

K45304201: สาขาวิชาสถิติประยุกต์

คำสำคัญ: การวิเคราะห์การจำแนก / การวิเคราะห์การถดถอยโลจิสติก / โรคเลปโตสไปโรสิส

พินดา คงแจ่ม: การศึกษาเปรียบเทียบตัวแบบในการวินิจฉัยโรคเลปโตสไปโรสิส (A COMPARATIVE STUDY OF LEPTOSPIROSIS DIAGNOSIS MODELS) อาจารย์ผู้ควบคุมวิทยานิพนธ์ : ศศ.ดร.กมลชนก พานิชการ และ ศ.พญ.ยุพิน ศุภธรรมมงคล.110หน้า. ISBN 974-11-6146-8

วัตถุประสงค์ของการวิจัยเพื่อเปรียบเทียบประสิทธิภาพของตัวแบบที่ได้จากการวิเคราะห์การจำแนกและการวิเคราะห์การถดถอยโลจิสติก โดยพิจารณาการจำแนก 3 กรณี คือ (1) จำแนกเป็นคนที่ได้รับการยืนยันการวินิจฉัยเป็นโรคเลปโตสไปโรสิส กับกลุ่มที่วินิจฉัยเป็นโรคอื่นๆ (2) จำแนกเป็นคนที่ได้รับการยืนยันการวินิจฉัยเป็นโรคเลปโตสไปโรสิสกับกลุ่มที่ได้รับการยืนยันการวินิจฉัยเป็นโรคในกลุ่ม Rickettsiosis (3) จำแนกเป็นคนที่ได้รับการยืนยันการวินิจฉัยเป็นโรคเลปโตสไปโรสิส กลุ่มที่ได้รับการยืนยันการวินิจฉัยเป็นโรคในกลุ่ม Rickettsiosis และกลุ่มที่วินิจฉัยเป็นโรคอื่นๆ

เกณฑ์ของ WHO ที่แพทย์นำมาใช้ช่วยในการวินิจฉัยโรคเลปโตสไปโรสิสอาจมีประสิทธิภาพไม่เพียงพอจึงศึกษาการนำเทคนิคทางสถิติ 2 เทคนิคคือ การวิเคราะห์การจำแนกและการวิเคราะห์การถดถอยโลจิสติกมาสร้างตัวแบบเพื่อช่วยในการวินิจฉัยโรคเพื่อความเร็วและแม่นยำในการวินิจฉัย โดยใช้ข้อมูลคนไข้ในโครงการวิจัยจำนวน 559, 149 และ 559 ค่าสังเกตสำหรับกรณี 1, 2 และ 3 ตามลำดับ ตัวแปรอิสระ 27 ตัวจากข้อมูล 4 ส่วนคือ ข้อมูลส่วนตัวของผู้ป่วย อาการทางคลินิกจากประวัติ การตรวจร่างกายแรกพบและผลตรวจทางห้องปฏิบัติการแรกพบ ขั้นตอนการศึกษาจะทำการสร้างตัวแบบจากแต่ละเทคนิคและจำแนกค่าสังเกตเข้ากลุ่ม จากนั้นคำนวณร้อยละการจำแนกเข้ากลุ่มถูกต้องรวมเพื่อเปรียบเทียบประสิทธิภาพของตัวแบบที่ได้จากเทคนิคทั้งสอง

ผลการวิจัยเมื่อพิจารณาเปอร์เซ็นต์การจัดข้อมูลเข้ากลุ่มถูกต้อง เฉพาะกรณีที่คนไข้ได้รับการยืนยันการวินิจฉัยว่าเป็นโรคเลปโตสไปโรสิสสามารถสรุปได้ดังต่อไปนี้ (1) กรณีที่ 1 การวิเคราะห์การถดถอยโลจิสติกมีประสิทธิภาพดีกว่าคิดเป็น 79.9% (2) กรณีที่ 2 การวิเคราะห์การถดถอยโลจิสติกมีประสิทธิภาพดีกว่าคิดเป็น 86.5% (3) กรณีที่ 3 การวิเคราะห์การจำแนกมีประสิทธิภาพดีกว่าคิดเป็น 70.1% ผลการวิจัยเมื่อพิจารณาเปอร์เซ็นต์การจัดข้อมูลเข้ากลุ่มถูกต้องรวม สามารถสรุปได้ดังต่อไปนี้ (1) กรณีที่ 1 การวิเคราะห์การจำแนกมีประสิทธิภาพดีกว่าคิดเป็น 74.6% (2) กรณีที่ 2 ทั้งสองเทคนิคมีประสิทธิภาพเท่าเทียมกันคิดเป็น 79.0% (3) สำหรับกรณีที่ 3 การวิเคราะห์การถดถอยโลจิสติกมีประสิทธิภาพดีกว่าคิดเป็น 62.1% เมื่อเปรียบเทียบผลการจัดข้อมูลเข้ากลุ่มของกรณีที่ 1 พบว่าตัวแบบที่สร้างขึ้นมีประสิทธิภาพสูงกว่าการใช้เกณฑ์ขององค์การอนามัยโลกวัดโดยใช้เปอร์เซ็นต์การจัดข้อมูลเข้ากลุ่มถูกต้องรวม

พิจารณาตัวแปรที่ถูกนำเข้าสมการ พบว่าตัวแปรที่สามารถจำแนกกลุ่มได้อย่างมีประสิทธิภาพสำหรับทั้ง 3 กรณีมี 5 ตัวคือ เยื่อตาอักเสบ (X_{15}), กดเจ็บกล้ามเนื้อ (X_{17}), PMN>75% (X_{22}), Total Bilirubin>3 mg/dl (X_{25}) และ Abnormal SGOT/SGPT/Alkaline phosphatase (X_{26})

ตัวแบบที่ถูกสร้างขึ้นถูกนำมาตรวจสอบโดยนำไปใช้กับข้อมูลชุดใหม่ ซึ่งเป็นคนไข้ที่ทำการวิจัยและเก็บข้อมูลในโรงพยาบาลและช่วงเวลาเดียวกัน พบว่าเมื่อพิจารณาเปอร์เซ็นต์การจัดข้อมูลเข้ากลุ่มถูกต้องเฉพาะกรณีที่คนไข้ยืนยันการวินิจฉัยเป็นโรคเลปโตสไปโรสิสตัวแบบสำหรับการจำแนกกรณีที่ 1 และกรณีที่ 2 มีประสิทธิภาพค่อนข้างดี ส่วนตัวแบบสำหรับการจำแนกในกรณีที่ 3 ต้องทำการศึกษาเพิ่มเติมจึงจะนำมาใช้ได้

K45304201: MAJOR : APPLIED STATISTICS

KEYWORD : DISCRIMINANT ANALYSIS / LOGISTIC REGRESSION ANALYSIS / LEPTOSPIROSIS

PANIDA KONGJAM : A COMPARATIVE STUDY OF LEPTOSPIROSIS DIAGNOSIS MODELS. THESIS ADVISOR : ASST.PROF. KAMOLCHANOK PANISHKAN, Ph.D., AND PROF. YUPIN SUPUTTAMONGKOL. 110 pp. ISBN 974-11-6146-8

The objective of this research is to compare the efficiency of discriminant analysis models and logistic regression models. The study was divided into 3 parts : (I) classifying patients into confirmed leptospirosis group and another group, (II) classifying patients into confirmed leptospirosis and rickettsiosis groups, (III) classifying patients into confirmed leptospirosis, rickettsiosis groups and another group.

Physicians typically use the first two parts of WHO criteria for leptospirosis diagnosis. This may not appropriate for tropical areas such as Thailand. This study proposes the optional criteria using two statistical methods : discriminant analysis and logistic regression analysis. A total data of 559,149,559 patients were used for analyzing in part (I) , (II) and (III), respectively. The 27 independent variables were studied. They involved with 4 characteristics of patients: personal data, clinical presentations, physical examinations and laboratory results. The effectiveness of the proposed models was assessed by using percentage of the correct classification into leptospirosis group. The results of the study are as follows :

- (1) for part (I), the logistic regression model presented the highest correct classification percentage (79.9%)
- (2) for part (II), the logistic regression model presented the highest correct classification percentage (86.5%)
- (3) for part (III), the discriminant analysis model showed the highest correct classification percentage (70.1%).

By focusing on the percentage of correct classification, discriminant analysis model was the best in part (I) (74.6%). Both models performed similarly (79.0%) in part (II). Logistic regression model performed well in part (III) (62.1%). Overall, the important independent variables selected in models are Conjunctival injection (X_{15}), Calf tenderness (X_{17}), PMN>75% (X_{22}), Total Bilirubin>3 mg/dl (X_{25}) and Abnormal SGOT/SGPT/Alkaline phosphatase (X_{26}).

The proposed models were used for classifying new data sets. Results showed that the proposed models performed satisfactorily in part (I) and (II). However, in part (III) the models may need to be improved.