราะห์จำแนกกลุ่มจะได้ฟังก์ชันจำแนกกลุ่ม (Discriminant Function) หรือบางครั้งเรียกว่า สมการจำแนกกลุ่ม (กัลยา, 2551ข) ดังนี้

$$\hat{D} = b_0 + b_1x_1 + b_2x_2 + ... + b_px_p$$

เมื่อ $\hat{D}$ คือ Discriminant Score

b_i คือสัมประสิทธิ์ของสมการจำแนกกลุ่ม โดยที่ i = 1, 2, ..., p

x_i คือตัวแปรอิสระ ตัวที่ i โดยที่ i = 1, 2, ..., p

p คือ จำนวนตัวแปรจำแนกกลุ่ม

การประมาณค่าสัมประสิทธิ์ของสมการจำแนกกลุ่มมีเป้าหมายเพื่อทำให้ความแตกต่างระหว่างกลุ่มมีค่ามากที่สุด นั้นคือ ทำให้ค่าไอเก้นมีค่าสูงสุด หรือทำให้มีเปอร์เซ็นต์การจัดกลุ่มผิดมีค่าน้อยที่สุด ซึ่งค่าไอเก้น หรือที่เรียกว่า Discriminant Criterion หรือ Characteristic Roots หรือ Latent Roots เขียนแทนด้วยสัญลักษณ์ λ ซึ่งค่า λ อาจมีได้หลายค่า โดยที่จำนวนของค่า λ จะเท่ากับจำนวนกลุ่ม (k) ลบ ด้วย 1 นั่นคือ k-1 หรือเท่ากับจำนวนตัวแปรจำแนกกลุ่ม (p) แล้วแต่ว่าจำนวนใดจะน้อยกว่ากัน

6.1) วัตถุประสงค์ของการวิเคราะห์จำแนกกลุ่ม (กัลยา, 2551ก) เพื่อหาสาเหตุหรือปัจจัยที่ควรใช้ในการแบ่งกลุ่ม และเพื่อสร้างสมการจำแนกกลุ่ม (Discriminant Function) ที่ดีที่สุดจากข้อมูลที่มีอยู่ รวมทั้งเพื่อนำสมการจำแนกกลุ่มที่ได้มาใช้พยากรณ์หน่วยวิเคราะห์ใหม่ว่าสมควรจัดให้อยู่ในกลุ่มใด

6.2) ข้อสมมติเบื้องต้นของการวิเคราะห์จำแนกกลุ่ม (Hair *et al.*, 2010: p. 245) ได้แก่

6.2.1) ตัวแปรอิสระมีการแจกแจงปรกติหลายตัวแปร (Normality of Independent Variables)

6.2.2) เมตริกซ์ความแปรปรวนร่วมของตัวแปรอิสระของประชากรต้องเท่ากัน (Equal Dispersion Matrices)

6.2.3) ตัวแปรอิสระและตัวแปรตามมีความสัมพันธ์เชิงเส้น (Linearity of Relationships)

6.2.4) ตัวแปรอิสระไม่มีความสัมพันธ์เชิงเส้นแบบพหุ (Multicollinearity)

6.3) ลักษณะข้อมูลและการเตรียมข้อมูลที่จะนำมาวิเคราะห์ มีดังนี้

6.3.1) แบ่งกลุ่มตัวอย่างออกเป็นกลุ่มอย่างน้อย 2 กลุ่ม โดยแต่ละกลุ่มตัวอย่างต้องมีขนาดเท่ากันหรือใกล้เคียงกัน เพื่อให้ผลการวิเคราะห์และการสรุปผลมีความน่าเชื่อถือ (นันทา, 2558)

6.3.2) เลือกตัวแปรอิสระที่คาดว่าจะทำให้กลุ่มที่แบ่งไว้ในข้อ 6.3.1 แตกต่างกัน

6.3.3) ตัวแปรตามเป็นตัวแปรที่แบ่งเป็นกลุ่มๆ ตั้งแต่ 2 กลุ่มขึ้นไป มีระดับการวัดในมาตรานามบัญญัติ (Nominal Scale) หรือมาตราอันดับ (Ordinal Scale) และถ้าหากมีข้อมูลอยู่ในระดับอื่นให้แปลงข้อมูลเป็นตัวแปรเชิงกลุ่มก่อนที่จะนำไปวิเคราะห์

6.3.4) ตัวแปรอิสระเป็นตัวแปรต่อเนื่องที่มีระดับการวัดอยู่ในมาตราแบบช่วง (Interval Scale) หรือมาตราแบบอัตราส่วน (Ratio Scale) และถ้าหากมีข้อมูลอยู่ในระดับอื่นให้แปลงให้เป็นตัวแปรหุ่น (Dummy Coding) ก่อนที่จะนำไปวิเคราะห์

6.4) สถิติสำคัญของการวิเคราะห์จำแนกกลุ่ม (สมบัติ, 2552) ได้แก่

Veridian E-Journal, Science and Technology Silpakorn University
Volume 5 Number 3 May – June 2018 ISSN 2408 - 1248

6.4.1) ค่าไอเก้น (Eigenvalue) เป็นค่าที่แสดงอัตราส่วนการผันแปรระหว่างกลุ่มต่อการผันแปรภายในกลุ่ม ถ้าค่าไอเก้นมีค่าสูง แสดงว่าสมการจำแนกกลุ่มดีหรือมีค่าจำแนกสูงหรือค่าไอเก้นก็คือความแปรปรวนของคะแนนแปลงรูป Y ที่แปลงมาจาก $x_1, x_2, ..., x_p$ นั่นเอง

6.4.2) ค่าสหสัมพันธ์คาโนนิคอล (Canonical Correlation) เป็นสถิติซึ่งสามารถใช้ในการตัดสินความสำคัญของสมการจำแนกกลุ่มเป็นมาตรวัดความสัมพันธ์ของสมการกับกลุ่มของตัวแปร ซึ่งระบุการเป็นสมาชิกของกลุ่มนั้นๆ ของตัวแปรตาม โดยชี้ให้เห็นว่าการเป็นสมาชิกกลุ่มมีความสัมพันธ์กับสมการจำแนกกลุ่มที่หามาได้มากน้อยเพียงใด ดังนั้น ถ้าค่าสหสัมพันธ์คาโนนิคอลมีค่าสูง แสดงว่า การเป็นสมาชิกของกลุ่มสามารถอธิบายความผันแปรของตัวแปรกับสมการจำแนกกลุ่มได้มาก

6.4.3) ค่าวิลค์แลมบ์ดา (Wilks' Lambda) เป็นสถิติที่ใช้ทดสอบความแตกต่างระหว่างกลุ่ม (ทรงศักดิ์, 2551) และเป็นมาตรวัดอำนาจในการจำแนกกลุ่มของตัวแปรด้วย ถ้าค่าวิลค์แลมบ์ดามีค่ามาก (มากกว่า 0.5) ตัวแปรจะอธิบายการเป็นสมาชิกของกลุ่มได้น้อย ถ้าค่าวิลค์แลมบ์ดามีค่าน้อย (น้อยกว่า 0.5) ตัวแปรจะอธิบายการเป็นสมาชิกของกลุ่มได้มาก

6.5 ขั้นตอนของการวิเคราะห์จำแนกกลุ่ม (กัลยา, 2551ข) ประกอบด้วย

6.5.1) พิจารณาหรือกำหนดตัวแปรอิสระที่คาดว่าจะมีผลหรืออิทธิพลต่อการจำแนกกลุ่ม

6.5.2) เลือกตัวอย่างแต่ละกลุ่มเพื่อใช้เป็นตัวแทนของกลุ่ม

6.5.3) เก็บรวบรวมข้อมูลของตัวแปรอิสระที่เลือกไว้ในข้อ 6.5.1

6.5.4) สร้างเกณฑ์หรือสร้างสมการจำแนกกลุ่มโดยใช้ข้อมูลที่เก็บจากข้อ 6.5.2 และ 6.5.3 สำหรับหลักเกณฑ์ที่ใช้ จะนำหลักการของการวิเคราะห์การถดถอยและการวิเคราะห์ความแปรปรวน (ANOVA) โดยการสร้างสมการเชิงเส้นที่เป็นสมการแสดงความสัมพันธ์ระหว่างตัวแปรแบ่งกลุ่มหรือตัวแปรตาม (D) กับตัวแปรอิสระหรือประมาณค่าสัมประสิทธิ์ของสมการจำแนกกลุ่มที่ทำให้ค่าไอเก้นมีค่าสูงสุดหรือทำให้มีเปอร์เซ็นต์การจัดกลุ่มผิดน้อยที่สุด

6.5.5) นำเกณฑ์หรือสมการจำแนกกลุ่มที่ได้จากข้อที่ 6.5.4 มาพยากรณ์กลุ่มของหน่วยวิเคราะห์ใหม่

ผลการวิจัย

1) การประมาณช่วงความเชื่อมั่น 95% ของขนาดประชากรนกประจำถิ่นด้วยข้อมูลจากวิธี Capture-recapture ผลการวิเคราะห์ข้อมูลเป็นดังนี้

1.1) ข้อมูลปฐมภูมิที่เก็บรวบรวมได้จากวิธี Capture-recapture ในเขตห้ามล่าสัตว์ป่าบึงบอระเพ็ด จังหวัดนครสวรรค์ คือ จำนวนนกประจำถิ่นที่จับได้ในครั้งที่ 1 เท่ากับ 72 ตัว จำนวนนกประจำถิ่นที่จับได้ในครั้งที่ 2 เท่ากับ 62 ตัว และจำนวนนกประจำถิ่นที่จับได้ในครั้งที่ 3 เท่ากับ 61 ตัว โดยสถานีแหลมตาเส็งพบนกประจำถิ่นมากที่สุด คือ 68 ตัว รองลงมาคือ สถานีทางเข้าประมงเชิงพาณิชย์ 33 ตัว สถานีเนินระฆัง 29 ตัว สถานีแหลมนา 26 ตัว สถานีอุทยานนกน้ำ (สวนบัว) 23 ตัว และสถานีฝายเก่าพบนกประจำถิ่นน้อยที่สุด คือ 16 ตัว โดยที่ $f_1 = 190$, $f_2 = 5$, $f_3 = 0$ และ $n = f_1 + f_2 + f_3 = 195$ ซึ่งชนิดของนกประจำถิ่นที่เก็บรวบรวมได้จากวิธี Capture-recapture ประกอบด้วยนกประจำถิ่นจำนวน 19 ชนิด ได้แก่ นกเอี้ยงหงอน นกปรอดสวน

นกกระจาบธรรมดา นกกระจิบหญ้าสีเรียบ นกอีวาบตั๊กแตน นกกระจิบธรรมดา นกกระจาบทอง นกกระจิบหญ้าท้องเหลือง นกยางไฟหัวดำ นกกระติ๊ดสีอิฐ นกกระจาบอกลาย กระจาบธรรมดา นกกระจอกตาล นกอีแพรดแถบอกดำ นกจาบคาหัวเขียว นกกระเต็นอกขาว นกยางดำ นกหางนาค และนกหัวขวานด่างอกลายจุด

1.2) คำนวณค่าประมาณแบบจุด ค่าประมาณความคลาดเคลื่อนมาตรฐาน และช่วงความเชื่อมั่น 95% ของตัวประมาณขนาดประชากรนกประจำถิ่นทั้ง 3 ตัว ผลการวิเคราะห์แสดงดังตารางที่ 1

ตารางที่ 1 ค่าสถิติของตัวประมาณขนาดประชากรนกประจำถิ่นด้วยข้อมูลจากวิธี Capture-recapture

ตัวประมาณ	ค่าประมาณแบบจุด	ค่าประมาณความคลาดเคลื่อนมาตรฐาน	ค่าประมาณแบบช่วง (95%)
Zelterman	3803.355	1687.221	(496.402, 7110.308)
Chao	3802.000	1674.320	(2234.005, 6575.783)
Weighted	3805.000	1677.538	(517.026, 7092.974)

จากตารางที่ 1 พบว่า ตัวประมาณแบบถ่วงน้ำหนักเป็นตัวประมาณขนาดประชากรนกประจำถิ่นที่มีค่าประมาณแบบจุดสูงที่สุด โดยมีค่าเท่ากับ 3,805.000 รองลงมาคือ ตัวประมาณของ Zelterman มีค่าเท่ากับ 3,803.355 และตัวประมาณของ Chao เป็นตัวประมาณที่มีค่าประมาณแบบจุดต่ำที่สุด โดยมีค่าเท่ากับ 3,802.000

สำหรับค่าประมาณความคลาดเคลื่อนมาตรฐานของตัวประมาณขนาดประชากรนกประจำถิ่น พบว่า ตัวประมาณของ Chao มีค่าประมาณความคลาดเคลื่อนมาตรฐานต่ำที่สุด คือ 1,674.320 ตัวประมาณแบบถ่วงน้ำหนักมีค่าประมาณความคลาดเคลื่อนมาตรฐานเท่ากับ 1,677.538 และตัวประมาณของ Zelterman มีค่าประมาณความคลาดเคลื่อนมาตรฐานสูงที่สุด คือ 1,687.221

สำหรับค่าประมาณขนาดประชากรนกประจำถิ่นที่ระดับความเชื่อมั่น 95% พบว่า ตัวประมาณของ Zelterman มีช่วงความเชื่อมั่น 95% ของขนาดประชากรนกประจำถิ่นอยู่ระหว่าง 496.402 ถึง 7,110.308 ตัวประมาณของ Chao มีช่วงความเชื่อมั่น 95% ของขนาดประชากรนกประจำถิ่นอยู่ระหว่าง 2,234.005 ถึง 6,575.783 และตัวประมาณแบบถ่วงน้ำหนักมีช่วงความเชื่อมั่น 95% ของขนาดประชากรนกประจำถิ่นอยู่ระหว่าง 517.026 ถึง 7,092.974

2) การวิเคราะห์การถดถอยโลจิสติค

2.1) ตรวจสอบข้อสมมติเบื้องต้นของการวิเคราะห์การถดถอยโลจิสติค

2.1.1) เริ่มจากการสร้างกรอบแนวคิดการวิจัยและกำหนดระดับการวัดของตัวแปรอิสระและตัวแปรตาม โดยตัวแปรตาม คือ เพศนกกระติ๊ดแดง แบ่งออกเป็น 2 กลุ่ม ได้แก่ เพศผู้มีค่าเท่ากับ 0 และเพศเมียมีค่าเท่ากับ 1 และตัวแปรอิสระที่พิจารณามี 6 ตัว ได้แก่ ความยาวปีก (x_1), ความยาวท้ายทอยถึงปลายปาก (x_2), ความยาวปาก (x_3), ความยาวแข้ง (x_4), ความยาวหาง (x_5) และน้ำหนัก (x_6)

2.1.2) ค่าเฉลี่ยของความคลาดเคลื่อนสุ่มเป็นศูนย์ โดยตรวจสอบจากสถิติทดสอบ Z พบว่าค่าพี (P-value) เท่ากับ 1.000 ซึ่งมีค่ามากกว่าระดับนัยสำคัญ 0.05 จึงยอมรับสมมติฐานหลัก แสดงว่า ค่าเฉลี่ยของความคลาดเคลื่อนสุ่มมีค่าเท่ากับศูนย์อย่างมีนัยสำคัญทางสถิติที่ระดับนัยสำคัญ 0.05

Veridian E-Journal, Science and Technology Silpakorn University
Volume 5 Number 3 May – June 2018 ISSN 2408 - 1248

2.1.3) ตัวแปรอิสระต้องไม่มีความสัมพันธ์กันหรือไม่เกิดปัญหา Multicollinearity ในงานวิจัยนี้พิจารณาจากค่า Tolerance และค่า Variance Inflation Factor (VIF) ซึ่งผลการวิเคราะห์แสดงดังตารางที่ 2

ตารางที่ 2 ค่า Tolerance และค่า Variance Inflation Factor (VIF) ของตัวแปรอิสระ

ตัวแปรอิสระ	Tolerance	Variance Inflation Factor (VIF)
x_1 (ความยาวปีก)	0.763	1.311
x_2 (ความยาวท้ายทอยถึงปลายปาก)	0.723	1.383
x_3 (ความยาวปาก)	0.748	1.337
x_4 (ความยาวแข้ง)	0.752	1.330
x_5 (ความยาวหาง)	0.827	1.209
x_6 (น้ำหนัก)	0.710	1.408

จากตารางที่ 2 แสดงค่า Tolerance ของตัวแปรความยาวปีก (x_1), ความยาวท้ายทอยถึงปลายปาก (x_2), ความยาวปาก (x_3), ความยาวแข้ง (x_4), ความยาวหาง (x_5) และน้ำหนัก (x_6) ซึ่งพบว่ามีค่าเข้าใกล้ 1 นอกจากนี้ค่า VIF ของตัวแปรความยาวปีก (x_1), ความยาวท้ายทอยถึงปลายปาก (x_2), ความยาวปาก (x_3), ความยาวแข้ง (x_4), ความยาวหาง (x_5) และน้ำหนัก (x_6) พบว่าไม่ได้มีค่าเข้าใกล้ 10 จึงสรุปได้ว่า ตัวแปรอิสระทั้ง 6 ตัว ไม่มีความสัมพันธ์กันหรือไม่เกิดปัญหา Multicollinearity

2.1.4) การวิเคราะห์การถดถอยโลจิสติคจะต้องใช้ขนาดตัวอย่าง (n) มากกว่าการวิเคราะห์การถดถอยแบบปกติ โดยจะใช้ขนาดตัวอย่าง (n) คือ n ไม่น้อยกว่า 30p โดยที่ p คือ จำนวนตัวแปรอิสระ สำหรับงานวิจัยนี้มีจำนวนตัวแปรอิสระ 6 ตัว และ n เท่ากับ 186 ดังนั้น 186 ไม่น้อยกว่า 30(6) ซึ่งเป็นไปตามข้อตกลงเบื้องต้น

2.2) การเลือกตัวแปรอิสระเข้าสมการถดถอยโลจิสติค

2.2.1) แบ่งชุดข้อมูลออกเป็น 2 ส่วน ส่วนแรกเป็นชุดข้อมูลที่นำมาใช้สร้างสมการถดถอยโลจิสติค (ร้อยละ 80) และส่วนที่สองเป็นชุดข้อมูลที่นำมาใช้ตรวจสอบความถูกต้องของสมการถดถอยโลจิสติค (ร้อยละ 20)

2.2.2) จากการตรวจสอบจำนวนนกกระติ๊ดแดงทั้งหมด 186 ตัว พบว่า เป็นเพศผู้ 93 ตัว และเป็นเพศเมีย 93 ตัว ในการเลือกตัวแปรอิสระเข้าสู่สมการถดถอยโลจิสติคใช้วิธี Forward Stepwise และทำการทดสอบความสัมพันธ์ระหว่างชุดตัวแปรอิสระกับตัวแปรตาม ผลการวิเคราะห์แสดงดังตารางที่ 3

ตารางที่ 3 ผลการทดสอบความสัมพันธ์ระหว่างชุดตัวแปรอิสระกับตัวแปรตาม

	-2Log Likelihood (-2LL)	Chi-Square	df	P-value
Baseline	221.098			
Model	160.211	60.887	5	0.000

จากตารางที่ 3 แสดงผลการทดสอบความสัมพันธ์ระหว่างชุดตัวแปรอิสระกับการเป็นนกกระติ๊ดแดงเพศเมีย โดยพิจารณาจากข้อมูลส่วนที่ 1 ซึ่งพบว่า ถ้าสมการถดถอยโลจิสติคประกอบด้วยค่าคงที่เพียงอย่างเดียวจะให้ค่า -2LL เท่ากับ 221.098 และถ้าสมการถดถอยโลจิสติคประกอบด้วยค่าคงที่และชุดตัวแปรอิสระจะให้ค่า -2LL เท่ากับ 160.211 แสดงว่าสมการถดถอยโลจิสติคที่มีชุดตัวแปรอิสระรวมอยู่ด้วยมีความเหมาะสมมากกว่า

สมการถดถอยโลจิสติคที่มีค่าคงที่เพียงอย่างเดียว นั่นคือ มีตัวแปรอิสระอย่างน้อย 1 ตัว มีความสัมพันธ์กับตัวแปรตามอย่างมีนัยสำคัญทางสถิติที่ระดับนัยสำคัญ 0.05 (P-value = 0.000 < 0.05)

2.2.3) การทดสอบความมีนัยสำคัญของสัมประสิทธิ์การถดถอยแต่ละตัวของสมการถดถอยโลจิสติค เมื่อกำหนดให้ตัวแปรตาม คือ เพศนกกระติ๊ดแดงที่เป็นเพศเมียมีค่าเท่ากับ 1 (y=1) จะได้ผลการวิเคราะห์แสดงดังตารางที่ 4

ตารางที่ 4 ผลการทดสอบสัมประสิทธิ์การถดถอยแต่ละตัวของสมการถดถอยโลจิสติค

	b	S.E.	Wald-test	df	P-value
Constant	-58.683	16.638	12.439	1	0.000*
x_1 (ความยาวปีก)	0.570	0.177	10.415	1	0.001*
x_2 (ความยาวท้ายทอยถึงปลายปาก)	0.939	0.400	5.494	1	0.019*
x_4 (ความยาวแข้ง)	-1.327	0.309	18.447	1	0.000*
x_5 (ความยาวหาง)	1.340	0.262	26.087	1	0.000*
x_6 (น้ำหนัก)	-2.180	0.703	9.620	1	0.002*

* คือ P-value < 0.05 แสดงว่ามีนัยสำคัญทางสถิติที่ระดับ 0.05

จากตารางที่ 4 แสดงการทดสอบสัมประสิทธิ์การถดถอยแต่ละตัวที่มีผลต่อการทำนายเพศนกกระติ๊ดแดง ซึ่งพบว่า ความยาวปีก (x_1), ความยาวท้ายทอยถึงปลายปาก (x_2), ความยาวแข้ง (x_4), ความยาวหาง (x_5) และน้ำหนัก (x_6) มีค่าพี (P-value) น้อยกว่า 0.05 จึงปฏิเสธสมมติฐานหลัก นั่นคือ ตัวแปรอิสระทั้ง 5 ตัว มีความสัมพันธ์กับการทำนายเพศนกกระติ๊ดแดงที่ระดับนัยสำคัญ 0.05 ดังนั้นสมการถดถอยโลจิสติคในการพยากรณ์ความน่าจะเป็นสำหรับการทำนายเพศนกกระติ๊ดแดง โดยมีกลุ่มนกกระติ๊ดแดงที่เป็นเพศผู้เป็นระดับอ้างอิงของตัวแปรตาม คือ

$$\text{logit} = \log\left(\frac{P_y}{Q_y}\right) = \log\left(\frac{P_y}{1-P_y}\right) = -58.683 + 0.570x_1 + 0.939x_2 - 1.327x_4 + 1.340x_5 - 2.180x_6$$

เมื่อ P_y คือ ความน่าจะเป็นที่นกกระติ๊ดแดงจะเป็นเพศเมียหรือ y=1

2.3) การตรวจสอบความถูกต้องของสมการถดถอยโลจิสติค ซึ่งสมการถดถอยโลจิสติคที่ได้มีค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) เท่ากับ 0.142 เมื่อตรวจสอบความถูกต้องของสมการถดถอยโลจิสติคโดยใช้ข้อมูลส่วนที่ 2 ผลการวิเคราะห์แสดงดังตารางที่ 5

ตารางที่ 5 ผลการตรวจสอบความถูกต้องของสมการถดถอยโลจิสติคโดยใช้ข้อมูลส่วนที่ 2

เพศนกกระติ๊ดแดง	**จำนวนที่สังเกตได้ (Observed Frequency)**	**ผลการทำนายเพศนกกระติ๊ดแดง**	
		ทำนายถูก (%)	**ทำนายผิด (%)**
เพศผู้	25	23 (92.00)	2 (8.00)
เพศเมีย	21	19 (90.48)	2 (9.52)
รวม	46	42 (91.30)	4 (8.70)

Veridian E-Journal, Science and Technology Silpakorn University
Volume 5 Number 3 May – June 2018 ISSN 2408 - 1248

จากตารางที่ 5 แสดงผลการตรวจสอบความถูกต้องของสมการถดถอยโลจิสติค โดยใช้ข้อมูลส่วนที่ 2 พบว่า ในกลุ่มนกเพศผู้เดิมมี 25 ตัว แต่จากการทำนายโดยใช้สมการถดถอยโลจิสติค พบว่า ทำนายได้ถูกต้อง 23 ตัว คิดเป็นร้อยละ 92.00 ส่วนในกลุ่มนกเพศเมียเดิมมี 21 ตัว แต่จากการทำนายโดยใช้สมการถดถอยโลจิสติค พบว่า ทำนายได้ถูกต้อง 19 ตัว คิดเป็นร้อยละ 90.48 เมื่อคิดรวมนกกระติ๊ดแดงทั้งหมด 46 ตัว พบว่าสมการถดถอยโลจิสติคสามารถทำนายได้ถูกต้องคิดเป็นร้อยละ 91.30

3) การวิเคราะห์จำแนกกลุ่ม

3.1) ตรวจสอบข้อสมมติเบื้องต้นของการวิเคราะห์จำแนกกลุ่ม

3.1.1) ตัวแปรอิสระต้องมีการแจกแจงปรกติหลายตัวแปร ซึ่งทำการตรวจสอบโดยพิจารณาจาก ข้อมูลค่านอกเกณฑ์หลายตัวแปร (Multivariate Outliers) จากการวิเคราะห์ Mahalanobis Distances พบว่า ค่า Mahalanobis Distances เท่ากับ 9.782 มีค่าน้อยกว่าเมื่อนำไปเปรียบเทียบกับค่าไคกำลังสองที่องศาเสรีเท่ากับ 6 คือ 12.592 แสดงว่า มีความน่าจะเป็นที่ตัวแปรอิสระจะมีการแจกแจงปรกติหลายตัวแปร

3.1.2) เมตริกซ์ความแปรปรวนร่วมของตัวแปรอิสระของประชากรต้องเท่ากัน พิจารณาจากค่า Box's M พบว่าค่าพี (P-value) เท่ากับ 0.000 ซึ่งน้อยกว่าระดับนัยสำคัญ 0.05 แสดงว่าความแปรปรวนร่วมของประชากรทุกกลุ่มไม่เท่ากันที่ระดับนัยสำคัญ 0.05 ซึ่งไม่เป็นไปตามข้อตกลงเบื้องต้น ดังนั้นในการวิเคราะห์ข้อมูลด้วยโปรแกรม SPSS จะใช้คำสั่ง Separated Groups เพื่อให้มีการแยกวิเคราะห์ความแปรปรวนร่วมในแต่ละกลุ่ม

3.1.3) ตัวแปรอิสระและตัวแปรตามต้องมีความสัมพันธ์เชิงเส้น โดยตรวจสอบจากสัมประสิทธิ์สหสัมพันธ์พอยท์ไบซีเรียล (Point Biserial Correlation Coefficient) (สารภี, 2556) เมื่อพิจารณาจากค่าพี (P-value) พบว่า ตัวแปรความยาวท้ายทอยถึงปลายปาก (x_2) ไม่มีความสัมพันธ์เชิงเส้นกับตัวแปรเพศนกกระติ๊ดแดง ที่ระดับนัยสำคัญ 0.05 ดังนั้น ในการวิเคราะห์จึงตัดตัวแปรความยาวท้ายทอยถึงปลายปาก (x_2) ออกจากการวิเคราะห์

3.1.4) ตัวแปรอิสระไม่มีความสัมพันธ์เชิงเส้นแบบพหุ โดยพิจารณาจากค่า Tolerance และค่า Variance Inflation Factor (VIF) ซึ่งผลการวิเคราะห์แสดงดังตารางที่ 6

ตารางที่ 6 ค่า Tolerance และค่า Variance Inflation Factor (VIF) ของตัวแปรอิสระ

ตัวแปรอิสระ	Tolerance	Variance Inflation Factor (VIF)
x_1 (ความยาวปีก)	0.861	1.162
x_3 (ความยาวปาก)	0.896	1.116
x_4 (ความยาวแข้ง)	0.755	1.324
x_5 (ความยาวหาง)	0.872	1.147
x_6 (น้ำหนัก)	0.727	1.376

จากตารางที่ 6 แสดงค่า Tolerance ของตัวแปรความยาวปีก (x_1), ความยาวปาก (x_3), ความยาวแข้ง (x_4), ความยาวหาง (x_5) และน้ำหนัก (x_6) ซึ่งพบว่ามีค่าเข้าใกล้ 1 นอกจากนี้ค่า VIF ของตัวแปรความยาวปีก (x_1), ความยาวปาก (x_3), ความยาวแข้ง (x_4), ความยาวหาง (x_5) และน้ำหนัก (x_6) พบว่าไม่ได้มีค่าเข้าใกล้ 10 จึงสรุปได้ว่า ตัวแปรอิสระทั้ง 5 ตัว ไม่มีความสัมพันธ์เชิงเส้นแบบพหุ (Multicollinearity)

3.2) การสร้างสมการจำแนกกลุ่ม

3.2.1) แบ่งชุดข้อมูลออกเป็น 2 ส่วน ส่วนแรกเป็นชุดข้อมูลที่นำมาใช้สร้างสมการจำแนกกลุ่ม (ร้อยละ 80) และส่วนที่สองเป็นชุดข้อมูลที่นำมาใช้ตรวจสอบความถูกต้องของสมการจำแนกกลุ่ม (ร้อยละ 20)

3.2.2) การทดสอบความเท่ากันของค่าเฉลี่ยของตัวแปรอิสระแต่ละกลุ่ม คือ กลุ่มนกเพศผู้และกลุ่มนกเพศเมีย โดยพิจารณาจากสถิติทดสอบ Wilks' Lambda ซึ่งผลการวิเคราะห์แสดงดังตารางที่ 7

ตารางที่ 7 ค่าสถิติการทดสอบความเท่ากันของค่าเฉลี่ยของตัวแปรอิสระแต่ละกลุ่ม

ตัวแปรอิสระ	Wilks' Lambda	F-test	P-value
x_1 (ความยาวปีก)	0.982	3.303	0.071
x_3 (ความยาวปาก)	0.970	5.624	0.019*
x_4 (ความยาวแข้ง)	0.816	41.514	0.000*
x_5 (ความยาวหาง)	0.826	38.661	0.000*
x_6 (น้ำหนัก)	0.901	20.229	0.000*

* คือ P-value < 0.05 แสดงว่ามีนัยสำคัญทางสถิติที่ระดับ 0.05

จากตารางที่ 7 แสดงค่าสถิติการทดสอบค่าเฉลี่ยของตัวแปรอิสระต่างๆ ที่แบ่งกลุ่มตัวแปรตาม โดยใช้สถิติทดสอบ Wilks' Lamda พบว่า มีค่าเฉลี่ยของตัวแปรความยาวปากใน 2 กลุ่มแตกต่างกัน (P-value = 0.019) รวมทั้งตัวแปรความยาวแข้ง (P-value = 0.000) ตัวแปรความยาวหาง (P-value = 0.000) และตัวแปรน้ำหนัก (P-value = 0.000) ที่มีค่าเฉลี่ยใน 2 กลุ่มแตกต่างกันที่ระดับนัยสำคัญ 0.05 สำหรับค่าเฉลี่ยของตัวแปรความยาวปีกใน 2 กลุ่มไม่แตกต่างกันที่ระดับนัยสำคัญ 0.05 (P-value = 0.071)

3.2.3) การนำเสนอค่าไอเก้นและค่าสหสัมพันธ์คาโนนิคอลของการตรวจสอบสมการจำแนกกลุ่ม ผลการวิเคราะห์แสดงดังตารางที่ 8

ตารางที่ 8 ค่าไอเก้นและค่าสหสัมพันธ์คาโนนิคอลของการตรวจสอบสมการจำแนกกลุ่ม

Function	Eigenvalue	% of Variance	Canonical Correlation
D (ค่าคะแนนจำแนกกลุ่ม)	2.670	100.0	0.933

จากตารางที่ 8 แสดงค่าไอเก้นและค่าสหสัมพันธ์คาโนนิคอล เนื่องจากตัวแปรตาม คือ เพศนกกระติ๊ดแดงมี 2 กลุ่ม จึงทำให้มีค่าไอเก้นเพียงค่าเดียว คือ เท่ากับ 2.670 ถือว่ามีค่าค่อนข้างสูง แสดงว่าสมการจำแนกกลุ่มสามารถจำแนกได้ค่อนข้างดี และมีค่าสหสัมพันธ์คาโนนิคอลเท่ากับ 0.933 ซึ่งถือว่าค่อนข้างสูง แสดงว่าสมการจำแนกกลุ่มนั้นสามารถจำแนกประชากรออกเป็นกลุ่มๆ ได้ดี เมื่อนำค่าสหสัมพันธ์คาโนนิคอลมายกกำลังสองจะได้เท่ากับ 0.8705 ซึ่งค่าที่ได้แสดงให้เห็นว่าตัวแปรอิสระในสมการจำแนกกลุ่มสามารถอธิบายความแปรปรวนของตัวแปรตามได้ร้อยละ 87.05

Veridian E-Journal, Science and Technology Silpakorn University
Volume 5 Number 3 May – June 2018 ISSN 2408 - 1248

3.2.4) การทดสอบความมีนัยสำคัญของสมการจำแนกกลุ่ม ผลการวิเคราะห์แสดงดังตารางที่ 9

ตารางที่ 9 ค่าสถิติการทดสอบความมีนัยสำคัญของสมการจำแนกกลุ่ม

Test of Function(s)	Wilks' Lambda	Chi-square	df	P-value
D (ค่าคะแนนจำแนกกลุ่ม)	0.399	93.360	4	0.000*

* คือ P-value < 0.05 แสดงว่ามีนัยสำคัญทางสถิติที่ระดับ 0.05

จากตารางที่ 9 แสดงค่าสถิติการทดสอบความมีนัยสำคัญของสมการจำแนกกลุ่ม พบว่าค่าวิลค์แลมบ์ดามีค่าเท่ากับ 0.399 ถือว่ามีค่าน้อย (น้อยกว่า 0.5) แสดงว่าสมการจำแนกกลุ่มที่ได้จะเป็นสมการจำแนกกลุ่มที่ดี สำหรับค่าไคกำลังสองมีค่าเท่ากับ 93.360 และค่าพี (P-value) เท่ากับ 0.000 ซึ่งมีค่าน้อยกว่าระดับนัยสำคัญ 0.05 แสดงว่า สมการจำแนกกลุ่มที่ได้สามารถจำแนกออกเป็น 2 กลุ่มได้อย่างมีนัยสำคัญทางสถิติที่ระดับนัยสำคัญ 0.05

3.2.5) ค่าสัมประสิทธิ์ของสมการจำแนกกลุ่ม ผลการวิเคราะห์แสดงดังตารางที่ 10

ตารางที่ 10 ค่าสัมประสิทธิ์ของสมการจำแนกกลุ่ม

	เพศนกกระติ๊ดแดง	
	เพศผู้	**เพศเมีย**
Constant	-1,020.309	-995.903
x_1 (ความยาวปีก)	25.241	24.764
x_4 (ความยาวแข้ง)	6.625	8.140
x_5 (ความยาวหาง)	23.751	22.668
x_6 (น้ำหนัก)	5.170	7.259

จากตารางที่ 10 แสดงค่าสัมประสิทธิ์ของสมการจำแนกกลุ่มโดยแยกเป็นกลุ่มตามวิธีของ Fisher's ซึ่งจำนวนสมการจำแนกกลุ่มจะมีเท่ากับจำนวนกลุ่ม กล่าวคือ 2 สมการ และสามารถเขียนสมการจำแนกกลุ่มเพศนกได้ดังนี้

สมการจำแนกของกลุ่มนกเพศผู้ คือ $\hat{D}_1 = -1{,}020.309 + 25.241x_1 + 6.625x_4 + 23.751x_5 + 5.170x_6$

สมการจำแนกของกลุ่มนกเพศเมีย คือ $\hat{D}_2 = -995.903 + 24.764x_1 + 8.140x_4 + 22.668x_5 + 7.259x_6$

สำหรับการพยากรณ์หน่วยวิเคราะห์ว่าควรจัดอยู่ในกลุ่มใด สามารถทำได้โดยการแทนค่าของตัวแปรอิสระ ทั้ง 4 ตัวของหน่วยวิเคราะห์นั้นๆ ลงในสมการจำแนกกลุ่มทั้ง 2 สมการ ถ้าสมการจำแนกกลุ่มใดมีค่ามากกว่าก็จัดหน่วยวิเคราะห์นั้นๆ อยู่ในกลุ่มนั้น

3.3) การตรวจสอบความถูกต้องของสมการจำแนกกลุ่ม ซึ่งสมการจำแนกกลุ่มที่ได้มีค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) เท่ากับ 0.169 เมื่อตรวจสอบความถูกต้องของสมการจำแนกกลุ่มโดยใช้ข้อมูลส่วนที่ 2 ผลการวิเคราะห์แสดงดังตารางที่ 11

ตารางที่ 11 ผลการตรวจสอบความถูกต้องของสมการจำแนกกลุ่มโดยใช้ข้อมูลส่วนที่ 2

เพศนกกระติ๊ดแดง	จำนวนที่สังเกตได้ (Observed Frequency)	ผลการทำนายเพศนกกระติ๊ดแดง	
		ทำนายถูก (%)	ทำนายผิด (%)
เพศผู้	25	22 (88.00)	3 (12.00)
เพศเมีย	21	18 (85.71)	3 (14.29)
รวม	46	40 (86.96)	6 (13.04)

จากตารางที่ 11 แสดงผลการตรวจสอบความถูกต้องของสมการจำแนกกลุ่ม โดยใช้ข้อมูลส่วนที่ 2 พบว่า ในกลุ่มนกเพศผู้เดิมมี 25 ตัว แต่จากการทำนายโดยใช้สมการจำแนกกลุ่ม พบว่า ทำนายได้ถูกต้อง 22 ตัว คิดเป็นร้อยละ 88.00 ส่วนในกลุ่มนกเพศเมียเดิมมี 21 ตัว แต่จากการทำนายโดยใช้สมการจำแนกกลุ่ม พบว่า ทำนายได้ถูกต้อง 18 ตัว คิดเป็นร้อยละ 85.71 เมื่อคิดรวมนกกระติ๊ดแดงทั้งหมด 46 ตัว พบว่าสมการจำแนกกลุ่มสามารถทำนายได้ถูกต้องคิดเป็นร้อยละ 86.96

สรุปผลและอภิปรายผล

1) การประมาณช่วงความเชื่อมั่น 95% ของขนาดประชากรนกประจำถิ่นในเขตห้ามล่าสัตว์ป่าบึงบอระเพ็ด จังหวัดนครสวรรค์ ด้วยข้อมูลจากวิธี Capture-recapture ผลการวิเคราะห์พบว่า ตัวประมาณแบบถ่วงน้ำหนักให้ค่าประมาณแบบจุดสูงที่สุดคือ 3,805.000 ซึ่งมีค่าประมาณความคลาดเคลื่อนมาตรฐานเท่ากับ 1,677.538 และมีช่วงความเชื่อมั่น 95% อยู่ระหว่าง 517.026 ถึง 7,092.974 สำหรับตัวประมาณของ Zelterman และตัวประมาณของ Chao ให้ค่าประมาณแบบจุดใกล้เคียงกับตัวประมาณแบบถ่วงน้ำหนัก โดยมีค่าเท่ากับ 3,803.355 และ 3,802.000 ตามลำดับ แต่ตัวประมาณของ Chao มีความกว้างของช่วงความเชื่อมั่น 95% แคบกว่าตัวประมาณของ Zelterman และตัวประมาณแบบถ่วงน้ำหนัก ดังนั้น ในการประมาณขนาดประชากรนกประจำถิ่นดังกล่าวนี้ ถ้าเป็นการประมาณแบบจุดและต้องการนำค่าประมาณแบบจุดไปใช้ในการอนุมานเรื่องอื่นๆ ต่อไป ควรใช้ตัวประมาณของ Chao เนื่องจากมีค่าประมาณความคลาดเคลื่อนมาตรฐานต่ำที่สุด หรือถ้าเป็นการประมาณแบบช่วงที่ระดับความเชื่อมั่น 95% ควรเลือกใช้ช่วงความเชื่อมั่น 95% ของ Chao เนื่องจากมีความกว้างของช่วงความเชื่อมั่นแคบที่สุดเมื่อเปรียบเทียบกับวิธีอื่นๆ ที่ได้ศึกษาในงานวิจัยนี้

2) การสร้างตัวแบบการทำนายเพศนกประจำถิ่นในเขตห้ามล่าสัตว์ป่าบึงบอระเพ็ด จังหวัดนครสวรรค์ สำหรับงานวิจัยนี้เป็นการสร้างตัวแบบทำนายเพศนกกระติ๊ดแดง ซึ่งเป็นนกประจำถิ่นชนิดหนึ่ง ที่พบในเขตห้ามล่าสัตว์ป่าบึงบอระเพ็ด โดยใช้ตัวแปรอิสระ 6 ตัว ได้แก่ ความยาวปีก (x_1), ความยาวท้ายทอยถึงปลายปาก (x_2), ความยาวปาก (x_3), ความยาวแข้ง (x_4), ความยาวหาง (x_5) และน้ำหนัก (x_6) ซึ่งใช้วิธีการทางสถิติในการสร้างตัวแบบ 2 วิธี ได้แก่ การวิเคราะห์การถดถอยโลจิสติค (Logistic Regression Analysis) และการวิเคราะห์จำแนกกลุ่ม (Discriminant Analysis) ผลการวิเคราะห์พบว่า การวิเคราะห์การถดถอย โลจิสติคมีอำนาจจำแนกสูงกว่าการวิเคราะห์จำแนกกลุ่ม ซึ่งสามารถจำแนกกลุ่มนกกระติ๊ดแดงเพศผู้ได้ถูกต้องร้อยละ 92.00 และสามารถจำแนกกลุ่มนกกระติ๊ดแดงเพศเมียได้ถูกต้องร้อยละ 90.48 โดยภาพรวมสามารถจำแนกกลุ่มได้ถูกต้องร้อยละ 91.30 และในการคัดเลือกตัวแปรอิสระเข้าสู่สมการถดถอยโลจิสติคโดยวิธี Forward Stepwise ได้ตัวแปรอิสระ

Veridian E-Journal, Science and Technology Silpakorn University
Volume 5 Number 3 May – June 2018 ISSN 2408 - 1248

จำนวน 5 ตัว ได้แก่ ความยาวปีก (x_1), ความยาวท้ายทอยถึงปลายปาก (x_2), ความยาวแข้ง (x_4), ความยาวหาง (x_5) และน้ำหนัก (x_6) โดยสามารถสร้างสมการถดถอยโลจิสติคได้ดังนี้

$$\text{logit} = \log\left(\frac{P_y}{Q_y}\right) = \log\left(\frac{P_y}{1-P_y}\right) = -58.683 + 0.570x_1 + 0.939x_2 - 1.327x_4 + 1.340x_5 - 2.180x_6$$

สำหรับการวิเคราะห์จำแนกกลุ่ม พบว่า การวิเคราะห์จำแนกกลุ่มมีอำนาจจำแนกน้อยกว่าการวิเคราะห์การถดถอยโลจิสติคเพียงเล็กน้อย กล่าวคือ การวิเคราะห์จำแนกกลุ่มสามารถจำแนกกลุ่มนกกระติ๊ดแดงเพศผู้ได้ถูกต้องร้อยละ 88.00 และสามารถจำแนกกลุ่มนกกระติ๊ดแดงเพศเมียได้ถูกต้องร้อยละ 85.71 โดยภาพรวมสามารถจำแนกกลุ่มได้ถูกต้องร้อยละ 86.96 ซึ่งการวิเคราะห์จำแนกกลุ่ม ตัวแปรอิสระที่ถูกคัดเลือกเข้าสมการจำแนกกลุ่มโดยวิธี Stepwise มีจำนวน 4 ตัว ได้แก่ ความยาวปีก (x_1), ความยาวแข้ง (x_4), ความยาวหาง (x_5) และน้ำหนัก (x_6) โดยสามารถแบ่งกลุ่มเพศนกกระติ๊ดแดงด้วยสมการจำแนกกลุ่มซึ่งแยกเป็นกลุ่มตามวิธีของ Fisher's และสามารถเขียนสมการจำแนกกลุ่มได้ดังนี้

สมการจำแนกของกลุ่มนกเพศผู้ คือ $\hat{D}_1 = -1{,}020.309 + 25.241x_1 + 6.625x_4 + 23.751x_5 + 5.170x_6$

สมการจำแนกของกลุ่มนกเพศเมีย คือ $\hat{D}_2 = -995.903 + 24.764x_1 + 8.140x_4 + 22.668x_5 + 7.259x_6$

ดังนั้น ในการจำแนกเพศนกประจำถิ่นควรใช้การวิเคราะห์การถดถอยโลจิสติคเนื่องจากมีอำนาจจำแนกได้ถูกต้องสูงกว่าการวิเคราะห์จำแนกกลุ่ม

เอกสารอ้างอิง

ภาษาไทย

กัลยา วานิชย์บัญชา. (2551ก). การวิเคราะห์ข้อมูลหลายตัวแปร. พิมพ์ครั้งที่ 3. โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย, จุฬาลงกรณ์มหาวิทยาลัย, กรุงเทพฯ.

_____ . (2551ข). การวิเคราะห์สถิติขั้นสูงด้วย SPSS for Windows. พิมพ์ครั้งที่ 6. โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย, จุฬาลงกรณ์มหาวิทยาลัย, กรุงเทพฯ.

ไชยรัตน์ เบิร์ดฟาร์ม. การแยกเพศนก. แหล่งที่มา: http://chairatbirdfarm.blogspot.com/2009/12/separate-sex-of-birds.html, 27 พฤศจิกายน 2560.

ทรงศักดิ์ ภูสีอ่อน. (2551). การประยุกต์ใช้ SPSS วิเคราะห์ข้อมูลงานวิจัย. พิมพ์ครั้งที่ 2. ประสานการพิมพ์, กาฬสินธุ์.

ธงชัย งามประเสริฐวงศ์, นิพาดา เรือนแก้ว ดิษยทัต และ กำธร ธีรคุปต์. คู่มือสื่อการสอนวิชาชีววิทยา. สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน และคณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย. กรุงเทพฯ.

นันทา คุมคณะ. (2558). “การระบุเพศในประชากรไทย โดยวิเคราะห์จากการวัดกระดูกขากรรไกรล่าง.” **Veridian E-Journal, Science and Technology Silpakorn University.** ปีที่ 2, ฉบับที่ 1 (มกราคม-มิถุนายน): 51 – 64.

ประชุม สุวัตถี. (2552). การสำรวจตัวอย่าง: การชักตัวอย่างและการวิเคราะห์. พิมพ์ครั้งที่ 1. โครงการส่งเสริมและพัฒนาเอกสารวิชาการ สถาบันบัณฑิตพัฒนบริหารศาสตร์, กรุงเทพฯ.

ปภากร อินทร์โท่ล่. (2556). “การวิเคราะห์ถดถอยโลจิสติกเพื่อการพยากรณ์การพบเชื้อซาลโมเนลล่าเอนเทอริกาในสุกร.” วารสารวิชาการ **Veridian E-Journal** ปีที่ 6, ฉบับที่ 3 (กันยายน-ธันวาคม): 949 – 959.

ไพรัช ทาบสีแพร. (2527). การแก้ไขปรับปรุงวิธีจับซ้ำ. วารสารวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น. 12(4): 241-254.

ยุทธ ไกรวรรณ์. (2555). หลักการและการใช้การวิเคราะห์การถดถอยโลจิสติคสำหรับการวิจัย. วารสารวิจัย มหาวิทยาลัยเทคโนโลยีราชมงคลศรีวิชัย. 4(1): 1-12.

สารภี รังสีโกศัย, นัยนา พิพัฒน์วณิชชา และ รวีวรรณ เผ่ากัณหา. (2556). ปัจจัยที่มีความสัมพันธ์กับพฤติกรรมการเตรียมตัวเพื่อเผชิญกับภาวะใกล้ตายและความตายของผู้สูงอายุในชมรมผู้สูงอายุ จังหวัดปัตตานี. วารสารวิทยาลัยพยาบาลสงขลานครินทร์. 33(1): 43-56.

สำนักงานนโยบายและแผนทรัพยากรธรรมชาติและสิ่งแวดล้อม. (2546). ความหลากหลายทางชีวภาพในพื้นที่ชุ่มน้ำบึงบอระเพ็ด. กระทรวงทรัพยากรธรรมชาติและสิ่งแวดล้อม. กรุงเทพฯ. 146 หน้า.

สมบัติ ท้ายเรือคำ. (2552). เอกสารประกอบคำสอนวิชาสถิติขั้นสูงสำหรับการวิจัยทางการศึกษา. ประสานการพิมพ์, กาฬสินธุ์.

Veridian E-Journal, Science and Technology Silpakorn University
Volume 5 Number 3 May – June 2018 ISSN 2408 - 1248

ภาษาต่างประเทศ

Böhning, D. (2008). A simple variance formula for population size estimators by conditioning. Statistical Methodology. 5: 410-423.

Burnham, K. P. and Overton, W. S. (1978). Estimation of the size of a closed population when capture probabilities vary among animals. Biometrika. 65: 625-633.

Burns, N. and Grove, S.K. (1993). The Practice of Nursing Research: Conduct, Critique and Utilization. W.B.Saunders Company, Philadelphia.

Chao, A. (1989). Estimating population size for sparse data in capture-recapture experiments. Biometrics. 45: 427-438.

Chapman, D. G. (1951). Some properties of the hypergeometric distribution with applications to zoological censuses. University of California publications in Statistics. 1: 131-160.

Darroch, J. N. (1958). The multiple recapture census. I. Estimation of a closed population. Biometrika. 45: 343-359.

Hair, Joseph F. and others. (2010). Multivariate data analysis : 7th ed. New Jersey: Pearson Prentice Hall.

Lincoln, F. C. (1930). Calculating waterfowl abundance on the basis of banding returns. U.S. Department of Agricultural Circular. 118: 1-4

Peterson, G. G. J. (1896). The yearly immigration of young place into the limford from the german sea. Report of Danish Biology Statistics. 6: 1-48.

Zelterman, D. (1988). Robust estimator in truncated discrete distributions with application to capture-recapture experiments. Journal of Statistical Planning and Inference. 18: 225-237.