รุ้งรวี อำนาจตระกูล 2555: การตรวจสอบค่าผิดปกติสำหรับข้อมูลหลายตัวแปรโดยใช้องค์ประกอบรอง ปริญญาปรัชญาดุษฎีบัณฑิต (สถิติ) สาขาสถิติ ภาควิชาสถิติ อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก: อาจารย์อำไพ ทองธีรภาพ, Ph.D. 87 หน้า

การศึกษาครั้งนี้มีวัตถุประสงค์เพื่อสร้างตัวสถิติที่ใช้ตรวจสอบค่าผิดปกติสำหรับข้อมูลหลายตัวแปรโดยใช้องค์ประกองรอง กรณีที่ค่าผิดปกติเกิดจากความสัมพันธ์ของตัวแปรไม่สอดคล้องกัน โดยศึกษาหลักการ และทฤษฎีที่เกี่ยวข้องกับวิธีการวิเคราะห์องค์ประกอบหลัก รวมทั้งสมบัติการแจกแจงแบบปกติ ทำให้ได้ตัวสถิติ R_{2z}^2 และ R_{3z}^2 ซึ่งใช้เพียงองค์ประกอบรอง 2 และ 3 องค์ประกอบ ตามลำดับ พบว่า ตัวสถิติทั้งสองมีการแจกแจงแบบไคกำลังสอง ที่มีองศาอิสระเท่ากับ 1 เมื่อทำการเปรียบเทียบประสิทธิภาพของตัวสถิติ R_{2z}^2 และ R_{3z}^2 กับตัวสถิติ $d_{2i,(2)}^2$ และ $d_{2i,(3)}^2$ ที่นำเสนอโดย Hawkins และตัวสถิติ d^2 จากวิธี Mahalanobis distance ในการตรวจสอบค่าผิดปกติสำหรับข้อมูลหลายตัวแปร จะทดลองกับข้อมูลจำลองภายใต้สถานการณ์ข้อมูลมีการแจกแจงแบบปกติหลายตัวแปร กำหนดจำนวนตัวแปร $p = 5, 10(2)20$ ขนาดข้อมูลตัวอย่าง $n = 30, 40, 50, 75, 100$ และร้อยละของค่าผิดปกติ 4 ระดับ ได้แก่ 10, 20, 30 และ 40 รวมทั้งหมด 140 สถานการณ์ ในแต่ละสถานการณ์ทำซ้ำจำนวน 1,000 ครั้ง

จากผลการศึกษา พบว่า ตัวสถิติ R_{2z}^2 จะเหมาะสำหรับตรวจสอบค่าผิดปกติเมื่อมีจำนวนตัวแปรตั้งแต่ 16 ตัวแปรขึ้นไป และขนาดข้อมูลตัวอย่างมากกว่า 50 ชุดข้อมูล แต่ตัวสถิติ R_{3z}^2 จะมีประสิทธิภาพต่ำกว่าทุกตัวสถิติในหลายกรณี จึงยังไม่เหมาะที่จะนำไปใช้ในการตรวจสอบค่าผิดปกติ ส่วนตัวสถิติ $d_{2i,(2)}^2$ จะเหมาะสำหรับตรวจสอบค่าผิดปกติเมื่อมีจำนวนตัวแปรไม่เกิน 10 ตัวแปร แต่เมื่อมีจำนวนตัวแปรอยู่ระหว่าง 12 ถึง 16 ตัวแปร ตัวสถิติ $d_{2i,(3)}^2$ จะเหมาะสำหรับตรวจสอบค่าผิดปกติมากกว่า สำหรับตัวสถิติ d^2 จะเหมาะสำหรับตรวจสอบค่าผิดปกติเมื่อมีจำนวนตัวแปรตั้งแต่ 14 ตัวแปรขึ้นไป อย่างไรก็ตามทุกตัวสถิติจะมีประสิทธิภาพใกล้เคียงกัน โดยตัวสถิติ R_{2z}^2 และ R_{3z}^2 จะมีรูปแบบการคำนวณที่ไม่ยุ่งยาก สามารถคำนวณได้ง่าย และรวดเร็ว

____________________ ____________________ ___ / ___ / ___

ลายมือชื่อนิสิต ลายมือชื่ออาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์