

รักรณิ นุตรชน 2551: การศึกษาประสิทธิภาพของวิธีการทดสอบการแจกแจงปกติพหุ 6 วิธี  
 ปริญญาวิทยาศาสตรมหาบัณฑิต (สถิติ) สาขาสถิติ ภาควิชาสถิติ ภาควิชาการที่ปรึกษา:  
 ผู้ช่วยศาสตราจารย์บุญอ้อม โจนที, Ph.D. 241 หน้า

การศึกษานี้มีวัตถุประสงค์เพื่อศึกษาประสิทธิภาพของวิธีการทดสอบการแจกแจงปกติพหุ 6 วิธี ได้แก่ สถิติ S สถิติ K สถิติ T สถิติ  $W_F$  สถิติ HZ และสถิติ O โดยทำการศึกษาทั้งกรณีทราบค่าและไม่ทราบค่าพารามิเตอร์ โดยการจำลองข้อมูลแต่ละสถานการณ์ด้วยเทคนิคมอนติคาร์โล จำนวน 5,000 ซ้ำ เมื่อกำหนดให้ข้อมูลมีการแจกแจง 3 แบบ คือ การแจกแจงปกติพหุ การแจกแจงลิอองนอร์มอลพหุ และการแจกแจงสตีเวน-ทีพหุ เมื่อตัวอย่างมีขนาดเล็ก ( $n=20$ ) ขนาดกลาง ( $n=30$ ) และขนาดใหญ่ ( $n=40, 50$ ) โดยที่จำนวนตัวแปร ( $p$ ) เท่ากับ 2, 3, 4 และศึกษาที่ระดับนัยสำคัญของการทดสอบ ( $\alpha$ ) 2 ระดับ คือ 0.05 และ 0.10 ผลการศึกษาแบ่งเป็น 2 ส่วนดังนี้

1. ความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ I พบว่า สถิติ S, K และ T สามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้ดี กรณีทราบค่าพารามิเตอร์ ส่วนสถิติ  $W_F$  สามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้ดี เมื่อตัวอย่างมีขนาดเล็ก ทั้งกรณีทราบค่าและไม่ทราบค่าพารามิเตอร์ สถิติ O สามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้ดี เมื่อตัวอย่างมีขนาดเท่ากับ 30, 50 ทั้งกรณีทราบค่าและไม่ทราบค่าพารามิเตอร์ ขณะที่สถิติ HZ สามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้ดีเกือบทุกกรณี ทั้งกรณีทราบค่าและไม่ทราบค่าพารามิเตอร์

2. การเปรียบเทียบอำนาจการทดสอบแบ่งเป็น 2 กรณี คือ 1) กรณีทราบค่าพารามิเตอร์ เมื่อประชากรมีการแจกแจงลิอองนอร์มอลพหุ สถิติ  $W_F$  ให้อำนาจการทดสอบสูงสุด เมื่อตัวอย่างมีขนาดเล็ก ส่วนสถิติ HZ ให้อำนาจการทดสอบสูงสุด เมื่อตัวอย่างมีขนาดกลางและใหญ่ และเมื่อประชากรมีการแจกแจงสตีเวน-ทีพหุ สถิติ K ให้อำนาจการทดสอบสูงสุดทุกกรณี 2) กรณีไม่ทราบค่าพารามิเตอร์ สำหรับประชากรมีการแจกแจงลิอองนอร์มอลพหุ เมื่อตัวอย่างขนาดเล็ก สถิติ  $W_F$  ให้อำนาจการทดสอบสูงสุด และตัวอย่างขนาดกลาง สถิติ  $W_F$  และ HZ ให้อำนาจการทดสอบใกล้เคียงกัน สำหรับตัวอย่างขนาดใหญ่ สถิติ HZ, S, T และ O ให้อำนาจการทดสอบใกล้เคียงกัน ส่วนกรณีประชากรมีการแจกแจงสตีเวน-ทีพหุ เมื่อตัวอย่างมีขนาดเล็ก สถิติ  $W_F$  ให้อำนาจการทดสอบสูงสุด และเมื่อตัวอย่างมีขนาดกลาง สถิติ  $W_F$  และ HZ ให้อำนาจการทดสอบสูงสุดใกล้เคียงกัน ในขณะที่ตัวอย่างมีขนาดใหญ่ สถิติ T ให้อำนาจการทดสอบสูงสุด

จากการศึกษานี้ พบว่า สถิติ HZ สามารถควบคุมความคลาดเคลื่อนประเภทที่ I ได้ และมีอำนาจการทดสอบสูงที่สุดในเกือบทุกสถานการณ์ของข้อมูลที่ศึกษา ดังนั้นสถิติ HZ เป็นวิธีการทดสอบการแจกแจงปกติพหุที่เหมาะสม

Rukmanee Butchon 2008: The Efficient Study of the Six Multivariate Normal Distribution Tests. Master of Science (Statistics), Major Field: Statistics, Department of Statistics. Thesis Advisor: Assistant Professor Boonorm Chomtee, Ph.D. 241 pages.

The objectives of this research are to compare controlling type I error and power of the six multinormality tests: S, K, T, WF, HZ and O in both known and unknown parameters for the three distributions: multivariate normal, multivariate lognormal and multivariate student-t distributions for fixed parameters; 2, 3, 4 variables ( $p$ ); 20, 30, 40, 50 sample sizes ( $n$ ); 0.05, 0.1 levels of significance ( $\alpha$ ). The data were simulated by Monte Carlo Technique and 5,000 replications for each situation. The results are as follows:

1. For Type I error: S, K, T statistics can be rightly control type I error when known parameters.  $W_F$  statistic can be properly control type I error both known and unknown parameters when  $n=20$ . O is a good statistic for controlling type I error both known and unknown parameters when  $n=30, 50$ . However, HZ statistic can be nicely control type I error for almost situations.

2. For Power of the test: it can be concluded to 2 parts. Part 1: For known parameter cases, when data are multivariate lognormal distribution,  $W_F$  statistic provides the highest power of the test when  $n=20$ . HZ statistic gives the maximum power of the test when  $n=30, 40, 50$ . When data are multivariate student-t distribution, K statistic is the best statistic for all cases. Part 2: For unknown parameter cases, when data are multivariate lognormal distribution,  $W_F$  statistic gives the highest power of the test when  $n=20$ . However, both  $W_F$  and HZ statistics provide the high power of the test when  $n=30$ . In addition, HZ, S, T and O statistics are the maximum power of the test at  $n=40, 50$ . In case of data are multivariate student-t distribution,  $W_F$  statistic is the best when  $n=20$ . Also, both  $W_F$  and HZ statistics are the best when  $n=30$ . However, T statistic is the best when  $n=40, 50$ .

In summary, HZ statistic is recommended as an optimal test for multivariate normal distribution testing because it can be nicely control type I error and provides the highest power of the test in almost situations.