ใบรับรองวิทยานิพนธ์

บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์

ปรัชญาดุษฎีบัณฑิต (สถิติ)

ปริญญา

สถิติ — **สาขา**

สถิติ — **ภาควิชา**

เรื่อง สถิติทดสอบเพื่อเปรียบเทียบฟังก์ชันการถดถอยแบบไม่อิงพารามิเตอร์

Test Statistics for Comparison of Nonparametric Regression Functions

นามผู้วิจัย นางสาวอัญชลี ทองกำเหนิด

ได้พิจารณาเห็นชอบโดย

ประธานกรรมการ .....

(อาจารย์อำไพ ทองธีรภาพ, Ph.D.)

กรรมการ .....

(ผู้ช่วยศาสตราจารย์บุญอ้อม โฉมที, Ph.D.)

กรรมการ .....

(รองศาสตราจารย์อุษณีย์ ลีรวัฒน์, วท.ด.)

หัวหน้าภาควิชา .....

(อาจารย์อำไพ ทองธีรภาพ, Ph.D.)

บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์รับรองแล้ว

.....

(รองศาสตราจารย์กัญจนา ธีระกุล, D.Agr.)

คณบดีบัณฑิตวิทยาลัย

วันที่ ..... **เดือน** ..... **พ.ศ.** .....

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

วิทยานิพนธ์

เรื่อง

สถิติทดสอบเพื่อเปรียบเทียบฟังก์ชันการถดถอยแบบไม่อิงพารามิเตอร์

Test Statistics for Comparison of Nonparametric Regression Functions

โดย

นางสาวอัญชลี ทองกำเหนิด

เสนอ

บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์
เพื่อความสมบูรณ์แห่งปริญญาปรัชญาดุษฎีบัณฑิต (สถิติ)
พ.ศ. 2554

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

อัญชลี ทองกำเหนิด 2554: สถิติทดสอบเพื่อเปรียบเทียบฟังก์ชันการถดถอยแบบไม่อิงพารามิเตอร์ ปริญญาปรัชญาดุษฎีบัณฑิต (สถิติ) สาขาสถิติ ภาควิชาสถิติ อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก: อาจารย์อำไพ ทองธีรภาพ, Ph.D. 158 หน้า

การศึกษาในครั้งนี้ มีวัตถุประสงค์เพื่อสร้างสถิติทดสอบการเท่ากันของฟังก์ชันการถดถอย k ฟังก์ชัน และศึกษาเกี่ยวกับการแจกแจงของสถิติทดสอบภายใต้สมมติฐานหลัก (H_0) และสมมติฐานทางเลือก (H_1) โดยสมมติฐานของการทดสอบอยู่ในรูปของ $H_0 : m_1 = m_2 = m_3 = ... = m_k$ และ $H_1: m_i \neq m_j$ สำหรับ i, j บางค่า $i \neq j,\ i, j \in \{1, ... , k\},\ k \geq 2$ เมื่อ m_j คือฟังก์ชันการถดถอยแบบไม่อิงพารามิเตอร์ที่มีตัวแบบอยู่ในรูปของ $Y_{ij} = m_j(X_{ij}) + \sigma_j(X_{ij})\varepsilon_{ij},\ i = 1, ... , n_j, j = 1, ... , k,$ $m_j(X_{ij}) = E(Y_{ij}|X_{ij})$ คือฟังก์ชันการถดถอยแบบไม่อิงพารามิเตอร์ที่ประมาณจากวิธีการของ Nadaraya – Watson Estimator (NW), $\sigma_j^2(X_{ij}) = Var(Y_{ij}|X_{ij})$ คือความแปรปรวนและ ε_{ij} คือความคลาดเคลื่อน การสร้างสถิติทดสอบมีแนวคิดพื้นฐานมาจากการเปรียบเทียบความแตกต่างระหว่างตัวประมาณฟังก์ชันการแจกแจงของความคลาดเคลื่อนในแต่ละประชากร ในการประมาณค่าวิกฤตของสถิติทดสอบใช้หลักการของบูตสแตรป นอกจากนั้นในการศึกษาครั้งนี้ได้ศึกษาอำนาจการทดสอบ และความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 ของสถิติทดสอบเมื่อตัวอย่างมีขนาดเล็ก (n=20) ขนาดกลาง (n=50) ขนาดใหญ่ (n = 100) และการกำหนดการ Shift ของฟังก์ชันสำหรับตัวแปรอิสระ X ที่ทำให้สมมติฐานทางเลือกเป็นจริง โดยพิจารณาใน 3 ลักษณะคือ Constant Shift, Affine Shift และ Quadratic Shift

ผลการศึกษาพบว่าสถิติทดสอบ $U_{KS}^1, U_{KU}^1, U_{KS}^2, U_{KU}^2$ ทั้ง 4 รูปแบบเป็นฟังก์ชันของ $U_j(y) = f_{\varepsilon j}(y)p_j\sum_{r=1}^{k}\sum_{i=1}^{n_r}\frac{Y_{ir} - m(X_{ir})}{\sigma_j(X_{ir})}\left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j)}{p_j}\right),\ j = 1, ... , k,$ $-\infty < y < \infty$ โดยที่ $U_j(y)$ มีการแจกแจงแบบปกติ นอกจากนั้นสถิติทดสอบทั้ง 4 รูปแบบมีอำนาจการทดสอบสูงขึ้นเมื่อขนาดของตัวอย่างเพิ่มขึ้น ในขณะเดียวกันการควบคุมความคลาดเคลื่อนประเภทที่ 1 จะทำได้ดีเมื่อตัวอย่างมีขนาดใหญ่

_____ _____ ___ / ___ / ___

ลายมือชื่อนิสิต ลายมือชื่ออาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

Unchalee Tonggumnead 2011: Test Statistics for Comparison of Nonparametric Regression Functions. Doctor of Philosophy (Statistics), Major Field: Statistics, Department of Statistics. Thesis Advisor: Mrs. Umpai Thongteeraparp, Ph.D. 158 pages.

The purpose of this study is to construct the test statistics for testing the equality of k regression function, and to study about the distribution of the test statistics under the null hypothesis (H_0) and alternative hypothesis (H_1). The hypothesis is as follows : $H_0: m_1 = m_2 = m_3 = ... = m_k$ vs $H_1: m_i \neq m_j$, for some $i \neq j$, $i,j \in \{1,...,k\}$, $k \geq 2$. m_j is non – parametric regression function, the model is as follows : $Y_{ij} = m_j(X_{ij}) + \sigma_j(X_{ij})\varepsilon_{ij}$, $i = 1, ... ,n_j$, $j = 1, ... ,k$, $m_j(X_{ij}) = E(Y_{ij}|X_{ij})$ is non – parametric regression function that estimate from Nadaraya – Watson Estimator (NW), $\sigma_j^2(X_{ij}) = Var(Y_{ij}|X_{ij})$ is the conditional variance, and ε_{ij} is the error variable. The critical values of the test statistics can be approximated by bootstrap procedure. In addition, the power of the test and controlling the probability of type I error present in the case of small sample sizes (n =20), moderate sample sizes (n = 50), large sample sizes (n=100) and specify the shift function of independent variable X under the alternative hypothesis in three types: Constant Shift, Affine Shift and Quadratic Shift.

The result findings were as follows: The test statistics $U_{KS}^1, U_{KU}^1, U_{KS}^2, U_{KU}^2$ are the function of $U_j(y) = f_{\varepsilon j}(y)p_j \sum_{r=1}^{k}\sum_{i=1}^{n_r} \frac{Y_{ir} - m(X_{ir})}{\sigma_j(X_{ir})}\left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j)}{p_j}\right)$, $j = 1, ... ,k$, $-\infty < y < \infty$, and $U_j(y)$ have normal distribution. In addition, power of the test base on $U_{KS}^1, U_{KU}^1, U_{KS}^2, U_{KU}^2$ performed best for the large sample size, and mostly these four test statistics could control the probability of type I error in all situations for the large sample size.

_____ / ___ / ___

Student's signature Thesis Advisor's signature

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

กิตติกรรมประกาศ

ผู้วิจัยขอกราบขอบพระคุณ อาจารย์ ดร. อำไพ ทองธีรภาพ ประธานกรรมการที่ปรึกษาวิทยานิพนธ์ ที่ได้ช่วยเหลือในการวางแผนงานวิจัย ตลอดจนให้คำปรึกษาแนะนำและตรวจแก้ไขข้อบกพร่องต่าง ๆ ของวิทยานิพนธ์ฉบับนี้ ขอกราบขอบพระคุณ รองศาสตราจารย์ ดร. อุษณีย์ ลีรวัฒน์ ผู้ช่วยศาสตราจารย์ ดร. บุญอ้อม โฉมที กรรมการที่ปรึกษาวิทยานิพนธ์ร่วม รองศาสตราจารย์ ประสิทธิ์ พยัคฆพงษ์ ประธานการสอบปากเปล่าขั้นสุดท้าย และผู้ช่วยศาสตราจารย์ ดร. กุศยา ปลั่งพงษ์พันธ์ ที่กรุณาให้คำปรึกษาและตรวจแก้ไขวิทยานิพนธ์จนกระทั่งเสร็จสมบูรณ์

ขอกราบขอบพระคุณอาจารย์ภาควิชาสถิติทุกท่าน ที่ได้อบรมสั่งสอนและมอบความรู้อันเป็นประโยชน์อย่างยิ่งในการนำไปใช้ประโยชน์ ขอบคุณ อาจารย์นิกร แสงงาม ที่ให้คำแนะนำและให้กำลังใจกับผู้วิจัยมาด้วยดีเสมอมา

ด้วยความดีหรือประโยชน์อันใดเนื่องจากวิทยานิพนธ์เล่มนี้ ขอบมอบแด่คุณพ่อ คุณแม่ ผู้ให้ชีวิต ให้กำลังใจ และให้โอกาสที่ดีกับผู้วิจัยตลอดมา

อัญชลี ทองกำเหนิด

กันยายน 2554

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

สารบัญ

หน้า

สารบัญ	(1)
สารบัญตาราง	(2)
สารบัญภาพ	(6)
คำอธิบายสัญลักษณ์และคำย่อ	(11)
คำนำ	1
วัตถุประสงค์	7
การตรวจเอกสาร	9
อุปกรณ์และวิธีการ	31
อุปกรณ์	31
วิธีการ	31
ผลและวิจารณ์	85
ผล	85
วิจารณ์	127
สรุปและข้อเสนอแนะ	133
สรุป	133
ข้อเสนอแนะ	135
เอกสารและสิ่งอ้างอิง	136
ภาคผนวก	139
ภาคผนวก ก ความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1	140
ภาคผนวก ข การเขียนโปรแกรม	143
ประวัติการศึกษาและการทำงาน	158

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

สารบัญตาราง

ตารางที่		หน้า
1	เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานหลักของตัวแบบที่ 1) - 4) สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \sigma_2(x) = 0.50$ เมื่อการแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim N(0,1), \varepsilon_2 \sim N(0,1)$	86
2	เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Constant Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \sigma_2(x) = 0.50$ เมื่อการแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim N(0,1), \varepsilon_2 \sim N(0,1)$	89
3	เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Affine Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \sigma_2(x) = 0.50$ เมื่อการแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim N(0,1), \varepsilon_2 \sim N(0,1)$	91
4	เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Quadratic Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \sigma_2(x) = 0.50$ เมื่อการแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim N(0,1), \varepsilon_2 \sim N(0,1)$	93
5	เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานหลักของตัวแบบที่ 1) - 4) สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \sigma_2(x) = 0.50$ เมื่อการแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim$ Exponential (1) - 1, $\varepsilon_2 \sim$ Exponential (1) - 1	96

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

สารบัญตาราง (ต่อ)

ตารางที่		หน้า
6	เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Constant Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \sigma_2(x) = 0.50$ เมื่อการแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim$ Exponential (1) - 1, $\varepsilon_2 \sim$ Exponential (1) - 1	99
7	เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Affine Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \sigma_2(x) = 0.50$ เมื่อการแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim$ Exponential (1) - 1, $\varepsilon_2 \sim$ Exponential (1) - 1	101
8	เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Quadratic Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \sigma_2(x) = 0.50$ เมื่อการแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim$ Exponential (1) - 1, $\varepsilon_2 \sim$ Exponential (1) - 1	103
9	เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานหลักของตัวแบบที่ 1) - 4) สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \sigma_2(x) = 0.50, \sigma_3(x) = 0.75$ เมื่อการแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim N(0,1), \varepsilon_2 \sim N(0,1), \varepsilon_3 \sim N(0,1)$	106
10	เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Constant Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \sigma_2(x) = 0.50, \sigma_3(x) = 0.75$ เมื่อการแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim N(0,1), \varepsilon_2 \sim N(0,1), \varepsilon_3 \sim N(0,1)$	109

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

สารบัญตาราง (ต่อ)

ตารางที่		หน้า
11	เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Affine Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25,\ \sigma_2(x) = 0.50, \sigma_3(x) = 0.75$ เมื่อการแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim N(0,1),\ \varepsilon_2 \sim N(0,1),\ \varepsilon_3 \sim N(0,1)$	111
12	เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Quadratic Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25,\ \sigma_2(x) = 0.50, \sigma_3(x) = 0.75$ เมื่อการแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim N(0,1),\ \varepsilon_2 \sim N(0,1),\ \varepsilon_3 \sim N(0,1)$	113
13	เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานหลักของตัวแบบที่ 1) - 4) สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25,\ \sigma_2(x) = 0.50, \sigma_3(x) = 0.75$ เมื่อการแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim$ Exponential (1) - 1, $\varepsilon_2 \sim$ Exponential (1) - 1, $\varepsilon_3 \sim$ Exponential (1) - 1	116
14	เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณีConstant Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \sigma_2(x) = 0.50, \sigma_3(x) = 0.75$ เมื่อการแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim$ Exponential (1) - 1, $\varepsilon_2 \sim$ Exponential (1) - 1, $\varepsilon_3 \sim$ Exponential (1) - 1	119
15	เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Affine Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25,\ \sigma_2(x) = 0.50, \sigma_3(x) = 0.75$ เมื่อการแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim$ Exponential (1) - 1, $\varepsilon_2 \sim$ Exponential (1) - 1, $\varepsilon_3 \sim$ Exponential (1) - 1	121

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

สารบัญตาราง (ต่อ)

ตารางที่		หน้า
16	เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Quadratic Shift สำหรับสถิติทดสอบ U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU} ในกรณีที่ $\sigma_1(x) = 0.25$, $\sigma_2(x) = 0.50$, $\sigma_3(x) = 0.75$ เมื่อการแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim$ Exponential (1) - 1, $\varepsilon_2 \sim$ Exponential (1) - 1, $\varepsilon_3 \sim$ Exponential (1) - 1	123

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

สารบัญภาพ

ภาพที่		หน้า
1	ความแตกต่างของการถดถอยแบบอิงพารามิเตอร์และการถดถอยแบบไม่อิงพารามิเตอร์	2
2	ค่าเฉลี่ยของความคลาดเคลื่อนเมื่อ $E(\varepsilon_i) = 0$	5
3	ค่าเฉลี่ยของความคลาดเคลื่อนภายใต้สมมติฐานหลักเมื่อ $E(\varepsilon_{i0}) > 0$	5
4	แผนภาพการกระจายระหว่าง log (รายได้เฉลี่ยต่อครัวเรือน) และ log (ค่าใช้จ่ายเฉลี่ยต่อครัวเรือน) ปี พ.ศ. 2547 (วงกลม) ปี พ.ศ. 2549 (สี่เหลี่ยม) และปี พ.ศ. 2550 (สามเหลี่ยม) โดยประมาณเส้นโค้งการถดถอยระหว่าง log (รายได้เฉลี่ยต่อครัวเรือน) และ log (ค่าใช้จ่ายเฉลี่ยต่อครัวเรือน) ปี พ.ศ. 2547 ด้วยเส้นประ ปี พ.ศ. 2549 ด้วยเส้นประจุดประ และ ปี พ.ศ. 2550 ด้วยเส้นทึบ	84
5	เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานหลักของตัวแบบที่ 1) - 4) สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \sigma_2(x) = 0.50$ การแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim N(0,1), \varepsilon_2 \sim N(0,1)$ เมื่อระดับนัยสำคัญเท่ากับ 0.05 และ 0.10	87
6	เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Constant Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \sigma_2(x) = 0.50$ การแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim N(0,1), \varepsilon_2 \sim N(0,1)$ เมื่อระดับนัยสำคัญเท่ากับ 0.05 และ 0.10	90
7	เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Affine Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \sigma_2(x) = 0.50$ การแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim N(0,1), \varepsilon_2 \sim N(0,1)$ เมื่อระดับนัยสำคัญเท่ากับ 0.05 และ 0.10	92

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

สารบัญภาพ (ต่อ)

ภาพที่		หน้า
8	เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Quadratic Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \sigma_2(x) = 0.50$ การแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim N(0,1), \varepsilon_2 \sim N(0,1)$ เมื่อระดับนัยสำคัญเท่ากับ 0.05 และ 0.10	94
9	เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานหลักของตัวแบบที่ 1) - 4) สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \sigma_2(x) = 0.50$ การแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim$ Exponential (1) - 1, $\varepsilon_2 \sim$ Exponential (1) - 1 เมื่อระดับนัยสำคัญเท่ากับ 0.05 และ 0.10	97
10	เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Constant Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \sigma_2(x) = 0.50$ การแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim$ Exponential (1) - 1, $\varepsilon_2 \sim$ Exponential (1) - 1 เมื่อระดับนัยสำคัญเท่ากับ 0.05 และ 0.10	100
11	เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Affine Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \sigma_2(x) = 0.50$ การแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim$ Exponential (1) - 1, $\varepsilon_2 \sim$ Exponential (1) - 1 เมื่อระดับนัยสำคัญเท่ากับ 0.05 และ 0.10	102
12	เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Quadratic Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \sigma_2(x) = 0.50$ เมื่อการแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim$ Exponential (1) - 1, $\varepsilon_2 \sim$ Exponential (1) - 1 เมื่อระดับนัยสำคัญเท่ากับ 0.05 และ 0.10	104

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

สารบัญภาพ (ต่อ)

ภาพที่		หน้า
13	เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานหลักของตัวแบบที่ 1) - 4) สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \sigma_2(x) = 0.50, \sigma_3(x) = 0.75$ เมื่อการแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim N(0,1), \varepsilon_2 \sim N(0,1), \varepsilon_3 \sim N(0,1)$ ที่ระดับนัยสำคัญเท่ากับ 0.05 และ 0.10	107
14	เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Constant Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \sigma_2(x) = 0.50, \sigma_3(x) = 0.75$ เมื่อการแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim N(0,1), \varepsilon_2 \sim N(0,1), \varepsilon_3 \sim N(0,1)$ ที่ระดับนัยสำคัญเท่ากับ 0.05 และ 0.10	110
15	เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Affine Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \sigma_2(x) = 0.50, \sigma_3(x) = 0.75$ เมื่อการแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim N(0,1), \varepsilon_2 \sim N(0,1), \varepsilon_3 \sim N(0,1)$ ที่ระดับนัยสำคัญเท่ากับ 0.05 และ 0.10	112
16	เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Quadratic Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \sigma_2(x) = 0.50, \sigma_3(x) = 0.75$ เมื่อการแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim N(0,1), \varepsilon_2 \sim N(0,1), \varepsilon_3 \sim N(0,1)$ ที่ระดับนัยสำคัญเท่ากับ 0.05 และ 0.10 เมื่อระดับนัยสำคัญเท่ากับ 0.05 และ 0.10	114

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

สารบัญภาพ (ต่อ)

ภาพที่		หน้า
17	เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานหลักของตัวแบบที่ 1) - 4) สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \sigma_2(x) = 0.50, \sigma_3(x) = 0.75$ เมื่อการแจกแจงของความคลาดเคลื่อน ε_1 ~ Exponential (1) - 1, ε_2 ~ Exponential (1) - 1, ε_3 ~ Exponential (1) - 1 ที่ระดับนัยสำคัญเท่ากับ 0.05 และ 0.10	117
18	เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Constant Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \sigma_2(x) = 0.50, \sigma_3(x) = 0.75$ เมื่อการแจกแจงของความคลาดเคลื่อน ε_1 ~ Exponential (1) - 1, ε_2 ~ Exponential (1) - 1, ε_3 ~ Exponential (1) - 1 ที่ระดับนัยสำคัญเท่ากับ 0.05 และ 0.10	120
19	เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Affine Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \sigma_2(x) = 0.50, \sigma_3(x) = 0.75$ เมื่อการแจกแจงของความคลาดเคลื่อน ε_1 ~ Exponential (1) - 1, ε_2 ~ Exponential (1) - 1, ε_3 ~ Exponential (1) - 1 ที่ระดับนัยสำคัญเท่ากับ 0.05 และ 0.10	122
20	เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Quadratic Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \sigma_2(x) = 0.50, \sigma_3(x) = 0.75$ เมื่อการแจกแจงของความคลาดเคลื่อน ε_1 ~ N(0,1), ε_2 ~ N(0,1), ε_3 ~ N(0,1) ที่ระดับนัยสำคัญเท่ากับ 0.05 และ 0.10 เมื่อระดับนัยสำคัญเท่ากับ 0.05 และ 0.10	124

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

สารบัญภาพ (ต่อ)

ภาพที่		หน้า
21	ค่า P - values ที่พารามิเตอร์ปรับให้เรียบ (h) เท่ากับ 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40 สำหรับสถิติทดสอบ U^1_{KS} (◆) และ U^1_{KU} (■)	126

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

คำอธิบายสัญลักษณ์และคำย่อ

$F_{\varepsilon j}(y)$ = ฟังก์ชันการแจกแจงสะสมของตัวอย่าง (Empirical distribution function) ของความคลาดเคลื่อนของประชากรที่ j

$F_{\varepsilon j0}(y)$ = ฟังก์ชันการแจกแจงสะสมของตัวอย่างของความคลาดเคลื่อนของประชากรที่ j เมื่อสมมติฐานหลักเป็นจริง

$F_{\varepsilon}(y)$ = ฟังก์ชันการแจกแจงสะสมของตัวอย่างของค่าเฉลี่ยของความคลาดเคลื่อน

$F_{\varepsilon 0}(y)$ = ฟังก์ชันการแจกแจงสะสมของตัวอย่างของค่าเฉลี่ยของความคลาดเคลื่อนเมื่อสมมติฐานหลักเป็นจริง

$m_j(X_{ij})$ = ฟังก์ชันการถดถอยของ Y บน X โดยที่ $m_j(X_{ij}) = E(Y_{ij}|X_{ij})$

$\sigma_j^2(X_{ij})$ = ความแปรปรวน โดยที่ $\sigma_j^2(X_{ij}) = Var(Y_{ij}|X_{ij})$

ε_{ij} = ความคลาดเคลื่อน โดยที่ $E(\varepsilon_{ij}|X_{ij}) = 0$ และ $Var(\varepsilon_{ij}) = 1$

$\bar{Y}_j$ = ค่าเฉลี่ยของตัวแปรตาม Y จากประชากรกลุ่มที่ j,

$$\bar{Y}_j = \frac{\sum_{i=1}^{n_j} Y_{ij}}{n_j}$$

$\hat{m}_j(X_{ij})$ = ตัวประมาณฟังก์ชันการถดถอย $m_j(X_{ij})$

$\hat{m}(X_{ij})$ = ตัวประมาณฟังก์ชันการถดถอย $m(X_{ij})$ เมื่อสมมติฐานหลักเป็นจริง

$\hat{\sigma}_j^2(X_{ij})$ = ตัวประมาณของความแปรปรวน

$W_{ij}(X_{ij},h)$ = Nadaraya – Watson Weight: $W_{ij}(X_{ij},h) = \dfrac{K\left(\dfrac{x - X_{ij}}{h}\right)}{\sum_{i=1}^{n_j} K\left(\dfrac{x - X_{ij}}{h}\right)}$

K = Kernel function

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

คำอธิบายสัญลักษณ์และคำย่อ (ต่อ)

h = พารามิเตอร์ปรับให้เรียบ

x = ตำแหน่งที่ต้องการประมาณฟังก์ชันการถดถอย

X_{ij} = ตัวแปรอิสระค่าที่ i กลุ่มที่ j, $i = 1, \ldots, n_j$, $j = 1, \ldots, k$

Y_{ij} = ตัวแปรตามค่าที่ i กลุ่มที่ j, $i = 1, \ldots, n_j$, $j = 1, \ldots, k$

$\hat{U}(y)$ = k – dimensional process $\hat{U}(y) = (\hat{U}_1(y), \ldots, \hat{U}_k(y))^t$, $-\infty < y < \infty$, เมื่อ $\hat{U}_j(y) = n_j^{1/2}(\hat{F}_{\varepsilon j0}(y) - \hat{F}_{\varepsilon j}(y))$, $j = 1, \ldots, k$

n_j = ขนาดตัวอย่างกลุ่มที่ j

n = ขนาดของตัวอย่างทั้งหมด $n = \sum_{j=1}^{k} n_j$, $i = 1, \ldots, n_j$, $j = 1, \ldots, k$

p_j = ค่าสัดส่วนของตัวอย่าง, $p_j = \frac{n_j}{n}$

$f_j(x)$ = ฟังก์ชันความหนาแน่นน่าจะเป็น (Probability density function) ของ X_j

$f_{mix}(x)$ = mixture ของฟังก์ชันความหนาแน่นน่าจะเป็นของ X_j โดย $f_{mix}(x) = \sum_{j=1}^{k} p_j f_j(x)$

$\hat{m}(x)$ = ตัวประมาณฟังก์ชันการถดถอย $m_n(x)$ เมื่อ $m_n(x) = m_o(x) + n^{-1/2} S(x)$

$m_o(x)$ = ฟังก์ชันการถดถอยภายใต้สมมติฐานหลัก

$S(x)$ = ฟังก์ชันของ x ที่เปลี่ยนไป (Shift), $S(x) = \sum_{j=1}^{k} p_j \frac{f_j(x)}{f_{mix}(x)} s_j(x)$

$\hat{m}_{jn}(x)$ = ตัวประมาณฟังก์ชันการถดถอย $m_{jn}(x)$ เมื่อ $m_{jn}(x) = m_o(x) + n^{-1/2} s_j(x)$

$s_j(x)$ = ฟังก์ชันของ x กลุ่มที่ j ที่เปลี่ยนไป

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

คำอธิบายสัญลักษณ์และคำย่อ (ต่อ)

$f_{\varepsilon j}(y)$ = ฟังก์ชันความหนาแน่นน่าจะเป็นของความคลาดเคลื่อนกลุ่มที่ j, โดย $f_{\varepsilon j}(y) = F'_{\varepsilon j}(y)$

$f_{\varepsilon j0}(y)$ = ฟังก์ชันความหนาแน่นน่าจะเป็นของความคลาดเคลื่อนภายใต้สมมติฐานหลัก โดย $f_{\varepsilon j0}(y) = F'_{\varepsilon j0}(y)$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

สถิติทดสอบเพื่อเปรียบเทียบฟังก์ชันการถดถอยแบบไม่อิงพารามิเตอร์

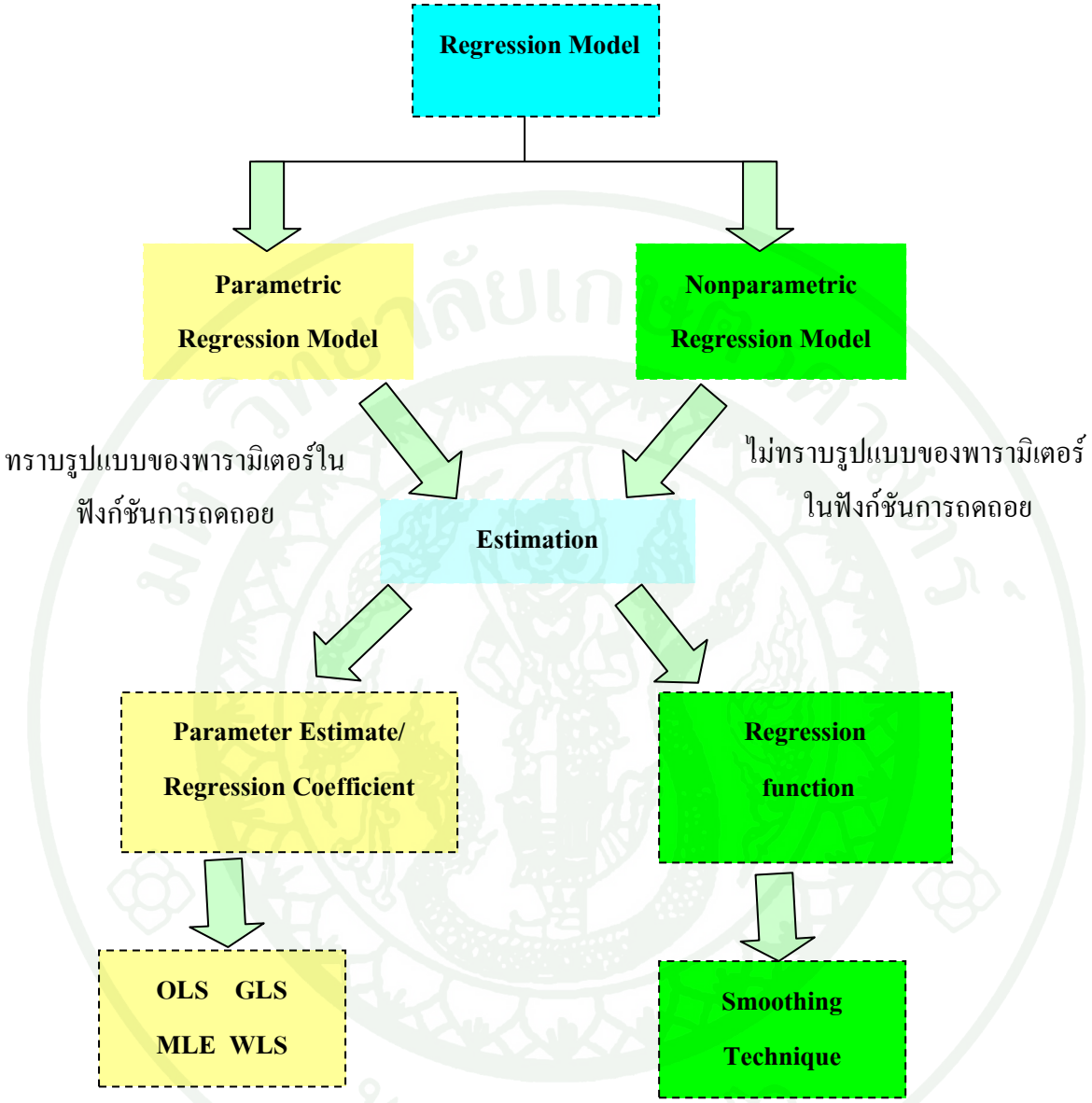
Test Statistics for Comparison of Nonparametric Regression Functions

คำนำ

การวิเคราะห์การถดถอย (Regression Analysis) เป็นการศึกษาความสัมพันธ์ระหว่างตัวแปร โดย เรียกตัวแปรที่สนใจศึกษาว่า ตัวแปรตาม (Dependent Variable หรือ Response Variable) และเรียกตัวแปรอื่นที่แทนปัจจัยที่ทำให้ค่าของตัวแปรตามเปลี่ยนแปลงว่า ตัวแปรอิสระ (Independent Variable หรือ Predictor Variable) การวิเคราะห์การถดถอย เมื่อแบ่งตามตัวแบบการถดถอยสามารถแบ่งได้เป็น 2 ประเภท ได้แก่ ตัวแบบการถดถอยเชิงเส้น (Linear Regression Model) และตัวแบบการถดถอยไม่เชิงเส้น (Nonlinear Regression Model) โดยตัวแบบเชิงเส้น หมายถึง ตัวแบบที่สามารถเขียนได้ในแบบเชิงเส้นของพารามิเตอร์ (Linear in Parameter) และตัวแบบไม่เชิงเส้น หมายถึง ตัวแบบที่ไม่สามารถเขียนในแบบเชิงเส้นของพารามิเตอร์ได้ (Nonlinear in Parameter) ตัวแบบการถดถอยไม่เชิงเส้นบางตัวแบบสามารถแปลงให้เป็นตัวแบบการถดถอยเชิงเส้นได้ (Intrinsically Linear) ในขณะที่บางตัวแบบไม่สามารถแปลงเป็นตัวแบบการถดถอยเชิงเส้นได้ (Nonintrinsically Linear) ซึ่งในการประมาณค่าพารามิเตอร์กรณีการวิเคราะห์การถดถอยไม่เชิงเส้นจะมีความยุ่งยากกว่า จำเป็นต้องใช้วิธีการทางคณิตศาสตร์มาช่วย ได้แก่ Numerical Search เช่น วิธีของ Gauss – Newton, Steepest Descent, Marquardt เป็นต้น

การวิเคราะห์การถดถอยเมื่อแบ่งตามวิธีการในการวิเคราะห์ สามารถแบ่งได้เป็น 2 ประเภทใหญ่ ๆ คือ การวิเคราะห์การถดถอยแบบอิงพารามิเตอร์ (Parametric Regression) และการวิเคราะห์การถดถอยแบบไม่อิงพารามิเตอร์ (Nonparametric Regression) การถดถอยแบบอิงพารามิเตอร์ทราบรูปแบบของพารามิเตอร์ในฟังก์ชันการถดถอย แต่ไม่ทราบค่าของพารามิเตอร์ในฟังก์ชันการถดถอย และสามารถประมาณค่าพารามิเตอร์ได้ด้วยวิธีการทางพาราเมตริก ในขณะที่การถดถอยแบบไม่อิงพารามิเตอร์ไม่ทราบรูปแบบของพารามิเตอร์ในฟังก์ชันการถดถอย และไม่ประมาณค่าพารามิเตอร์ แต่จะประมาณเส้นโค้งการถดถอย (Regression Curve) โดยทั่วไปมักใช้เทคนิคการทำให้เรียบ (Smoothing Technique) ความแตกต่างของการถดถอยแบบอิงพารามิเตอร์และการถดถอยแบบไม่อิงพารามิเตอร์ แสดงดังภาพที่ 1

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



ภาพที่ 1 ความแตกต่างของการถดถอยแบบอิงพารามิเตอร์และการถดถอยแบบไม่อิงพารามิเตอร์

การวิเคราะห์การถดถอยนั้นในความเป็นจริงแล้ว ความสัมพันธ์ระหว่างตัวแปรอิสระ X และตัวแปรตาม Y ไม่ใช่ความสัมพันธ์ต่อกันโดยสมบูรณ์ แต่จะมีส่วนของความคลาดเคลื่อน (Error Term / Noise) เข้ามาเกี่ยวข้อง ดังนั้นตัวแบบการถดถอยแบบไม่อิงพารามิเตอร์สามารถกำหนดได้โดยให้ $(\underline{X}_j, \underline{Y}_j)$ เป็นเวกเตอร์สุ่ม k เวกเตอร์ที่เป็นอิสระกัน $i = 1, \ldots, n_j$, $j = 1, \ldots, k$ ดังนั้นสามารถเขียนฟังก์ชันการถดถอยได้ดังนี้

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

$$Y_{ij} = m_j(X_{ij}) + \sigma_j(X_{ij})\varepsilon_{ij} \qquad i = 1, \ldots, n_j,\ j = 1, \ldots, k \qquad (1)$$

เมื่อ $m_j(X_{ij})$ คือ ฟังก์ชันการถดถอยของ Y บน X โดยที่
$m_j(X_{ij}) = E(Y_{ij}|X_{ij})$

$\sigma_j^2(X_{ij})$ คือ ความแปรปรวน โดยที่ $\sigma_j^2(X_{ij}) = Var(Y_{ij}|X_{ij})$

ε_{ij} คือ ความคลาดเคลื่อน โดยที่ $E(\varepsilon_{ij}|X_{ij}) = 0$ และ $Var(\varepsilon_{ij}) = 1$

สำหรับปัญหาที่สนใจศึกษาเกี่ยวกับการวิเคราะห์การถดถอยเป็นส่วนใหญ่ คือการเปรียบเทียบค่าพารามิเตอร์แต่ละค่าที่สนใจ อย่างไรก็ตามหากต้องการที่จะเปรียบเทียบฟังก์ชันการถดถอย $m_j(X_{ij})$ หรืออาจกล่าวได้ว่าเป็นการเปรียบเทียบความแตกต่างระหว่างข้อมูล k กลุ่ม เมื่อมีตัวแปรอิสระ X เข้ามาเกี่ยวข้อง สามารถทำได้โดยกำหนดให้สมมติฐานของการทดสอบอยู่ในรูปของ

$$H_0: m_1 = m_2 = m_3 = \ldots = m_k \text{ เมื่อ } k \geq 2 \qquad (2)$$

$$H_1: m_i \neq m_j \text{ สำหรับ } i, j \text{ บางค่า } i \neq j,\ i, j \in \{1, \ldots, k\} \qquad (3)$$

ในการทดสอบสมมติฐานดังกล่าวข้างต้นโดยวิธีการทางพาราเมตริกสามารถประมาณค่าพารามิเตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุด (Maximum - Likelihood Method) เนื่องจากทราบรูปแบบของพารามิเตอร์ในฟังก์ชันการถดถอย $m_j(X_{ij})$ ดังนั้นสามารถทดสอบสมมติฐานได้โดยกำหนดให้ฟังก์ชันการถดถอยฟังก์ชันที่ 1 $m_1(x) = m_1(x, \beta_1)$ ในที่นี้สามารถระบุได้ว่าตัวพารามิเตอร์ที่ไม่ทราบค่าคือ β_1 ซึ่งสามารถศึกษาได้ดีที่สุดในกรณีของตัวแบบเชิงเส้น (Linear Model) และสามารถขยายไปถึงตัวแบบ $m(x) = \varphi(x^t\beta_1)$ เมื่อ φ ถูกเรียกว่าลิงค์ฟังก์ชัน (Link Function) ซึ่งลักษณะดังกล่าวนี้มักจะเรียกว่าตัวแบบเชิงเส้นวางนัยทั่วไป (Generalized Linear Model : GLM) ในทำนองเดียวกันการประมาณค่าพารามิเตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุดสำหรับฟังก์ชันการถดถอยฟังก์ชันที่ 2 คือ $m_2(x) = m_2(x, \beta_2)$ จะได้ว่าการทดสอบสมมติฐาน $H_0: m_1 = m_2$ คือ การทดสอบสมมติฐาน $H_0: \beta_1 = \beta_2$ เมื่อ β_1, β_2 คือค่าสัมประสิทธิ์การถดถอยของฟังก์ชันการถดถอยที่ 1 และ 2 ตามลำดับ

เนื่องจากการทดสอบการเท่ากันของฟังก์ชันการถดถอยทั้ง 2 ฟังก์ชันนั้น ทราบรูปแบบของพารามิเตอร์ในฟังก์ชันการถดถอย จึงสามารถทำการทดสอบได้จากค่าพารามิเตอร์ในแต่ละ

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

ฟังก์ชันการถดถอยนั้น แต่ถ้าหากไม่ทราบรูปแบบของพารามิเตอร์ในฟังก์ชันการถดถอย $m_j(X_{ij})$ หรือความความคลาดเคลื่อนมีการแจกแจงไม่เป็นไปตามข้อกำหนดของวิธีการทางพาราเมตริก วิธีการทางนอนพาราเมตริกเป็นอีกทางเลือกหนึ่งที่สามารถนำมาใช้แทนวิธีการทางพาราเมตริกได้ ดังนั้นในการวิจัยนี้จึงสนใจศึกษาสถิติทดสอบที่ใช้ในการทดสอบสมมติฐาน

$H_0: m_1 = m_2 = m_3 = ... = m_k$ เมื่อ $k \geq 2$ และ $H_1: m_i \neq m_j$ สำหรับ i, j บางค่า $i \neq j$, $i, j \in \{1, ..., k\}$ ซึ่งมีแนวคิดคือ การเท่ากันของเส้นโค้งการถดถอย คือการเท่ากันของฟังก์ชันการแจกแจงของความความคลาดเคลื่อน (ε_{ij}) กล่าวคือ เมื่อใดก็ตามที่ฟังก์ชันการแจกแจงของความคลาดเคลื่อนของฟังก์ชันการถดถอยที่ทำการเปรียบเทียบไม่มีความแตกต่างกัน เส้นโค้งการถดถอยที่ทำการเปรียบเทียบย่อมไม่แตกต่างกัน ดังนั้นสถิติทดสอบที่ต้องการศึกษาจึงอยู่ในรูปของการทดสอบโดยใช้ฟังก์ชันการแจกแจงของความคลาดเคลื่อน

ซึ่งในการวิจัยนี้กำหนดตัวประมาณความคลาดเคลื่อน (ε_{ij}) ของประชากรกลุ่มที่ j ดังนี้

$$\frac{\overline{Y}_j - \hat{m}_j(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}} \qquad \text{โดยที่} \quad E(\varepsilon_{ij}|X_{ij}) = 0 \quad \text{และ} \quad Var(\varepsilon_{ij}) = 1$$

และกำหนดให้ $\dfrac{\overline{Y}_j - \hat{m}(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}}$ แทนตัวประมาณความคลาดเคลื่อน (ε_{ij}) เมื่อสมมติฐานหลักเป็นจริง

เมื่อ $\overline{Y}_j$ คือ ค่าเฉลี่ยของ Y จากประชากรกลุ่มที่ j, $\overline{Y}_j = \dfrac{\sum_{i=1}^{n_j} Y_{ij}}{n_j}$

$\hat{m}_j(X_{ij})$ คือ ตัวประมาณฟังก์ชันการถดถอย $m_j(X_{ij})$

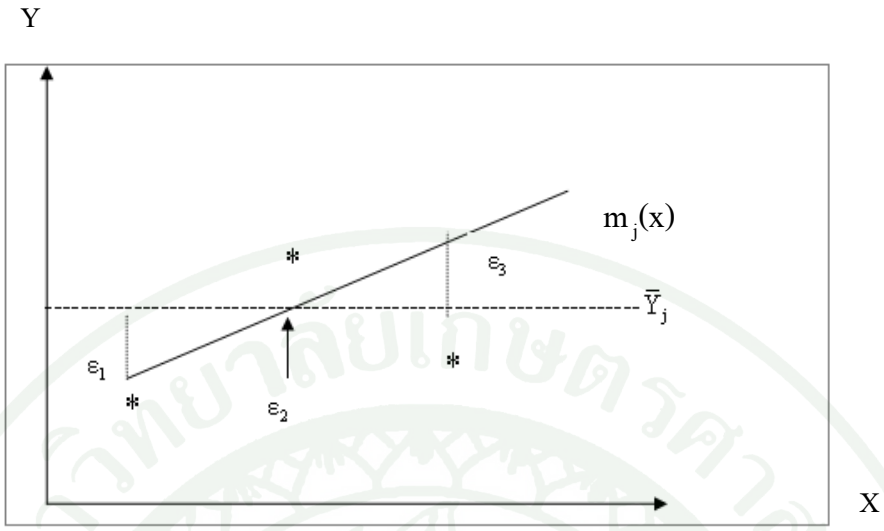
$\hat{m}(X_{ij})$ คือ ตัวประมาณฟังก์ชันการถดถอย $m_j(X_{ij})$ เป็นการประมาณฟังก์ชันการถดถอยเมื่อสมมติฐานหลักเป็นจริง

$\hat{\sigma}_j(X_{ij})$ คือ ตัวประมาณฟังก์ชันของความแปรปรวน $\sigma_j(X_{ij})$

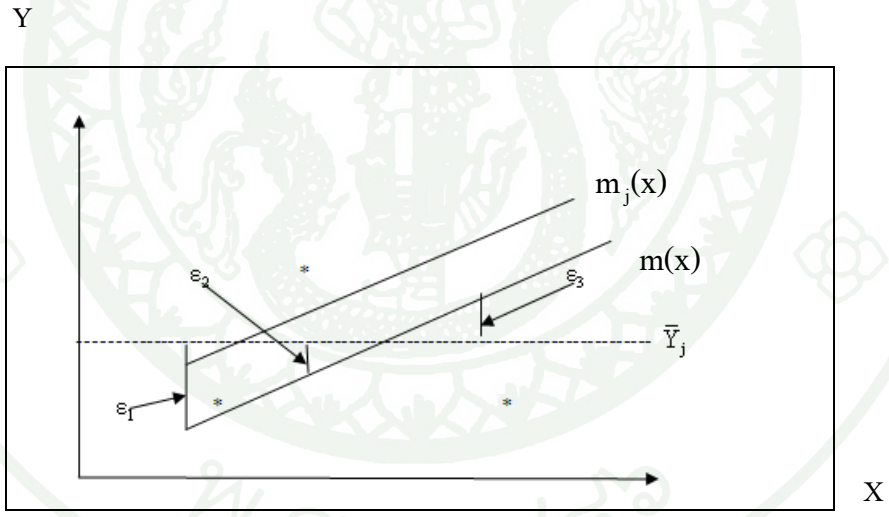
n_j คือ ขนาดตัวอย่างที่สุ่มจากประชากรกลุ่มที่ j

คุณสมบัติของค่าเฉลี่ยของความคลาดเคลื่อนแสดงดังภาพที่ 2 – 3

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



ภาพที่ 2 ค่าเฉลี่ยของความคลาดเคลื่อนเมื่อ $E(\varepsilon_i) = 0$



ภาพที่ 3 ค่าเฉลี่ยของความคลาดเคลื่อนภายใต้สมมติฐานหลักเมื่อ $E(\varepsilon_{i0}) > 0$

จากภาพที่ 2 แสดงให้เห็นว่าเมื่อกำหนดให้ $E(\varepsilon_i)$ แทนค่าเฉลี่ยของความคลาดเคลื่อนจากเส้นโค้งการถดถอยเส้นที่ j $m_j(x)$ จะได้ว่า $E(\varepsilon_i) = 0$ ในขณะเดียวกัน จากภาพที่ 3 เมื่อกำหนดให้ $E(\varepsilon_{i0})$ แทนค่าเฉลี่ยของความคลาดเคลื่อนจาก Common Regression Function $m(x)$ ซึ่งเป็นฟังก์ชันการถดถอยเมื่อสมมติฐานหลักเป็นจริงจะได้ว่า $E(\varepsilon_{i0}) > 0$ หลังจากได้ตัวประมาณความคลาดเคลื่อน (ε_{ij}) ทำการสร้างฟังก์ชันการแจกแจงสะสมของตัวอย่างของตัวประมาณความ

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

คลาดเคลื่อนของประชากรกลุ่มที่ j $F_{\varepsilon j}(y)$ และฟังก์ชันการแจกแจงสะสมของตัวอย่างของตัวประมาณความคลาดเคลื่อนของประชากรกลุ่มที่ j ภายใต้สมมติฐานหลักเป็นจริง $F_{\varepsilon j0}(y)$ จากนั้นสร้างสถิติทดสอบเพื่อใช้ในการทดสอบสมมติฐาน $H_0: m_1 = m_2 = m_3 = ... = m_k$, $k \geq 2$ และ $H_1: m_i \neq m_j$ สำหรับ i, j บางค่า $i \neq j$, $i,j \in \{1, ... , k\}$ หลังจากนั้นจึงทำการศึกษาเกี่ยวกับการแจกแจงของสถิติทดสอบเมื่อสมมติฐานหลักและสมมติฐานทางเลือกเป็นจริง

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

วัตถุประสงค์

1. เพื่อสร้างสถิติทดสอบการเท่ากันของฟังก์ชันการถดถอย k ฟังก์ชัน กรณีที่ไม่ทราบรูปแบบของฟังก์ชันการถดถอย โดยสมมติฐานของการทดสอบอยู่ในรูป $H_0: m_1 = m_2 = m_3 = ... = m_k$ เมื่อ $k \geq 2$ และ $H_1: m_i \neq m_j$ สำหรับ i, j บางค่า $i \neq j,\ i, j \in \{1, ... ,k\}$

2. เพื่อศึกษาเกี่ยวกับการแจกแจงของสถิติทดสอบเท่ากันของฟังก์ชันการถดถอย k ฟังก์ชันภายใต้สมมติฐานหลัก $H_0: m_1 = m_2 = m_3 = ... = m_k$ เมื่อ $k \geq 2$ และสมมติฐานทางเลือก $H_1: m_i \neq m_j$ สำหรับ i, j บางค่า $i \neq j,\ i, j \in \{1, ... ,k\}$ เป็นจริง

ประโยชน์ที่คาดว่าจะได้รับ

1. ได้สถิติทดสอบเพื่อใช้ในการทดสอบการเท่ากันของฟังก์ชันการถดถอย k ฟังก์ชัน $(k \geq 2)$ โดยสมมติฐานของการทดสอบอยู่ในรูป $H_0: m_1 = m_2 = m_3 = ... = m_k$ และ $H_1: m_i \neq m_j$ สำหรับ i, j บางค่า $i \neq j,\ i, j \in \{1, ... ,k\}$ ในกรณีที่ฟังก์ชันการถดถอยปราศจากการกำหนดรูปแบบของพารามิเตอร์

2. ทราบการแจกแจงของสถิติทดสอบการเท่ากันของฟังก์ชันการถดถอย k ฟังก์ชัน ในกรณีที่ฟังก์ชันการถดถอยปราศจากการกำหนดรูปแบบของพารามิเตอร์ ภายใต้สมมติฐานหลัก $H_0: m_1 = m_2 = m_3 = ... = m_k$ และสมมติฐานทางเลือก $H_1: m_i \neq m_j$ สำหรับ i,j บางค่า $i \neq j,\ i, j \in \{1, ... ,k\}$ เป็นจริง

3. เพื่อเป็นแนวทางในการพัฒนาสถิติเพื่อใช้ในการทดสอบการเท่ากันของฟังก์ชันการถดถอย k ฟังก์ชัน $(k \geq 2)$ ในกรณีที่ฟังก์ชันการถดถอยปราศจากการกำหนดรูปแบบของพารามิเตอร์

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

ขอบเขตของการวิจัย

ในการวิจัยครั้งนี้เป็นการพัฒนาสถิติทดสอบ เพื่อใช้ทดสอบสมมติฐาน

$H_0: m_1 = m_2 = m_3 = ... = m_k$ เมื่อ $k \geq 2$ และ
$H_1: m_i \neq m_j$ สำหรับ i,j บางค่า $i \neq j,\ i, j \in \{1, ... ,k\}$

ในกรณีที่ฟังก์ชันการถดถอยปราศจากการกำหนดรูปแบบของพารามิเตอร์ โดยมีขอบเขตของการวิจัยดังนี้

1. สถิติที่พัฒนาขึ้นเพื่อใช้ในการทดสอบสมมติฐานอยู่ในรูปของการทดสอบโดยใช้ฟังก์ชันการแจกแจงของความคลาดเคลื่อน ทำการศึกษาเกี่ยวกับการแจกแจงของสถิติทดสอบภายใต้สมมติฐานหลัก $H_0: m_1 = m_2 = m_3 = ... = m_k$ เมื่อ $k \geq 2$ และ $H_1: m_i \neq m_j$ สำหรับ i,j บางค่า $i \neq j,\ i, j \in \{1, ... ,k\}$ เป็นจริง

2. ประมาณฟังก์ชันการถดถอยโดยวิธีการของ Nadaraya-Watson Estimator (NW) ซึ่งฟังก์ชันการถดถอยที่ใช้ในการทดสอบการเท่ากันอยู่ใน 4 ลักษณะคือ ฟังก์ชันคงที่ ฟังก์ชันเชิงเส้นตรง ฟังก์ชันเอ็กโพเนนเชียล และฟังก์ชันตรีโกณมิติ

3. กำหนดฟังก์ชันการถดถอยเมื่อสมมติฐานทางเลือกเป็นจริง ได้แก่ฟังก์ชันการถดถอยในรูปของ $m_{jn}(x) = m_0(x) + n^{-1/2}s(x)$ โดยศึกษาเมื่อ S(x) แบ่งเป็น 3 ลักษณะคือ

1) $S_1(x) = 1$ (Constant Shift)
2) $S_2(x) = 1 + x$ (Affine Shift)
3) $S_3(x) = 1 + x + x^2$ (Quadratic Shift)

4. ศึกษาประสิทธิภาพของสถิติทดสอบและความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 โดยพิจารณาจากอำนาจการทดสอบ เมื่อขนาดของตัวอย่างมีขนาดเล็ก $n = 20$ ขนาดกลาง $n = 50$ และขนาดใหญ่ $n = 100$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

การตรวจเอกสาร

ในส่วนนี้กล่าวถึงเอกสารและทฤษฎีที่เกี่ยวข้องโดยเสนอเนื้อหาและขอบข่ายดังต่อไปนี้

1. การประมาณฟังก์ชันการถดถอยภายใต้วิธีการทางนอนพาราเมตริก
2. สถิติทดสอบที่เกี่ยวกับฟังก์ชันการแจกแจงสะสมของตัวอย่าง
3. การเปรียบเทียบระหว่างฟังก์ชันการถดถอย k ฟังก์ชัน $(k = 2)$ ภายใต้วิธีการทางพาราเมตริก
4. งานวิจัยที่เกี่ยวข้อง

การประมาณฟังก์ชันการถดถอยภายใต้วิธีการทางนอนพาราเมตริก

ถ้าให้ (X_{ij}, Y_{ij}), $i = 1, \ldots, n_j$, $j = 1, \ldots, k$ และ $n = \sum_{j=1}^{k} n_j$ เป็นตัวอย่างสุ่มจากประชากร k กลุ่ม ซึ่งเป็นอิสระต่อกันและมีการแจกแจงเดียวกัน $E|Y_{ij}| < \infty$ ดังนั้นจะได้ว่าค่าคาดหวังแบบมีเงื่อนไข (Conditional Expectation) ที่เกิดขึ้นคือ $E|Y_{ij}|X_{ij}| = m_j(X_{ij})$ ในทางสถิติฟังก์ชัน m_j เรียกว่าฟังก์ชันการถดถอย (Regression Function) ดังนั้นเมื่อไม่ทราบ m_j จึงจำเป็นต้องประมาณจากตัวอย่างสุ่มที่เป็นอิสระกันจากประชากร k กลุ่ม

$$
\begin{aligned}
&(X_{11}, Y_{11}), \ldots, (X_{n_1 1}, Y_{n_1 1}) \\
&(X_{12}, Y_{12}), \ldots, (X_{n_2 2}, Y_{n_2 2}) \\
&\quad\vdots \qquad\qquad\quad \vdots \\
&(X_{1k}, Y_{1k}), \ldots, (X_{n_k k}, Y_{n_k k})
\end{aligned}
$$

Stone (1977) กล่าวว่าโดยทั่วไปตัวประมาณของ m_j จะอยู่ในรูปของ

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

$$\hat{m}_1(x) = \sum_{i=1}^{n_1} Y_{1i} W_{1n_i}(x)$$

$$\hat{m}_2(x) = \sum_{i=1}^{n_2} Y_{2i} W_{2n_i}(x)$$

$$\vdots \qquad\qquad \vdots$$

$$\hat{m}_k(x) = \sum_{i=1}^{n_k} Y_{ki} W_{kn_i}(x)$$

เมื่อ x คือ จุดที่ต้องการประมาณฟังก์ชันการถดถอย

X_{ij} คือ ตัวแปรอิสระค่าที่ i กลุ่มที่ j, $i = 1, \ldots, n_j$, $j = 1, \ldots, k$

โดยที่ $W_{jn_i}(x) = W_{jn_i}(x, X_{11}, \ldots, X_{1n_i}) \in \square$ คือ ค่าถ่วงน้ำหนัก (Weight) ซึ่งมีสมบัติว่า

$$\sum_{i=1}^{n_1} W_{1n_i}(x) = \sum_{i=1}^{n_2} W_{2n_i}(x) = \ldots = \sum_{i=1}^{n_k} W_{kn_i}(x) = 1$$

เนื่องจากในการศึกษานี้ต้องการประมาณฟังก์ชันการถดถอยมากกว่าประมาณตัวพารามิเตอร์ที่ไม่ทราบค่า $W_{jn_i}(x)$ จะขึ้นอยู่กับ x หรืออาจกล่าวได้ว่า $W_{jn_i}(x)$ จะให้น้ำหนักมากกับ X_{ij} ที่อยู่ใกล้กับจุดที่ต้องการประมาณฟังก์ชันการถดถอย (x) และในขณะเดียวกันจะให้น้ำหนักน้อยกับ X_{ij} ที่อยู่ไกลจากจุดที่ต้องการประมาณฟังก์ชันการถดถอยออกไป

วิธีการประมาณฟังก์ชันการถดถอยที่นิยมใช้กันอย่างแพร่หลายมีอยู่ด้วยกันหลายวิธีในที่นี้ได้เสนอวิธีการประมาณฟังก์ชันการถดถอยที่มีประสิทธิภาพ 2 วิธีคือ 1) Nadaraya – Watson estimator (NW) และ 2) k – Nearest Neighbor Estimator (k – nn)

1) วิธี Nadaraya-Watson Estimator (NW)

Hardle (1993) ได้กล่าวถึง Nadaraya-Watson estimator (NW) โดยให้ K แทน kernel function ซึ่งมีลักษณะสมมาตร h คือ พารามิเตอร์ปรับให้เรียบเรียกว่า bandwidth หรือ window - width ซึ่งมีค่าเป็นบวก ($h > 0$) ดังนั้นสามารถประมาณฟังก์ชันการถดถอยได้ดังนี้

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

$$\hat{m}_j(x) = \frac{\sum_{i=1}^{n_j} K\left(\frac{X_{ij} - x}{h}\right) Y_{ij}}{\sum_{i=1}^{n_j} K\left(\frac{X_{ij} - x}{h}\right)} = \sum_{i=1}^{n_j} W_{jn_i}(x) Y_{ij} ,$$

$i = 1, \ldots, n_j$, $j = 1, \ldots, k$ เมื่อค่าถ่วงน้ำหนักคือ $W_{jn_i}(x)$

$$W_{jn_i}(x) = \frac{K\left(\frac{X_{ij} - x}{h}\right)}{\sum_{i=1}^{n_j} K\left(\frac{X_{ij} - x}{h}\right)}$$

2) วิธี k – Nearest Neighbor Estimator (k - nn)

Stone (1977) ได้เสนอแนะให้ใช้ค่าถ่วงน้ำหนักโดยกำหนดให้ $R = R_n$, $n = 1, \ldots, k$ โดย R_n เป็นระยะทางที่ใกล้ที่สุดระหว่าง x และ X_{ij} ตัวที่ n จำนวน k ตัว โดยค่าถ่วงน้ำหนัก $W_{jn_i}(x)$ อยู่ในรูป

$$W_{jn_i}(x) = \frac{K\left(\frac{X_{ij} - x}{R}\right)}{\sum_{i=1}^{n_j} K\left(\frac{X_{ij} - x}{R}\right)}$$

ดังนั้นตัวประมาณ k – nearest neighbor estimator (k - nn) จึงอยู่ในรูป

$$\hat{m}_j(x) = \frac{\sum_{i=1}^{n_j} K\left(\frac{X_{ij} - x}{R}\right) Y_{ij}}{\sum_{i=1}^{n_j} K\left(\frac{X_{ij} - x}{R}\right)} = \sum_{i=1}^{n_j} W_{jn_i}(x) Y_{ij}$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

วิธีการของ Nadaraya-Watson Estimator (NW) และ k – Nearest Neighbor Estimator (k - nn) มีลักษณะที่คล้ายคลึงกัน แตกต่างกันเพียงพารามิเตอร์ปรับให้เรียบในวิธีการของ NW คือ h และพารามิเตอร์ปรับให้เรียบในวิธีการของ k - nn คือ R_n เมื่อ R_n เป็นระยะทางที่ใกล้ที่สุดระหว่าง x และ X_{ij} ตัวที่ n จำนวน k ตัว แต่อย่างไรก็ตามพารามิเตอร์ปรับให้เรียบจากทั้ง 2 วิธีนี้มีความหมายในลักษณะเดียวกันคือ เมื่อ h และ R_n มีค่าสูง ฟังก์ชันการถดถอยที่ได้จะเรียบ (Smooth) แต่ไม่เหมาะสมกับข้อมูล ในทางตรงกันข้ามเมื่อ h และ R_n มีค่าต่ำ ฟังก์ชันการถดถอยจะเหมาะสมกับข้อมูล (Well - Fitting) แต่ไม่เรียบ ดังนั้นในการใช้วิธีการทางนอนพาราเมตริกประมาณฟังก์ชันการถดถอยจะต้องเป็นการเลือก (Trade Off) ระหว่างความเรียบและความเหมาะสมกับข้อมูล

เมื่อพิจารณาค่าเฉลี่ยกำลังสองของความคลาดเคลื่อน (Mean Square Error : MSE) $MSE\left[\hat{m}(x)\right] = E\left[(\hat{m}(x) - m(x))^2\right]$ พบว่า $MSE\left[\hat{m}(x)\right]$ สามารถเขียนให้อยู่ในรูปของความสัมพันธ์ระหว่าง ความเอนเอียง (Bias) และความแปรปรวน (Variance) ได้ดังนี้ $MSE\left[\hat{m}(x)\right] = (Bias\left[\hat{m}(x)\right])^2 + Var\left[\hat{m}(x)\right]$ ซึ่งในการประมาณฟังก์ชันการถดถอยนั้น ถ้าให้ความสำคัญกับความเหมาะสมของฟังก์ชันการถดถอยกับข้อมูล (h และ R_n มีค่าต่ำ) ความแปรปรวนจะมีค่าสูง ในทางตรงกันข้ามถ้าให้ความสำคัญกับความเรียบ (h และ R_n มีค่าสูง) ความเอนเอียงจะมีค่าสูง หรืออาจกล่าวได้ว่าต้องเป็นการเลือกระหว่างความเอนเอียงและความแปรปรวน (Takezawa, 1999)

สถิติทดสอบที่เกี่ยวกับฟังก์ชันการแจกแจงสะสมของตัวอย่าง

การทดสอบลักษณะต่าง ๆ ของประชากร ซึ่งอาจเป็นการทดสอบสัดส่วนของลักษณะต่าง ๆ ของประชากรว่าเป็นไปตามที่คาดไว้หรือไม่ หรือการทดสอบการแจกแจงของประชากรว่ามีการแจกแจงต่าง ๆ ตามที่คาดไว้หรือไม่ เรียกการทดสอบลักษณะนี้ว่าการทดสอบภาวะสารูปสนิทดี (Goodness of Fit Test) ซึ่งลักษณะการทดสอบภาวะสารูปสนิทดีอย่างหนึ่ง คือ การทดสอบโดยใช้ฟังก์ชันการแจกแจงของตัวอย่าง ซึ่งสถิติทดสอบที่เป็นที่รู้จัก ได้แก่ สถิติทดสอบ KS (Kolmogorov-Smirnov test) สถิติทดสอบ V (Kuiper V test) สถิติทดสอบ C (Pyke C test) สถิติทดสอบ B (Brunk B test) สถิติทดสอบ D (Derbin D test) สถิติทดสอบ W^2 (Cramer-Von Mises W^2 test) สถิติทดสอบ M^2 (Derbin M^2 test) และสถิติทดสอบ U^2 (Watson U^2 test) (Hegazy and Green, 1975)

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

ฟังก์ชันการแจกแจงของตัวอย่างของตัวอย่าง $\{x_j\}$, $j = 1, \ldots, n$ สามารถกำหนดได้โดย

$$F(x) = \frac{1}{n} (\text{number of } x_j \leq x)$$

$$= \frac{1}{n}\sum_{j=1}^{n} I(x_j \leq x)$$

เมื่อ I คือ Indicator Function

ถ้าตัวอย่างถูกแบ่งเป็นตัวอย่างย่อย (Sub Sample) จำนวน i กลุ่มเพื่อสร้างฟังก์ชันการแจกแจงของตัวอย่าง F_i ดังนั้นฟังก์ชันการแจกแจงของตัวอย่างรวม (Pooled Sample) สามารถอธิบายได้โดย

$$F = \frac{1}{n}\sum_{i} n_i F_i$$

เมื่อ n_i คือ จำนวนค่าสังเกตในกลุ่มที่ i

n คือ จำนวนค่าสังเกตทั้งหมด

สถิติทดสอบที่เกี่ยวข้องกับฟังก์ชันการแจกแจงของตัวอย่างที่เกี่ยวข้องกับการศึกษาครั้งนี้ประกอบด้วยสถิติทดสอบ Kolmogorov - Smirnov test และ Kuiper test โดยมีรายละเอียดดังนี้

1. สถิติทดสอบ Kolmogorov - Smirnov test

Kolmogorov - Smirnov มีหลักการใช้วัดค่าเบี่ยงเบนสูงสุด (Maximum Deviation) ของฟังก์ชันการแจกแจงของตัวอย่างภายในแต่ละกลุ่มจากฟังก์ชันการแจกแจงของตัวอย่างรวม คำนวณได้โดย

$$KS = \max_j \sqrt{\frac{1}{n}\sum_i n_i \left(F_i(x_j) - F(x_j)\right)^2}, \quad \text{เมื่อ} \quad j = 1, \ldots, n$$

2. สถิติทดสอบ Kuiper V test คำนวณได้โดย

$$V = \max_j \left(F_i(x_j) - F(x_j)\right) - \min_j \left(F_i(x_j) - F(x_j)\right), \quad \text{เมื่อ} \quad j = 1, \ldots, n$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

การเปรียบเทียบระหว่างฟังก์ชันการถดถอย k ฟังก์ชัน (k = 2)
ภายใต้วิธีการทางพาราเมตริก

ในการทดสอบการเท่ากันของสัมประสิทธิ์การถดถอย ระหว่าง 2 กลุ่มประชากรที่เป็นอิสระต่อกันหรือการทดสอบการเท่ากันของฟังก์ชันการถดถอย 2 ฟังก์ชันที่เป็นอิสระต่อกัน ซึ่งทั้ง 2 ฟังก์ชันการถดถอยสามารถเปรียบเทียบกันได้บนเวกเตอร์ของตัวแปรตาม $\underline{Y}$ และเวกเตอร์ของตัวแปรอิสระ $\underline{X}$ โดย y_1 และ y_2 เป็นค่าวัดของตัวแปรตาม Y จากประชากรกลุ่มที่ 1 และประชากรกลุ่มที่ 2 ตามลำดับ x_1 และ x_2 เป็นค่าวัดของตัวแปรอิสระ X จากประชากรกลุ่มที่ 1 และประชากรกลุ่มที่ 2 ตามลำดับ ตัวอย่างเช่น การศึกษาเกี่ยวกับข้อมูลการผลิตจากสายการผลิตที่ 1 และสายการผลิตที่ 2 ให้ y_1 และ y_2 แทนค่าวัดของตัวแปรตามเป็นข้อมูลเกี่ยวกับจำนวนชิ้นงานที่มีข้อบกพร่องจากสายการผลิตที่ 1 และสายการผลิตที่ 2 ตามลำดับ ให้ x_1 และ x_2 แทนค่าวัดของตัวแปรอิสระเป็นข้อมูลเกี่ยวกับอัตราเร็วในการผลิตจากสายการผลิตที่ 1 และสายการผลิตที่ 2 ตามลำดับ ต้องการศึกษาว่าตัวแปรอิสระ X_1 จากสายการผลิตที่ 1 และตัวแปรอิสระ X_2 จากสายการผลิตที่ 2 ส่งผลต่อตัวแปรตาม Y_1 และ Y_2 แตกต่างกันหรือไม่อย่างไร ซึ่งถ้าข้อมูลดังกล่าวไม่มีตัวแปรอิสระ X การศึกษาดังกล่าวอาจเทียบได้กับการเปรียบเทียบค่าเฉลี่ยของประชากร k กลุ่ม ซึ่งอาจใช้หลักการของการวิเคราะห์ความแปรปรวน แต่ในที่นี้รูปแบบความสัมพันธ์ระหว่างตัวแปรตาม Y และตัวแปรอิสระ X อยู่ในรูปแบบของการวิเคราะห์การถดถอย ดังนั้นการทดสอบดังกล่าวสามารถทำได้โดยการเปรียบเทียบค่าสัมประสิทธิ์การถดถอยของแต่ละฟังก์ชันที่เป็นอิสระต่อกัน

Brame *et al.* (1998) ได้ศึกษาเกี่ยวกับการทดสอบสมมติฐานที่อยู่ในรูป $\delta : \theta_a - \theta_b = 0$ เมื่อ θ_a , θ_b คือ ค่าสัมประสิทธ์การถดถอยของฟังก์ชันที่ 1 และ 2 ตามลำดับ

ในการเปรียบเทียบค่าสัมประสิทธิ์การถดถอยของ 2 ฟังก์ชันการถดถอยที่เป็นอิสระกัน สิ่งที่ต้องทราบประกอบด้วย 2 ส่วนคือ

1) ตัวประมาณของความแตกต่างระหว่างค่าสัมประสิทธิ์การถดถอยที่ถูกสุ่มมาจากแต่ละประชากรที่เป็นอิสระต่อกัน

2) ส่วนเบี่ยงเบนมาตรฐานของตัวประมาณใน 1)

ซึ่งในการประมาณค่าพารามิเตอร์หรือค่าสัมประสิทธิ์การถอยได้ใช้วิธีภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Estimator : MLE) เนื่องจากวิธี MLE จะให้ตัวประมาณที่มีคุณสมบัติที่ดีดังต่อไปนี้

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

1. เป็นตัวประมาณที่คงเส้นคงวา (Consistent Estimator) กล่าวคือเมื่อขนาดของตัวอย่างเพิ่มขึ้น $(n \to \infty)$ $\lim_{n\to\infty} \Pr(|\hat{\theta} - \theta| > \varepsilon) = 0$ เมื่อ $\varepsilon > 0$ หรือกล่าวอีกนัยหนึ่งว่า $\hat{\theta}$ เป็นตัวประมาณที่คงเส้นคงวาของ θ ถ้า $\hat{\theta}$ ลู่เข้าหา θ ในเชิงความน่าจะเป็น (Converge in Probability) เมื่อ n มีค่ามาก ๆ หรือ $n \to \infty$

2. ตัวประมาณภาวะน่าจะเป็นสูงสุดเป็นตัวประมาณที่ได้จากการ Maximized Likelihood Function โดย

$$L(\underline{\theta}|y) = k(y)\Pr(y|\underline{\theta}) \propto \Pr(y|\underline{\theta})$$

ซึ่งแสดงให้เห็นว่า Likelihood Function แปรผันตามฟังก์ชันความหนาแน่นน่าจะเป็นโดยทั่วไปค่าคงที่ k(y) ไม่สามารถประมาณค่าได้ (Not Estimable) อย่างไรก็ตามค่าคงที่จะไม่เปลี่ยนแปลงเซตของค่าประมาณพารามิเตอร์ที่แตกต่างกัน ดังนั้นจึงเป็นไปได้ที่จะเปรียบเทียบ Likelihood ของค่าประมาณพารามิเตอร์ของเซตหนึ่งกับ Likelihood ของเซตอื่น

3. ภายใต้การสุ่มตัวอย่างเมื่อตัวอย่างมีขนาดใหญ่ การแจกแจงของตัวอย่างสุ่มที่ได้จากวิธีภาวะน่าจะเป็นสูงสุด มีการแจกแจงแบบปกติ โดยตัวประมาณที่คงเส้นคงวาของความแปรปรวนของ $\hat{\theta}$ กำหนดโดย

$$\hat{V}(\hat{\theta}) = \frac{-1}{n}\left(\frac{\partial^2 \log L(\underline{\theta}|y)}{\partial\underline{\theta}\partial\underline{\theta}'}\right)^{-1}$$

4. ถ้ามีตัวประมาณที่มีประสิทธิภาพภายใต้เงื่อนไขปกติ ตัวประมาณแบบภาวะน่าจะเป็นสูงสุด จะเป็นตัวประมาณที่มีประสิทธิภาพ

ดังนั้น เมื่อได้ค่าประมาณพารามิเตอร์ ซึ่งในที่นี้คือค่าสัมประสิทธิ์การถดถอย การทดสอบจะอยู่ในรูป $\delta = 0$ ซึ่งก็คือ $\theta_a = \theta_b$ หรือ $\theta_a - \theta_b = 0$ เมื่อ θ_a, θ_b ค่าสัมประสิทธิ์การถดถอยของฟังก์ชันที่ 1 และ 2 ตามลำดับ

Brame *et al.* (1998 *Cited* Johnson and Kotz 1970) ได้เสนอแนวคิดที่มีประโยชน์ต่อการทดสอบในที่นี้โดย ถ้า $X_1, X_2, ..., X_n$ เป็นตัวแปรสุ่มที่มีการแจกแจงแบบปกติ และเป็นอิสระต่อกัน ดังนั้นฟังก์ชันเชิงเส้น (Linear Function) ของตัวแปรเหล่านี้จะมีการแจกแจงแบบปกติด้วย

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

ดังนั้น $\hat{\delta} = \hat{\theta}_a - \hat{\theta}_b \sim N(\delta, \sigma_{\hat{\delta}}^2)$ เมื่อ δ คือค่าเฉลี่ยของผลต่างของตัวแปรสุ่ม $\theta_a - \theta_b$ และ $\sigma_{\hat{\delta}}^2$ คือความแปรปรวนผลต่างของตัวแปรสุ่ม $\theta_a - \theta_b$

โดยทั่วไปตัวประมาณความแปรปรวนผลต่างของตัวแปรสุ่ม $\theta_a - \theta_b$ ถูกกำหนดโดย

$$\sigma_{\hat{\delta}}^2 = \sigma_{\hat{\theta}_a}^2 + \sigma_{\hat{\theta}_b}^2 - 2COV(\theta_a, \theta_b)$$

ในกรณีของ 2 ตัวแปรสุ่มที่มีการแจกแจงแบบปกติและเป็นอิสระต่อกัน จะได้ว่า $\sigma_{\hat{\delta}}^2 = \sigma_{\hat{\theta}_a}^2 + \sigma_{\hat{\theta}_b}^2$ ในขณะเดียวกัน ส่วนเบี่ยงเบนมาตรฐานของผลต่างของตัวแปรสุ่ม $\theta_a - \theta_b$ คือ $\sigma_{\hat{\delta}} = \sqrt{\sigma_{\hat{\delta}}^2} = \sqrt{\sigma_{\hat{\theta}_a}^2 + \sigma_{\hat{\theta}_b}^2}$ ดังนั้นในการทดสอบสมมติฐาน $H_0: \delta = 0$ และ $H_1: \delta \neq 0$ สามารถทำได้โดยใช้สถิติทดสอบ

$$Z = \frac{\hat{\delta} - 0}{\sigma_{\hat{\delta}}}$$

ซึ่งการแปลผลสามารถทำได้เช่นเดียวกับการทดสอบสมมติฐานทางสถิติทั่วไป นอกจากนั้น Brame *et al.* (1998 *Cited* Koo and Hong 1980) ได้พิจารณาส่วนเบี่ยงเบนมาตรฐานของผลต่างของตัวแปรสุ่ม $\theta_a - \theta_b$ ในอีกลักษณะหนึ่งโดย

$$\sigma_{\hat{\delta}} = \sqrt{\sigma_{\hat{\delta}}^2} = \sqrt{\frac{\nu_a \sigma_{\hat{\theta}_a}^2 + \nu_b \sigma_{\hat{\theta}_b}^2}{\nu_a + \nu_b}}$$

เมื่อ ν_a และ ν_b คือ องศาความเป็นอิสระ (Degree of Freedom) ที่สอดคล้องกับข้อมูลกลุ่ม a และ b ตามลำดับ โดยสถิติทดสอบอยู่ในรูปของ

$$t = \frac{\hat{\delta} - 0}{\sqrt{\frac{\nu_a \sigma_{\hat{\theta}_a}^2 + \nu_b \sigma_{\hat{\theta}_b}^2}{\nu_a + \nu_b}}}$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

งานวิจัยที่เกี่ยวข้อง

การเปรียบเทียบระหว่างฟังก์ชันการถดถอย k ฟังก์ชันการถดถอย ($k \geq 2$) ภายใต้วิธีการทางนอนพาราเมตริกนั้น เนื่องจากไม่ทราบรูปแบบของฟังก์ชันการถดถอย m(.) หรืออาจกล่าวได้ว่าเป็นการประมาณค่าฟังก์ชันการถดถอยที่ปราศจากการกำหนดรูปแบบของพารามิเตอร์ การเปรียบเทียบระหว่างฟังก์ชันการถดถอยโดยพิจารณาจากผลต่างระหว่างสัมประสิทธิ์การถดถอยเช่นเดียวกับวิธีการทางพาราเมตริกซ์ไม่สามารถทำได้ ดังนั้นได้มีผู้เสนอแนวคิดต่าง ๆ เกี่ยวกับ การเปรียบเทียบระหว่างฟังก์ชันการถดถอย k ฟังก์ชันการถดถอย ($k \geq 2$) ดังนี้

Delgado (1993) ทดสอบการเท่ากันของฟังก์ชันการถดถอยโดยกำหนดให้ $\{(X_i, Y_i, Z_i),\ i = 1, \ldots, n\}$ เป็นตัวอย่างสุ่มจากตัวแปรสุ่ม $\{X, Y, Z\}$ ตัวแปรตาม Y และ Z สัมพันธ์กับตัวแปรอิสระ X ตามตัวแบบการถดถอย

$$E(Y|X = x) = m_1(x) \quad \text{และ} \quad E(Z|X = x) = m_2(x)$$

กำหนดให้ความคลาดเคลื่อนของการถดถอยคือ $Y - m_1(x)$ และ $Z - m_2(x)$ เป็นอิสระกับ X และอาจมีการแจกแจงที่แตกต่างกันได้ $E|m_1(x)|^2 < \infty$, $E|m_2(x)|^2 < \infty$, $0 < \sigma_Y^2 = E|Y - m_1(X)|^2 < \infty$ และ $0 < \sigma_Z^2 = E|Z - m_2(X)|^2 < \infty$

สมมติฐานของการทดสอบคือ

$H_0: m_1(x) = m_2(x)$ สำหรับ $x \in X$ ทุกค่า

$H_1: m_1(x) \neq m_2(x)$ สำหรับ $x \in X$ บางค่า

ในการศึกษานี้ X เป็นตัวแปรสุ่มที่มีค่าอยู่ระหว่าง 0 ถึง 1 ความคลาดเคลื่อนของการถดถอยเป็นอิสระกัน เมื่อพิจารณาตามสมมติฐานหลัก การทดสอบจะอยู่ในรูปของ

$$\sup_{-\infty < t < \infty} \left| \int_{-\infty}^{t} (m_1(x) - m_2(x)) f(x) dx \right| = 0 \qquad (4)$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

กำหนดให้ $D_i = Y_i - Z_i$ จะได้ว่า Weighted Empirical Process คือ

$$\sup_{-\infty< t < \infty}\left|n^{-1}\sum_{j=1}^{n} D_j I(X_j < t)\right|$$

ดังนั้นจะได้ว่าการประมาณค่าที่สอดคล้องกับสมการทางซ้ายมือของสมการที่ (4) เมื่อ I(A) คือ Indicator Function ของเหตุการณ์ A สถิติทดสอบของ Delgado (1993) จึงอยู่ในรูปของ

$$T_n = \left[\sum_{j=1}^{n-1}(D_{j+1} - D_j)\right]^{-1/2} \sup_{-\infty < t < \infty}\left|\sum_{j=1}^{n} D_j I(X_j < t)\right| \qquad (5)$$

ซึ่งจะมีค่าใหญ่ภายใต้สมมติฐานทางเลือกและจะมีค่าเล็กภายใต้สมมติฐานหลัก และจากการประยุกต์ใช้ Kolmogorov ' s Law of Large Number (KLLN) จะได้ว่า

$(2n)^{-1}\sum_{j=1}^{n-1}(D_{j+1} - D_j)^2/2 \xrightarrow{\text{w. p. 1}} \sigma^2$ เมื่อ $n \to \infty$ ดังนั้นสมการ (5) จะอยู่ในรูปของ

$\sup_{-\infty< t < \infty}\left|(n\sigma^2)^{-1/2}\sum_{i=1}^{n} D_i I(X_i < t)\right|$ เพื่อที่จะตรวจสอบเกี่ยวกับ Asymptotic Distribution

ของสถิติทดสอบภายใต้สมมติฐานหลัก พบว่าภายใต้สมมติฐานหลักเป็นจริง

$$\sup_{-\infty< t < \infty}\left|(n\sigma^2)^{-1/2}\sum_{i=1}^{n} D_i I(X_i < t)\right| \xrightarrow{d} T = \sup_{0\le t\le 1}|B(t)|$$ เมื่อ $n \to \infty$ B(t)

คือ Standard Brownian Motion กำหนด T_α โดยที่ $Pr(T > T_\alpha) = \alpha$ ดังนั้น

$\lim_{n\to\infty} Pr(T_n > T_\alpha) = \alpha$ ภายใต้สมมติฐานหลักเป็นจริง และ $\lim_{n\to\infty} Pr(T_n > T_\alpha) = 1$

ภายใต้สมมติฐานทางเลือกเป็นจริง ดังนั้นสมมติฐานหลักจะถูกปฏิเสธที่ระดับนัยสำคัญ α เมื่อ $T_n > T_\alpha$ ค่าวิกฤตและ P – Value สามารถคำนวณได้จาก

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

$$\Pr(T > b) = 1 - 4\pi^{-1}\sum_{j=0}^{\infty}(-1)^{j}(2j+1)^{-1}\exp\left\{-(2j+1)^{2}\pi^{2}/(8b^{2})\right\},\ b > 0$$

ภายใต้สมมติฐานทางเลือก

$$T_n \xrightarrow{d} \sup_{0<t<1}\left|C(\sigma^{2})^{-1/2}\int_{0}^{t}|f(x)|\,f(x)dx + B(t)\right|$$ เมื่อ $n \to \infty$, C คือ ค่าคงที่

และ $\int_{0}^{t}|f(x)|\,f(x)dx = 0$ ก็ต่อเมื่อ $f(x) = 0$ สำหรับ x ทุกค่า

การทดสอบสมมติฐานสามารถขยายไปสู่การทดสอบการเท่ากันของฟังก์ชันการถดถอย k ฟังก์ชันได้ดังนี้

สมมติว่ามีตัวอย่างสุ่ม $\left\{(X_i, Y_1^{(i)}, \ldots, Y_n^{(p)}),\ i = 1, \ldots, n\right\}$ จากตัวแปรสุ่ม $(X, Y^{(1)}, \ldots, Y^{(p)})$ ตัวแปร $Y^{(1)}, \ldots, Y^{(p)}$ มีความสัมพันธ์กับ X ดังตัวแบบการถดถอย

$E(Y^{(k)}|X = x) = m_k(x),\ k = 1, \ldots, p$ ต้องการทดสอบสมมติฐานที่อยู่ในรูป

$H_0: m_k(x) = m_m(x)$ สำหรับ $m \neq k$ ทุกค่า $m, k = 1, \ldots, p$ และ $x \in \square$

$H_1: m_k(x) \neq m_m(x)$ สำหรับ $m \neq k$ บางค่า $m, k = 1, \ldots, p$ และ $x \in \square$

กำหนดให้ $\bar{Y}_j = p^{-1}\sum_{j=1}^{p}Y^{(k)}$, $D_j^k = Y_j^{(k)} - \bar{Y}_j$ และ

$$T_n^k = \left[\sum_{j=1}^{n-1}(D_{j+1}^k - D_j^k)^2/2\right]^{-1/2}\sup_{-\infty<t<\infty}\left|\sum_{j=1}^{n}D_j^k I(X < t)\right|$$

สถิติทดสอบคือ $T_n = \max_{1 \le k \le p} T_n^k$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

Young and Bowman (1995) ศึกษาเกี่ยวกับปัญหาการทดสอบการเท่ากันของฟังก์ชันการถดถอย k ฟังก์ชันโดยตัวแบบการถดถอยอยู่ในรูป

$$Y_{ij} = m_j(X_{ij}) + \sigma_j(X_{ij})\varepsilon_{ij} \qquad i = 1, \ldots, n_j,\ j = 1, \ldots, k$$

สมมติฐานของการทดสอบอยู่ในรูป

$$H_0: m_1 = m_2 = m_3 = \ldots = m_k \text{ เมื่อ } k \geq 2$$
$$H_1: m_i \neq m_j \text{ สำหรับ } i, j \text{ บางค่า } i \neq j,\ i, j \in \{1, \ldots, k\}$$

โดยข้อสมมติตลอดจนเงื่อนไขต่าง ๆ ของการทดสอบสมมติฐานประกอบไปด้วย

1. $N = \sum_{i=1}^{k} n_i$

$$\frac{n_i}{N} = k_i + O\left(\frac{1}{n}\right),\ i = 1,\ldots,k \text{ เมื่อ } k_1, \ldots, k_k \in (0,1)$$

2. ให้ $f_1(x), \ldots, f_k(x)$ แทน ฟังก์ชันความหนาแน่นน่าจะเป็นของตัวแปรอิสระ X_j $j = 1, \ldots, k$ ที่อยู่ในช่วง [0,1]

3. $\hat{m}_j(x) = \dfrac{\sum_{i=1}^{n_j} K\left(\dfrac{X_{ij} - x}{h}\right) Y_{ij}}{\sum_{i=1}^{n_j} K\left(\dfrac{X_{ij} - x}{h}\right)}$ แทนตัวประมาณ Nadaraya – Watson ของฟังก์ชันการถดถอยที่ j และ h คือพารามิเตอร์ปรับให้เรียบ

4. การประมาณ Common Regression ที่คำนวณจากตัวอย่างทั้งหมดอยู่ในรูปของ

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

$$\hat{m}(x) = \frac{\sum_{j=1}^{k}\sum_{i=1}^{n_j} K\left(\frac{x - X_{ij}}{h}\right) Y_{ij}}{\sum_{j=1}^{k}\sum_{i=1}^{n_j} K\left(\frac{x - X_{ij}}{h}\right)}$$

ซึ่งเป็นการประมาณฟังก์ชันการถดถอยสำหรับกรณีพารามิเตอร์ปรับให้เรียบ $h_i = h$, $i = 1, \ldots, k$ แนวคิดพื้นฐานของ Young and Bowman (1995) มาจากการวิเคราะห์ความแปรปรวนทางเดียว (One – Way Analysis of Variance) โดยสถิติในการทดสอบสมมติฐานอยู่ในรูป

$$T_n^{(2)} = \frac{1}{N}\sum_{j=1}^{k}\sum_{i=1}^{n_j}\left(\hat{m}(X_{ij}) - \hat{m}_j(X_{ij})\right)^2$$

เมื่อ $\hat{m}$, $\hat{m}_1, \ldots, \hat{m}_k$ คำนวณมาจากข้อสมมติข้อที่ 3 และ 4 ตามลำดับ

นอกจากนี้ Young and Bowman (1995) ได้แสดงให้เห็นว่าภายใต้สมมติฐานหลักหรือการเท่ากันของแต่ละฟังก์ชันการถดถอยเป็นจริง สถิติทดสอบ $T_n^{(2)}$ มีคุณสมบัติดังนี้

$$N\sqrt{h}\left(T_N^{(2)} + B_K^{(2)}h^{2r} - \frac{D_k^{(2)}}{Nh}\right) \xrightarrow{d} N\left(0, \beta_{k,2}^2\right) \quad \text{โดย}$$

$$D_k^{(2)} = \int_{-1}^{1} K^2(u)du \sum_{j=1}^{k}\int_0^1\left(1 - k_j\frac{f_j(x)}{\bar{R}(x)}\right)\sigma_j^2(x)\,dx$$

$\bar{R}$ แทน Convex Combination ภายใต้ฟังก์ชันความหนาแน่นน่าจะเป็น กล่าวคือ

$$\bar{R}(t) = \sum_{j=1}^{k} k_j f_j(x)$$

$$\beta_k^2 = k_r^2\int_0^1((m_1\bar{R})^{(r)} - m_1(\bar{R})^{(r)})^2(x)\frac{dt}{\bar{R}(x)} - k_r^2\sum_{j=1}^{k}k_j\int_0^1((m_r f_j)^{(r)} - m_1 f_j^{(r)})^2(x)\frac{dx}{f_j(x)}$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

$$\beta_{k,2}^2 = 2\int_{-1}^{1}(K*K)^2(u)du\left[\sum_{j=1}^{k}\int_{0}^{1}\sigma_j^4(x)\left(\frac{k_j f_j(x)}{\bar{R}(x)} - 1\right)^2 dx\right.$$

$$\left. + \sum_{j=1}^{k}\sum_{\substack{l=1 \\ l\neq j}}^{k}\int_{0}^{1}\sigma_j^2(x)\sigma_l^2(x)\frac{k_j f_j(x) k_l f_l(x)}{\bar{R}^2(x)}dx\right.$$

ภายใต้สมมติฐานทางเลือก $H_1: m_i \neq m_j$ สำหรับ i, j บางตัว $i \neq j$, $i, j \in \{1, \ldots, k\}$ สถิติทดสอบ $T_n^{(2)}$ มีคุณสมบัติดังนี้

$$\sqrt{N}\left(T_N^{(2)} - M_{k,2}^2\right) \xrightarrow{d} N\left(0, \gamma_{k,2}^2\right)$$

เมื่อ $$M_{k,2}^2 = \sum_{j=1}^{k}\sum_{\substack{l=1 \\ l<j}}^{k}\int_{0}^{1}\left(m_j(x) - m_l(x)\right)^2(x)\ \frac{k_j f_j(x) k_l f_l(tx}{\bar{R}(x)}\ dx$$

$$\gamma_{k,2}^2 = 4\sum_{j=1}^{k}\int_{0}^{1}\left(\sum_{\substack{l=1 \\ l\neq j}}^{k}(m_j(x) - m_l(x))\frac{k_l f_l(x)}{\bar{R}(x)}\right)^2 \sigma_j^2(x)\, k_j f_j(x)dx$$

Koul and Schick (1997) ศึกษาเกี่ยวกับปัญหาการทดสอบการเท่ากันของฟังก์ชันการถดถอย 2 ฟังก์ชันได้แก่ m_1 และ m_2 เมื่อสมมติฐานของการทดสอบอยู่ในรูป $H_0: m_1 = m_2$, $H_1: m_1 > m_2$ โดยพิจารณาค่าของตัวแปรอิสระ X ใน 2 ลักษณะคือ Common Design และ Distinct Design เมื่อ Common Design มีขนาดของตัวอย่าง $n_1 = n_2 = n$ กำหนดตัวแปรอิสระ X ให้ $X_{1,j} = X_{2,j} = X_j$, $j = 1, 2, \ldots, n$, $(X_1, \ldots, X_n$ เป็นอิสระต่อกัน) ในขณะที่ Distinct Design ตัวแปรอิสระ X อยู่ในรูป $X_{1,1}, X_{1,2}, \ldots, X_{1,n}$ เป็นอิสระกับ $X_{2,1}, X_{2,2}, \ldots, X_{2,n}$

สถิติทดสอบกรณี Common Design อยู่ในรูปของ

$$T_1 = \frac{1}{\sqrt{2n}}\sum_{j=1}^{n}\nu(X_j)\left(\psi(\hat{\varepsilon}_{1,j}) - \psi(\hat{\varepsilon}_{2,j})\right)$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

สถิติทดสอบกรณี Distinct Design อยู่ในรูปของ

$$T_2 = \sqrt{\frac{n_1 n_2}{n_1 + n_2}}\left(\frac{1}{n_1}\sum_{j=1}^{n_1} \nu(X_{k,j})\psi(\hat{\varepsilon}_{1,j}) - \frac{1}{n_2}\sum_{j=1}^{n_2} \nu(X_{2,t})\psi(\hat{\varepsilon}_{2,j}) \right)$$

เมื่อ ν , ψ แทน Measurable Function และ $\nu \geq 0$ โดยที่

$$\hat{\varepsilon}_{k,j} = Y_{k,j} - \hat{m}(X_{k,j}) \qquad j = 1,2, \ldots ,n_k \quad k = 1, 2$$

$$\hat{m}(x) = \frac{\sum_{k=1}^{2}\sum_{j=1}^{n_k} K\left(\frac{x - X_{ij}}{h}\right) Y_{ij}}{\sum_{k=1}^{2}\sum_{j=1}^{n_k} K\left(\frac{x - X_{ij}}{h}\right)}$$

Koul and Schick (1997) ได้แสดงเกี่ยวกับคุณสมบัติของสถิติทดสอบ T_1 และ T_2 ไว้ดังนี้

- กรณีของ Common Design (สถิติทดสอบคือ T_1) โดย

$T_1 = \frac{1}{\sqrt{2n}}\sum_{j=1}^{n} \nu(X_j)\left(\psi(\hat{\varepsilon}_{1,j}) - \psi(\hat{\varepsilon}_{2,j})\right) + O_{P_o}(1)$ คือ Asymptotic $N(0,\tau_1)$

ซึ่งสามารถประมาณ τ_1 ได้โดย

$$\hat{\tau}_1 = \frac{1}{n}\sum_{j=1}^{2} \nu^2(X_j)\frac{1}{2n}\sum_{j=1}^{n} (\psi(\hat{\varepsilon}_{1,j}) - \psi(\hat{\varepsilon}_{2,j}))^2$$

- กรณีของ Distinct Design (สถิติทดสอบคือ T_2) โดย

$$T_2 = \sqrt{\frac{n_1 n_2}{n_1+n_2}}\left(\frac{1}{n_1}\sum_{j=1}^{n_1} \nu(X_{k,j})\psi(\hat{\varepsilon}_{1,j}) - \frac{1}{n_2}\sum_{j=1}^{n_2} \nu(X_{2,t})\psi(\hat{\varepsilon}_{2,j}) \right) + O_{P_o}(1)$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

คือ Asymptotic $N(0,\tau_2)$ ซึ่งสามารถประมาณ τ_2 ได้โดย

$$\hat{\tau}_2 = \frac{1}{n_1 + n_2}\sum_{k=1}^{2}\sum_{j=1}^{n_k}\nu^2(X_{k,j})\psi^2(\hat{\varepsilon}_{k,j}) - \left(\frac{1}{n_1+n_2}\sum_{k=1}^{2}\sum_{j=1}^{n_k}\nu(X_{2,k})\psi(\hat{\varepsilon}_{2,k})\right)^2$$

Munk and Dette (1998) ได้สร้างสถิติทดสอบเพื่อเปรียบเทียบระหว่างฟังก์ชันการถดถอย 2 ฟังก์ชัน $m_1(x)$ และ $m_2(x)$ โดย $(X_i,Y_i) = (m_1(x) + \varepsilon_i,\ m_2(x) + \eta_i)$ สถิติทดสอบกำหนดโดย Weighted L^2 – Distance

$$M^2 = M^2(f,g) = \|m_1 - m_2\|^2 = \int_0^1 (m_1(x) - m_2(x))^2 dx$$

โดยสมมติฐานของการทดสอบอยู่ในรูป

$$H_0: M^2 = 0 \text{ และ } H_1: M^2 \neq 0$$

โดย M^2 สามารถเขียนได้ในรูป

$$M^2(m_1, m_2) = \sum_{i=0}^{m}\sum_{j=0}^{n}\int_0^1 I_{\left[x_{1,i},\, x_{1,i+1}\right)}(x) I_{\left[x_{2,j},x_{2,j+1}\right)}(x)(m_1(x) - m_2(x))^2 dx \qquad (6)$$

ซึ่งสามารถประมาณ M^2 ได้โดย

$$\hat{M}^2 = \sum_{i=0}^{m}\sum_{j=0}^{n}\lambda_{ij}(X_{i+1} - Y_{j+1})(X_i - Y_j)$$

เมื่อ $X_0 = X_1,\ Y_0 = Y_1,\ X_{m+1} = X_m$ และ $Y_{n+1} = Y_n \quad m, n \geq 1$

ในที่นี้ λ_{ij} กำหนดได้โดย

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

$$\lambda_{ij} = \int_0^1 I_{\left[x_{1,i},x_{1,i+1}\right)}(x)(I_{\left[x_{2,i},x_{2,j+1}\right)}(x)dx, \quad i = 1, \ldots, m \quad j = 1, \ldots, n$$

$$= (x_{1,i+1} \cap x_{2,j+1} - x_{1,i} \cup x_{2,t})I_{\left\{x_{1,i+1} \cap x_{2,j+1} > x_{1,i} \cup x_{2,j}\right\}}$$

นอกจากนั้น Munk and Dette (1998) ได้แสดงให้เห็นว่าเมื่อ $n \to \infty$, $m \to \infty$ และ

$\frac{m}{(n+m)} = k \in (0, \frac{1}{2})$ จะได้ว่า $(n+m)^{1/2}(\hat{M}^2 - M^2) \to N(0, \zeta^2)$ เมื่อ $N(0,\zeta^2)$ แทน

การแจกแจงแบบปกติที่มีค่าเฉลี่ยเท่ากับ 0 ความแปรปรวน ζ^2 โดยความแปรปรวน ζ^2 หาได้โดย

$$\zeta^2(M^2) = \zeta^2 = k^{-1}(\sigma_\varepsilon^4 + 4\sigma_\varepsilon^2 M^2) + (1 - k)^{-1}(\sigma_\eta^4 + 4\sigma_\eta^2 M^2) + 2\sigma_\varepsilon^2\sigma_\eta^2\iota(r,s)$$

$\sigma_\varepsilon^2, \sigma_\eta^2$ คือ ความแปรปรวนของฟังก์ชันการถดถอย $m_1(x)$ และ $m_2(x)$ ตามลำดับ

กรณีที่ $\frac{n}{m} = q \in \square$ จะได้ว่า

$$\zeta^2 = (q+1)(\sigma_\varepsilon^4 + 4\sigma_\varepsilon^2 M^2) + (1 + q)^{-1}(\sigma_\eta^4 + 4\sigma_\eta^2 M^2) + 2\sigma_\varepsilon^2\sigma_\eta^2(1 + q^{-1})$$

ถ้า $n = m$ จะได้ว่า

$$\zeta^2 = 2(\sigma_\varepsilon^2 + \sigma_\eta^2)^2 + 8M^2(\sigma_\varepsilon^2 + \sigma_\eta^2)$$

นอกจากนั้น Munk and Dette (1998) ได้เปรียบเทียบอำนาจการทดสอบกับวิธีการของ Delgado พบว่าวิธีของ Delgado มีประสิทธิภาพดีกว่าในกรณีที่ความแตกต่างระหว่าง $m_1(x)$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

และ $m_2(x)$ ที่อยู่ในรูปของ Linear เช่น $m_1(x) = x + 1$, $m_2(x) = x$ ดังนั้น $m_1(x) - m_2(x) = 1$ ในขณะที่วิธีการของ Munk and Delte จะมีประสิทธิภาพดีกว่าในกรณีที่ความแตกต่างระหว่าง $m_1(x)$ และ $m_2(x)$ อยู่ในรูปของ Oscillating Function เช่น $m_1(x) - m_2(x) = \sin(2\pi x)$

Neumeyer and Dette (2003) ได้เสนอสถิติทดสอบเพื่อเปรียบเทียบการเท่ากันของ ฟังก์ชันการถดถอย 2 ฟังก์ชัน ซึ่งแนวคิดพื้นฐานมาจากการหาผลต่างระหว่าง 2 ฟังก์ชันการแจกแจงสะสมของตัวอย่างของความคลาดเคลื่อน โดยสมมติฐานของการทดสอบอยู่ในรูป

$$H_0: m_1 = m_2 \quad \text{และ} \quad H_1: m_1 \neq m_2$$

จากตัวแบบการถดถอย

$$Y_{ij} = m_i(X_{ij}) + \sigma_i(X_{ij})\varepsilon_{ij}, \quad j = 1, \ldots, n_i, \ i = 1, 2$$

ได้กำหนดให้ตัวแปรอิสระ X_{ij} $(j = 1, \ldots, n_i)$ คือค่าสังเกตที่เป็นอิสระกันโดยมีฟังก์ชันความหนาแน่นน่าจะเป็น $f_i(x)$ ในช่วง $[0,1]$ $(i = 1, 2)$ และมีข้อสมมติเบื้องต้นเกี่ยวกับการทดสอบสมมติฐานดังนี้

1. $\hat{f}_i(x) = \frac{1}{n_i h}\sum_{j=1}^{n_i} K\left(\frac{x - X_{ij}}{h}\right)$ แทนตัวประมาณฟังก์ชันความหนาแน่นน่าจะเป็นของตัวแปรอิสระ X_{ij} $j = 1, \ldots, n_i, \ i = 1, 2$ และ $\hat{f}(x) = \frac{n_1}{N}\hat{f}_1(x) + \frac{n_2}{N}\hat{f}_2(x)$ แทนตัวประมาณฟังก์ชันความหนาแน่นน่าจะเป็นจากตัวแปรอิสระทั้ง 2 กลุ่ม ได้แก่ $X_{11}, \ldots, X_{1n_1}, X_{21}, \ldots, X_{2n_2}$

2. ในการศึกษาครั้งนี้กำหนดให้พารามิเตอร์ปรับให้เรียบ h มีขนาดเท่ากัน โดยมีคุณสมบัติ

$$h \to 0, \ Nh^{4d} \to 0, \ Nh^2 \to \infty$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

3. $\hat{f}(x) \xrightarrow{P} \hat{f}(x) = k_1 f_1(x) + k_2 f_2(x)$

4. ขณะที่ $n \to \infty$ $\frac{n_i}{N} \to k_i + O(\frac{1}{n})$, $i = 1, 2$ และ $k_i \in (0,1)$

5. ตัวประมาณ Nadaraya-Watson ของฟังก์ชันการถดถอยที่ประกอบไปด้วยตัวอย่างทั้ง 2 กลุ่ม กำหนดโดย

$$\hat{m}(x) = \frac{1}{Nh}\sum_{i=1}^{2}\sum_{j=1}^{n_i} K\left(\frac{x - X_{ij}}{h}\right) Y_{ij} \frac{1}{\hat{f}(x)}$$

$$= \frac{(n_1/N)\hat{f}_1(x)\hat{m}_1(x) + (n_2/N)\hat{f}_2(x)\hat{m}_2(x)}{\hat{f}(x)}$$

เมื่อ $\hat{m}_i(x) = \frac{1}{n_i h}\sum_{j=1}^{n_i} K\left(\frac{x - X_{ij}}{h}\right) Y_{ij} \frac{1}{\hat{f}_i(x)}$

6. ภายใต้สมมติฐานหลักของการทดสอบการเท่ากันของฟังก์ชันการถดถอยจะได้ว่า $H_0: m_1(x) = m_2(x) = m(x)$ สำหรับ $i = 1, 2$ โดยได้กำหนดความคลาดเคลื่อนใน 2 ลักษณะดังต่อไปนี้

$$e_{ij} = \frac{n_{3-i}}{N}\left(Y_{ij} - \hat{m}(X_{ij})\right)\hat{f}(X_{ij})\hat{f}_{3-i}(X_{ij}) \text{ และ}$$

$$f_{ij} = \frac{N}{n_i}\left(Y_{ij} - \hat{m}(X_{ij})\right)\hat{f}_i(X_{ij})$$

7. พิจารณา Marked Empirical Process โดย

$$\hat{R}_N^{(1)}(t) = \frac{1}{N}\sum_{j=1}^{n_1} e_{1j} I\{X_{1j} \le t\} - \frac{1}{N}\sum_{j=1}^{n_2} e_{2j} I\{X_{2j} \le t\}$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

$$\hat{R}_N^{(2)}(t) = \frac{1}{N}\sum_{j=1}^{n_1} f_{1j} I\left\{X_{1j} \le t\right\} - \frac{1}{N}\sum_{j=1}^{n_2} f_{2j} I\left\{X_{2j} \le t\right\}$$

เมื่อ $t \in [0,1]$ และ I{.} แทน Indicator Function สถิติทดสอบอยู่ในลักษณะต่อไปนี้

$$\sup_{t \in [0,1)} \left|\hat{R}_N^{(i)}(t)\right|$$

- ภายใต้สมมติฐานหลักของการเท่ากันของฟังก์ชันการถดถอย จะได้ว่า

$$E\left[\hat{R}_N^{(i)}(t)\right] = O\left(\frac{1}{Nh}\right) + O(h^{2d}) = o\left(\frac{1}{\sqrt{n}}\right), \quad l = 1,2$$

- ภายใต้สมมติฐานทางเลือกของการไม่เท่ากันของฟังก์ชันการถดถอย จะได้ว่า

$$E\left[\hat{R}_N^{(1)}(t)\right] = k_1 k_2 \int_0^t \left(m_1(x) - m_2(x)\right) f(x) f_1(x) f_2(x) dx + O(h^d)$$

$$E\left[\hat{R}_N^{(2)}(t)\right] = k_1 k_2 \int_0^t \left(m_1(x) - m_2(x)\right) dx + O(h^d)$$

ภายใต้สมมติฐานหลักของการเท่ากันของฟังก์ชันการถดถอย Marked Empirical Process $\sqrt{N}\hat{R}_N^{(1)}$ จะลู่เข้าสู่ Gaussian Process ที่มีค่าเฉลี่ยเท่ากับ 0 ฟังก์ชันความแปรปรวนร่วม

$$H^{(1)}(s,t) = \int_0^{s \wedge t} (\sigma_1^2(x) k_2 f_2(x) + \sigma_2^2(x) k_1 f_1(x)) k_1 f_1(x) k_2 f_2(x) f^2(x) dx$$

ในทำนองเดียวกันภายใต้สมมติฐานหลักของการเท่ากันของฟังก์ชันการถดถอย Marked Empirical Process $\sqrt{N}\hat{R}_N^{(2)}$ จะลู่เข้าสู่ Gaussian Process ที่มีค่าเฉลี่ยเท่ากับ 0 ฟังก์ชันความแปรปรวนร่วม

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

$$H^{(2)}(s,t) = \int_0^{s\wedge t} (\sigma_1^2(x)k_2f_2(x) + \sigma_2^2(x)k_1f_1(x))\frac{1}{k_1f_1(x)k_2f_2(x)}dx$$

Rardo-Fernandez (2007) ได้ศึกษาเกี่ยวกับสถิติทดสอบเพื่อเปรียบเทียบการเท่ากันของ k ฟังก์ชันการแจกแจงของความคลาดเคลื่อน โดยสมมติฐานของการทดสอบอยู่ในรูปของ

$$H_0: F_{\varepsilon 1} = F_{\varepsilon 2} = \ldots = F_{\varepsilon k}$$
$$H_1: F_{\varepsilon i} \neq F_{\varepsilon j} \quad \text{สำหรับ } i, j \text{ บางค่า } i \neq j,\ i,j \in \{1, \ldots, k\}$$

โดยสถิติทดสอบได้แก่

$$S_{CM} = \sum_{j=1}^{k}\int \hat{U}_j^2(y)d\hat{F}_\varepsilon(y)$$

เมื่อ $\hat{U}(y) = \left(\hat{U}_1(y), \ldots, \hat{U}_k(y)\right)^t,\ -\infty < y < \infty$

$$\hat{U}_j(y) = n_j^{1/2}\left(\hat{F}_\varepsilon(y) - \hat{F}_{\varepsilon_j}(y)\right),\ j = 1,2, \ldots, k$$

โดยตัวประมาณฟังก์ชันการแจกแจงความคลาดเคลื่อน $\hat{F}_{\varepsilon j}(y)$ คือ

$$\hat{F}_{\varepsilon j}(y) = \frac{1}{n_j}\sum_{i=1}^{n_j} I\left(\frac{Y_{ij} - \hat{m}_j(X_{ij})}{\hat{\sigma}_j(X_{ij})} \leq y\right)$$

และตัวประมาณฟังก์ชันการแจกแจงความคลาดเคลื่อน $\hat{F}_\varepsilon(y)$ ภายใต้สมมติฐานหลัก คือ

$$\hat{F}_\varepsilon(y) = \frac{1}{n}\sum_{j=1}^{k}\sum_{i=1}^{n_j} I\left(\frac{Y_{ij} - \hat{m}_j(X_{ij})}{\hat{\sigma}_j(X_{ij})} \leq y\right)$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

โดยที่ $\hat{m}_j(x)$ และ $\hat{\sigma}_j^2(x)$ แทนตัวประมาณฟังก์ชันการถดถอยและฟังก์ชันของความแปรปรวน โดย

$$\hat{m}_j(x) = \sum_{i=1}^{n_j} w_{ij}(x, h)Y_{ij}$$

$$\hat{\sigma}_j^2(x) = \sum_{i=1}^{n_j} w_{ij}(x, h)Y_{ij}^2 - \hat{m}_j^2(x)$$

$w_{ij}(x,h)$ คือ Nadaraya – Watson Weight คำนวณได้จาก

$$w_{ij}(x,h) = \frac{K\left(\frac{x - X_{ij}}{h}\right)}{\sum_{i'=1}^{n_j} K\left(\frac{x - X_{ij}}{h}\right)}$$

ผลจากการศึกษาการแจกแจงของสถิติทดสอบภายใต้สมมติฐานหลักพบว่า

k - Dimensional ของ $\hat{U}(y) = \left(\hat{U}_1(y), ..., \hat{U}_k(y)\right)^t$ จะลู่เข้าสู่ k - dimensional Gaussian Process $U(y) = \left(U_1(y), ... ,U_k(y)\right)^t$ ที่มีโครงสร้างของความแปรปรวนร่วม

$$Cov(U_j(y),U_{j'}(y')) = \sum_{l=1}^{k} p_j^{1/2} p_{j'}^{1/2} \left(1-\frac{I(l = j)}{p_j}\right)\left(1-\frac{I(l = j')}{p_{j'}}\right) E(\psi_l(X_l,Y_l,y)\psi_l(X_l,Y_l,y'))$$

สำหรับ $j = 1, \ldots ,k$

$$\psi_j(u,v,y) = I\left(\frac{v - m_j(u)}{\sigma_j(u)} \le y\right) - F_\varepsilon(y) + yf_\varepsilon(y)\frac{(v- m_j(u))^2 - \sigma_j^2(u)}{2\sigma_j^2(u)} + f_\varepsilon(y)\frac{v - m_j(u)}{\sigma_j(u)}$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

อุปกรณ์และวิธีการ

อุปกรณ์

1. คอมพิวเตอร์แลบท็อป Processor Intel (R) Core(TM) 2 Duo CPU P 7450 @ 2.13 GHZ, 1.58 GHZ หน่วยความจำ 2 GB

2. โปรแกรม SAS version 9.1

วิธีการ

การวิจัยในครั้งนี้เป็นการศึกษาเกี่ยวกับปัญหาการทดสอบการเท่ากันของฟังก์ชันการถดถอย k ฟังก์ชันโดยตัวแบบการถดถอยอยู่ในรูป

$$Y_{ij} = m_j(X_{ij}) + \sigma_j(X_{ij})\varepsilon_{ij},\ i = 1, \ldots, k,\ j = 1, \ldots, n_i$$

เมื่อ $m_j(X_{ij})$ คือ ฟังก์ชันการถดถอยของ Y บน X โดยที่
$m_j(X_{ij}) = E(Y_{ij}|X_{ij})$

$\sigma_j^2(X_{ij})$ คือ ความแปรปรวน โดยที่ $\sigma_j^2(X_{ij}) = Var(Y_{ij}|X_{ij})$

ε_{ij} คือ ความคลาดเคลื่อน โดยที่ $E(\varepsilon_{ij}|X_{ij}) = 0$ และ $Var(\varepsilon_{ij}) = 1$

สมมติฐานของการทดสอบอยู่ในรูป

$H_0: m_1 = m_2 = m_3 = \ldots = m_k$ เมื่อ $k \geq 2$

$H_1: m_i \neq m_j$ สำหรับ i, j บางค่า $i \neq j,\ i, j \in \{1, \ldots, k\}$

วิธีดำเนินการวิจัยประกอบไปด้วย

1) ศึกษาสถิติทดสอบการเท่ากันของฟังก์ชันการถดถอย k ฟังก์ชัน

2) การจำลองแบบ (Simulation)

3) การประยุกต์สถิติทดสอบการเท่ากันของฟังก์ชันการถดถอย k ฟังก์ชัน (k = 3) กับข้อมูลจริง โดยมีรายละเอียดดังต่อไปนี้

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

1. ศึกษาสถิติทดสอบการเท่ากันของฟังก์ชันการถดถอย k ฟังก์ชัน ประกอบไปด้วย

1.1 ศึกษาเกี่ยวกับการเท่ากันของการแจกแจงของความคลาดเคลื่อนซึ่งนำไปสู่การทดสอบสมมติฐาน $H_0: m_1 = m_2 = m_3 = ... = m_k$ เมื่อ $k \geq 2$ และ $H_1: m_i \neq m_j$ สำหรับ i, j บางค่า $i \neq j,\ i, j \in \{1, ..., k\}$

1.2 หาผลต่างระหว่างตัวประมาณการแจกแจงของความคลาดเคลื่อนภายใต้สมมติฐานหลักและตัวประมาณการแจกแจงของความคลาดเคลื่อนของฟังก์ชันการถดถอยฟังก์ชันที่ j $\hat{F}_{\varepsilon j0}(y) - \hat{F}_{\varepsilon j}(y)$ เมื่อสมมติฐานหลักเป็นจริง

1.3 ศึกษาเกี่ยวกับการแจกแจง ของ k – Dimensional Process $\hat{U}(y) = \left(\hat{U}_1(y), \hat{U}_2(y), ..., \hat{U}_k(y)\right)$ เมื่อสมมติฐานหลักเป็นจริงโดยใช้หลักการของ Cramer – Wold Device และ Donsker Theorem

1.4 ศึกษาการแจกแจงของสถิติทดสอบภายใต้สมมติฐานหลักโดยใช้ผลลัพธ์ที่ได้จากการศึกษาในข้อ 1.3 และ Continuous Mapping Theorem

1.5 หาผลต่างระหว่างตัวประมาณการแจกแจงของความคลาดเคลื่อนภายใต้สมมติฐานหลักและตัวประมาณการแจกแจงของความคลาดเคลื่อนของฟังก์ชันการถดถอยฟังก์ชันที่ j $\hat{F}_{\varepsilon j0}(y) - \hat{F}_{\varepsilon j}(y)$ เมื่อสมมติฐานทางเลือกเป็นจริง

1.6 ศึกษาเกี่ยวกับการแจกแจงของ k – Dimensional Process $\hat{U}(y) = \left(\hat{U}_1(y), \hat{U}_2(y), ..., \hat{U}_k(y)\right)$ เมื่อสมมติฐานทางเลือกเป็นจริงโดยใช้หลักการของ Donsker Theorem

1.7 ศึกษาการแจกแจงของ สถิติทดสอบภายใต้สมมติฐานทางเลือกโดยใช้ผลลัพธ์ที่ได้จากการศึกษาในข้อ 1.6 และ Continuous Mapping Theorem

การศึกษาสถิติทดสอบการเท่ากันของฟังก์ชันการถดถอย k ฟังก์ชัน ประกอบไปด้วยข้อตกลงต่าง ๆ ดังต่อไปนี้

ข้อตกลงสำหรับการศึกษาสถิติทดสอบการเท่ากันของฟังก์ชันการถดถอย k ฟังก์ชัน

1. กำหนดให้ $\dfrac{\overline{Y}_j - \hat{m}_j(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}}$ แทนตัวประมาณความคลาดเคลื่อนของประชากรที่ j โดยที่ $E(\varepsilon_{ij}|X_{ij}) = 0$ และ $Var(\varepsilon_{ij}) = 1$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

2. กำหนดให้ $\frac{\overline{Y}_j - \hat{m}(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}}$ แทนตัวประมาณความคลาดเคลื่อนเมื่อสมมติฐานหลักเป็นจริง

3. สำหรับ $j = 1, \ldots, k$, $\hat{m}_j(x) = \sum_{i=1}^{n_j} W_{ij}(x,h)Y_{ij}$

$$\hat{\sigma}_j^2(x) = \sum_{i=1}^{n_j} W_{ij}(x,h)Y_{ij}^2 - \hat{m}_j^2(x)$$

เมื่อ $w_{ij}(x, h) = \frac{K\left(\frac{x - X_{ij}}{h}\right)}{\sum_{i'=1}^{n_j} K\left(\frac{x - X_{ij}}{h}\right)}$ คือ Nadaraya – Watson Weight

K คือ Kernel Function

h คือ พารามิเตอร์ปรับให้เรียบ

4. กำหนดให้ $\hat{m}(x) = \sum_{j=1}^{k}\sum_{i=1}^{n_j} W_{ij}(x,h)Y_{ij}$ แทนตัวประมาณของ Common Regression Function $m(x) = m_1(x) = m_2(x) = \ldots = m_k(x)$ เมื่อสมมติฐานหลักเป็นจริง

5. เพื่อให้ง่ายต่อการศึกษาในที่นี้ใช้พารามิเตอร์ปรับให้เรียบ h เดียวกันในการประมาณ $\hat{m}(x)$, $\hat{m}_j(x)$ และ $\hat{\sigma}_j(x)$

6. สำหรับ $j = 1, \ldots, k$ ตัวประมาณการแจกแจงของความคลาดเคลื่อนคือ

$$\hat{F}_{\varepsilon j}(y) = \frac{1}{n_j}\sum_{i=1}^{n_j} I\left(\frac{\overline{Y}_j - \hat{m}_j(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}} \le y\right)$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

และตัวประมาณการแจกแจงความคลาดเคลื่อนเมื่อสมมติฐานหลักเป็นจริงคือ

$$\hat{F}_{\varepsilon j0}(y) = \frac{1}{n_j}\sum_{i=1}^{n_j} I\left(\frac{\overline{Y}_j - \hat{m}(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}} \le y\right)$$

7. ในที่นี้ได้เปรียบเทียบตัวประมาณการแจกแจงของความคลาดเคลื่อนในแต่ละประชากร โดยใช้ k – Dimensional Process ดังนี้

$$\hat{U}(y) = \left(\hat{U}_1(y), \hat{U}_2(y), \ldots, \hat{U}_k(y)\right), \quad -\infty < y < \infty$$

$$\hat{U}_j(y) = n_j^{1/2}\left(\hat{F}_{\varepsilon j0}(y) - \hat{F}_{\varepsilon j}(y)\right), \quad j = 1, \ldots, k$$

8. นอกจากตัวประมาณการแจกแจงของความคลาดเคลื่อนในข้อ 6 กำหนดตัวประมาณการแจกแจงของความคลาดเคลื่อนในอีกลักษณะหนึ่งดังนี้

$$\hat{F}_{\varepsilon}(y) = \frac{1}{n}\sum_{j=1}^{k}\sum_{i=1}^{n_j} I\left(\frac{\overline{Y}_j - \hat{m}_j(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}} \le y\right) \quad \text{และ}$$

$$\hat{F}_{\varepsilon 0}(y) = \frac{1}{n}\sum_{j=1}^{k}\sum_{i=1}^{n_j} I\left(\frac{\overline{Y}_j - \hat{m}(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}} \le y\right)$$

ในทำนองเดียวกัน Joint Process คือ $\hat{U}(y) = n^{1/2}\left(\hat{F}_{\varepsilon 0}(y) - \hat{F}_{\varepsilon}(y)\right), \quad -\infty < y < \infty$

9. กำหนดให้ $\frac{n_j}{n} \to p_j > 0$ จะได้ว่า $f_{mix}(x) = \sum_{j=1}^{k} p_j f_j(x)$ คือ Mixture ของฟังก์ชันความหนาแน่นน่าจะเป็นของ X_j เมื่อ $f_j(x)$ คือ ฟังก์ชันความหนาแน่นน่าจะเป็นของ X_j

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

10. ภายใต้สมมติฐานทางเลือก $\hat{m}(x)$ เป็นตัวประมาณของ

$$m_n(x) = m_0(x) + n^{-1/2}S(x)$$

เมื่อ $S(x) = \sum_{j=1}^{k} p_j \frac{f_j(x)}{f_{mix}(x)} s_j(x)$ และ $\hat{m}_{jn}(x)$ เป็นตัวประมาณของ

$$m_{jn}(x) = m_0(x) + n^{-1/2}s(x)$$

11. สำหรับ $j = 1, \ldots, k$

11.1 X_j เป็นตัวแปรสุ่มต่อเนื่อง $x \in \square_x$ และมีฟังก์ชันความหนาแน่นน่าจะเป็น $f_j(x)$

11.2 $f_j(x)$, $m_j(x)$ และ $\sigma_j(x)$ ต่อเนื่องและสามารถหาอนุพันธ์อันดับที่ 2 ได้

11.3 $\inf_{x\in\square_x} f_j(x) > 0$ และ $\inf_{x\in\square_x} \sigma_j(x) > 0$

12. สำหรับ $j = 1, \ldots, k$

12.1 $\frac{n_j}{n} \to p_j > 0$ ขณะ $n \to \infty$

12.2 $nh^4 \to 0$ และ $n_j h_n^{3+2\delta}(\log h_n^{-1}) \to \infty$ สำหรับ $\delta > 0$ ขณะที่ $h \to 0$

13. K เป็น Kernel ฟังก์ชันที่มีสมบัติ

13.1 $\int u^2 K(u)du < \infty$

13.2 $\int K(u)du = 1$

13.3 $\int uK(u)du = 0$

14. สำหรับ $j = 1, \ldots, k$

14.1 $F_j(y|x)$ เป็นฟังก์ชันที่ต่อเนื่องใน (x, y) และสามารถหาอนุพันธ์เทียบกับ y ได้

14.2 $F_j(y|x)$ เป็นฟังก์ชันที่ต่อเนื่องใน (x, y) และสามารถหาอนุพันธ์เทียบกับ x ได้

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

การศึกษาสถิติทดสอบการเท่ากันของฟังก์ชันการถดถอย k ฟังก์ชัน นำไปสู่ทฤษฎีบทต่าง ๆ ที่ต้องศึกษา ประกอบไปด้วยทฤษฎีบทที่ 1 เป็นทฤษฎีบทที่แสดงให้เห็นว่าการเท่ากันของการแจกแจงของความคลาดเคลื่อน คือการเท่ากันของฟังก์ชันการถดถอยซึ่งนำไปสู่การทดสอบสมมติฐาน $H_0: m_1 = m_2 = m_3 = ... = m_k$ และ $H_1: m_l \neq m_j$ สำหรับ i, j บางค่า $i \neq j,\ i,j \in \{1, ... , k\}$, ทฤษฎีบทที่ 2 เป็นการพิสูจน์เพื่อหาผลต่างระหว่างตัวประมาณการแจกแจงของความคลาดเคลื่อนภายใต้สมมติฐานหลักและตัวประมาณการแจกแจงของความคลาดเคลื่อนของฟังก์ชันการถดถอยฟังก์ชันที่ j $\hat{F}_{\varepsilon j0}(y) - \hat{F}_{\varepsilon j}(y)$ เมื่อสมมติฐานหลักเป็นจริง ทฤษฎีบทที่ 3 เป็นการพิสูจน์เพื่อหาการแจกแจงของ k – Dimensional Process ของ $\hat{U}(y) = \left(\hat{U}_1(y) , \hat{U}_2(y) , ... , \hat{U}_k(y)\right)$ เมื่อสมมติฐานหลักเป็นจริง โดยใช้หลักการของ Cramer – Wold Device และ Donsker Theorem ทฤษฎีบทที่ 4 เป็นการพิสูจน์เพื่อหาการแจกแจงของสถิติทดสอบภายใต้สมมุติฐานหลักโดยใช้ผลลัพธ์ที่ได้จากการพิสูจน์ทฤษฎีบทที่ 3 และ Continuous Mapping Theorem ทฤษฎีบทที่ 5 เป็นการพิสูจน์เพื่อหาผลต่างระหว่างตัวประมาณการแจกแจงของความคลาดเคลื่อนภายใต้สมมติฐานหลัก และตัวประมาณการแจกแจงของความคลาดเคลื่อนของฟังก์ชันการถดถอยฟังก์ชันที่ j $\hat{F}_{\varepsilon j0}(y) - \hat{F}_{\varepsilon j}(y)$ เมื่อสมมติฐานทางเลือกเป็นจริง ทฤษฎีบทที่ 6 เป็นการพิสูจน์เพื่อหาการแจกแจงของ k – Dimensional Process ของ $\hat{U}(y) = \left(\hat{U}_1(y) , \hat{U}_2(y) , ... , \hat{U}_k(y)\right)$ เมื่อสมมติฐานทางเลือกเป็นจริง โดยใช้หลักการของ Cramer – Wold Device และ Donsker Theorem และทฤษฎีบทที่ 7 เป็นการพิสูจน์เพื่อหาการแจกแจงของสถิติทดสอบภายใต้สมมติฐานทางเลือกโดยใช้ผลลัพธ์ที่ได้จากการพิสูจน์ทฤษฎีบทที่ 6 และ Continuous Mapping Theorem

ทฤษฎีบท 1.1 $F_{\varepsilon j0}(y) = F_{\varepsilon j}(y)$ เมื่อ $-\infty < y < \infty$ สำหรับทุก $j = 1 , 2 , ... , k$ ก็ต่อเมื่อ $m(x) = m_1(x) = ... = m_k(x)$ สำหรับทุก $x \in \square_x$

พิสูจน์ กำหนดให้ $F_{\varepsilon j0}(y) = F_{\varepsilon j}(y)$ ต้องแสดงว่า $m(x) = m_1(x) = ... = m_k(x)$ การที่การแจกแจงของความคลาดเคลื่อนเหมือนกัน แสดงว่าโมเมนต์ที่ 1 และโมเมนต์ที่ 2 ของแต่ละการแจกแจงต้องเท่ากัน

พิจารณาโมเมนต์ที่ 1 $E\left(\frac{\overline{Y}_j - m(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}}\right) = E\left(\frac{\overline{Y}_j - m_j(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}}\right) = 0$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

จะได้ว่า $m(x) = m_j(x)$

พิจารณาโมเมนต์ที่ 2

$$\text{Var}\left(\frac{\overline{Y}_j - m(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}}\right) = E\left[\left(\frac{\overline{Y}_j - m(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}}\right)^2\right] - \left[E\left(\frac{\overline{Y}_j - m(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}}\right)\right]^2$$

$$= E\left[\left(\frac{\overline{Y}_j - m(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}}\right)^2\right]$$

$$\text{Var}\left(\frac{\overline{Y}_j - m_j(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}}\right) = E\left[\left(\frac{\overline{Y}_j - m_j(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}}\right)^2\right] - \left[E\left(\frac{\overline{Y}_j - m_j(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}}\right)\right]^2$$

$$= E\left[\left(\frac{\overline{Y}_j - m_j(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}}\right)^2\right]$$

เนื่องจากโมเมนต์ที่ 2 ของแต่ละการแจกแจงเท่ากัน ดังนั้น

$$\text{Var}\left[\frac{\overline{Y}_j - m(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}}\right] = \text{Var}\left[\frac{\overline{Y}_j - m_j(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}}\right]$$

จะได้ว่า $m(x) = m_j(x)$

ในทางกลับกันกำหนดให้ $m(x) = m_1(x) = \ldots = m_k(x)$ ต้องแสดงว่า $F_{\varepsilon j0}(y) = F_{\varepsilon j}(y)$

พิจารณาโมเมนต์ที่ 1 เนื่องจาก $m(x) = m_j(x)$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

จะได้ว่า $$\left[\frac{\overline{Y}_j - m(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}}\right] = \left[\frac{\overline{Y}_j - m_j(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}}\right]$$

$$E\left[\frac{\overline{Y}_j - m(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}}\right] = E\left[\frac{\overline{Y}_j - m_j(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}}\right] = 0$$

ดังนั้นโมเมนต์ที่ 1 ของแต่ละการแจกแจงเท่ากัน

พิจารณาโมเมนต์ที่ 2 เนื่องจาก $m(x) = m_j(x)$

จะได้ว่า $$\left[\frac{\overline{Y}_j - m(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}}\right] = \left[\frac{\overline{Y}_j - m_j(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}}\right]$$

$$E\left[\frac{\overline{Y}_j - m(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}}\right]^2 = E\left[\frac{\overline{Y}_j - m_j(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}}\right]^2$$

ดังนั้นโมเมนต์ที่ 2 ของแต่ละการแจกแจงเท่ากัน

เนื่องจากโมเมนต์ที่ 1 และโมเมนต์ที่ 2 ของความคลาดเคลื่อน $\left[\frac{\overline{Y}_j - m(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}}\right]$ เท่ากับ $\left[\frac{\overline{Y}_j - m_j(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}}\right]$ จึงได้ว่า $F_{\varepsilon j0}(y)$ และ $F_{\varepsilon j}(y)$ มีการแจกแจงเหมือนกัน กล่าวคือ

$F_{\varepsilon j0}(y) = F_{\varepsilon j}(y)$ เพราะฉะนั้น $F_{\varepsilon j0}(y) = F_{\varepsilon j}(y)$ สำหรับทุก $j = 1, \ldots, k$ ก็ต่อเมื่อ

$m(x) = m_1(x) = \ldots = m_k(x)$ สำหรับทุก $x \in \square_x$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

ทฤษฎีบท 1.2 $F_{\varepsilon 0}(y) = F_{\varepsilon}(y)\ ,\ -\infty < y < \infty$ ก็ต่อเมื่อ $m(x) = m_1(x) = ... = m_k(x)$ สำหรับทุก $x \in \square_x$

พิสูจน์ กำหนดให้ $F_{\varepsilon 0}(y) = F_{\varepsilon}(y)$ ต้องแสดงว่า $m(x) = m_1(x) = ... = m_k(x)$ การที่การแจกแจงของความคลาดเคลื่อนเหมือนกัน แสดงว่าโมเมนต์ที่ 1 และโมเมนต์ที่ 2 ของแต่ละการแจกแจงต้องเท่ากัน

พิจารณาโมเมนต์ที่ 1 $$\sum_{j=1}^{k} p_j E\left(\frac{\overline{Y}_j - m(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}}\right) = \sum_{j=1}^{k} p_j E\left(\frac{\overline{Y}_j - m_j(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}}\right) = 0$$

จะได้ว่า $m(x) = m_j(x)$

พิจารณาโมเมนต์ที่ 2 เนื่องจาก

$$\sum_{j=1}^{k} p_j^2 E\left(\frac{\overline{Y}_j - m(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}}\right)^2 = \mathrm{Var}\sum_{j=1}^{k} p_j\left(\frac{\overline{Y}_j - m(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}}\right) + \sum_{j=1}^{k} p_j^2\left(E\left(\frac{\overline{Y}_j - m(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}}\right)\right)^2$$

$$\sum_{j=1}^{k} p_j^2\left(E\left(\frac{\overline{Y}_j - m(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}}\right)\right)^2 = 0$$

ดังนั้น $$\mathrm{Var}\sum_{j=1}^{k} p_j\left(\frac{\overline{Y}_j - m(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}}\right) = \sum_{j=1}^{k} p_j^2 E\left(\frac{\overline{Y}_j - m(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}}\right)^2$$

$$\sum_{j=1}^{k} p_j^2 E\left(\frac{\overline{Y}_j - m_j(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}}\right)^2 = \mathrm{Var}\sum_{j=1}^{k} p_j\left(\frac{\overline{Y}_j - m_j(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}}\right) + \sum_{j=1}^{k} p_j^2\left(E\left(\frac{\overline{Y}_j - m_j(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}}\right)\right)^2$$

$$\sum_{j=1}^{k} p_j^2\left(E\left(\frac{\overline{Y}_j - m_j(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}}\right)\right)^2 = 0$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

ดังนั้น $\text{Var}\sum_{j=1}^{k} p_j\left(\frac{\overline{Y}_j - m_j(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}}\right) = \sum_{j=1}^{k} p_j^2 E\left(\frac{\overline{Y}_j - m_j(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}}\right)^2$

เนื่องจากโมเมนต์ที่ 2 เท่ากันจะได้ว่า

$$\text{Var}\sum_{j=1}^{k} p_j\left(\frac{\overline{Y}_j - m(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}}\right) = \text{Var}\sum_{j=1}^{k} p_j\left(\frac{\overline{Y}_j - m_j(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}}\right)$$

จะได้ว่า $m(x) = m_j(x)$

ในทางกลับกันกำหนดให้ $m(x) = m_1(x) = \ldots = m_k(x)$ ต้องแสดงว่า $F_{\varepsilon 0}(y) = F_{\varepsilon}(y)$

พิจารณาโมเมนต์ที่ 1 เนื่องจาก $m(x) = m_j(x)$

$$\sum_{j=1}^{k} p_j\left(\frac{\overline{Y}_j - m(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}}\right) = \sum_{j=1}^{k} p_j\left(\frac{\overline{Y}_j - m_j(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}}\right)$$

$$\sum_{j=1}^{k} p_j E\left(\frac{\overline{Y}_j - m(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}}\right) = \sum_{j=1}^{k} p_j E\left(\frac{\overline{Y}_j - m_j(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}}\right) = 0$$

ดังนั้น โมเมนต์ที่ 1 ของแต่ละการแจกแจงเท่ากัน

พิจารณาโมเมนต์ที่ 2 เนื่องจาก $m(x) = m_j(x)$

$$\sum_{j=1}^{k} p_j\left(\frac{\overline{Y}_j - m(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}}\right) = \sum_{j=1}^{k} p_j\left(\frac{\overline{Y}_j - m_j(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}}\right)$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

$$\sum_{j=1}^{k} p_j^2 E\left(\frac{\overline{Y}_j - m(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}}\right)^2 = \sum_{j=1}^{k} p_j^2 E\left(\frac{\overline{Y}_j - m_j(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}}\right)^2$$

ดังนั้นโมเมนต์ที่ 2 ของแต่ละการแจกแจงเท่ากัน

เนื่องจากโมเมนต์ที่ 1 และโมเมนต์ที่ 2 ของค่าเฉลี่ยของความคลาดเคลื่อน

$\sum_{j=1}^{k} p_j \left[\frac{\overline{Y}_j - m(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}}\right]$ เท่ากับ $\sum_{j=1}^{k} p_j \left[\frac{\overline{Y}_j - m_j(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}}\right]$ จึงได้ว่า $F_{\varepsilon 0}(y)$ และ $F_{\varepsilon}(y)$ มี

การแจกแจงเหมือนกัน กล่าวคือ $F_{\varepsilon 0}(y) = F_{\varepsilon}(y)$ เพราะฉะนั้น $F_{\varepsilon 0}(y) = F_{\varepsilon}(y)$,

$-\infty < y < \infty$ ก็ต่อเมื่อ $m(x) = m_1(x) = ... = m_k(x)$ สำหรับทุก $x \in \square_x$

ทฤษฎีบทที่ 2 ประกอบไปด้วยบทแทรกที่ 1 – 3 โดยมีขั้นตอนและวิธีการพิสูจน์ดังต่อไปนี้

บทแทรก 1 ภายใต้ข้อตกลงที่ 1 – 13 สำหรับ $j = 1, ... ,k$ ใด ๆ จะได้ว่า

$$\int \frac{\hat{m}_j(x) - m_j(x)}{\sigma_j(x)/\sqrt{n_j}} f_j(x)dx = \frac{1}{n_j}\sum_{i=1}^{n_j} \frac{Y_{ij} - m_j(X_j)}{\sigma_j(X_{ij})/\sqrt{n_j}} + O_p(n_j h)^{-1}$$

พิสูจน์ Hardel (1993) ได้กำหนดให้ตัวประมาณของฟังก์ชันการถดถอยเท่ากับ

$$\hat{m}_j(x) = \sum_{i=1}^{n_j} W_{ij}(x,h) Y_{ij}$$

แต่ $$W_{ij}(x,h) = \frac{K(x - X_{ij})}{\hat{f}_j(x)}$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

เมื่อ $K(u) = (nh)^{-1}K(u/h)$

ตัวประมาณของฟังก์ชันความหนาแน่นน่าจะเป็นของ $f_j(x)$ คือ

$$\hat{f}_j(x) = \frac{1}{n_j h}\left[\sum_{i=1}^{n_j} K\left(\frac{x - X_{ij}}{h}\right)\right]$$

ดังนั้นตัวประมาณของ $m_j(x)$ คือ

$$\hat{m}_j(x) = \frac{1}{n_j h \hat{f}_j(x)}\sum_{i=1}^{n_j} K\left(\frac{x - X_{ij}}{h}\right)Y_{ij}$$

ดังนั้น $$\hat{m}_j(x) - m_j(x) = \frac{1}{n_j h \hat{f}_j(x)}\sum_{i=1}^{n_j} K\left(\frac{x - X_{ij}}{h}\right)\left(Y_{ij} - m_j(x)\right)$$

Wand and Jones (1995) ได้แสดงให้เห็นว่า $\hat{f}_j(x) - f_j(x) = O_p(n_j h)^{-1/2}$

$\hat{f}_j(x)\,(f_j(x))^{-1} - 1 = O_p(n_j h)^{-1/2}$ และ $\hat{m}_j(x) - m_j(x) = O_p(n_j h)^{-1/2}$

แทน $\hat{f}_j(x)$ ด้วย $f_j(x)$ จะได้ว่า

$$\hat{m}_j(x) - m_j(x) = \frac{1}{n_j h f_j(x)}\sum_{i=1}^{n_j} K\left(\frac{x - X_{ij}}{h}\right)\left(Y_{ij} - m_j(x)\right) + O_p(n_j h)^{-1}$$ ดังนั้น

$$\int \frac{\hat{m}_j(x) - m_j(x)}{\sigma_j(x)/\sqrt{n_j}} f_j(x)dx = \frac{1}{n_j h}\sum_{i=1}^{n_j}\int \frac{K\left(\frac{x - X_{ij}}{h}\right)(Y_{ij} - m_j(x))}{\sigma_j(x)/\sqrt{n_j}}\frac{f_j(x)}{f_j(x)}dx + O_p(n_j h)^{-1}$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

$$= \frac{1}{n_j h}\sum_{i=1}^{n_j}\int \frac{K\left(\frac{x - X_{ij}}{h}\right)(Y_{ij} - m_j(x))}{\sigma_j(x)/\sqrt{n_j}}dx + O_p(n_j h)^{-1}$$

$\int \frac{K\left(\frac{x - X_{ij}}{h}\right)(Y_{ij} - m_j(x))}{\sigma_j(x)/\sqrt{n_j}}dx$ ทำได้โดย

ขั้นที่ 1 ใช้เทคนิคการอินทิเกรตเปลี่ยนตัวแปร

ให้ $S_1(X_{ij}) = \left(\frac{Y_{ij} - m_j(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}}\right)$

$u = \left(\frac{x - X_{ij}}{h}\right)$

$hu = x - X_{ij}$

$x = hu + X_{ij}$

$h(du) = dx$

ดังนั้น $\int \frac{K\left(\frac{x - X_{ij}}{h}\right)(Y_{ij} - m_j(x))}{\sigma_j(x)/\sqrt{n_j}}dx = h\int K(u)S_1(hu + X_{ij})du$

ขั้นที่ 2 กระจายอนุกรมเทเลอร์อันดับ 2 ของ S_1 รอบ X_{ij} ดังนั้น

$$h\int K(u)S_1(hu + X_{ij})du = hS_1(X_{ij})\int K(u)du + h^2S_1'(X_{ij})\int uK(u)du + O_p(h)^3$$

จากข้อตกลง 13.2 และ 13.3

$$h\int K(u)S_1(hu + X_{ij})du = hS_1(X_{ij})(1) + h^2S_1'(X_{ij})(0) + O_p(h^3)$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

เมื่อ $$S_1(X_{ij}) = \left(\frac{Y_{ij} - m_j(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}} \right)$$

ดังนั้น $$h\int K(u)S_1(hu + X_{ij})du = h\frac{Y_{ij} - m_j(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}}$$

$$\int \frac{\hat{m}_j(x) - m_j(x)}{\sigma_j(x)/\sqrt{n_j}} f_j(x)dx = \frac{h}{n_j h}\sum_{i=1}^{n_j} \frac{Y_{ij} - m_j(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}} + O_p(n_j h)^{-1}$$

$$\int \frac{\hat{m}_j(x) - m_j(x)}{\sigma_j(x)/\sqrt{n_j}} f_j(x)dx = \frac{1}{n_j}\sum_{i=1}^{n_j} \frac{Y_{ij} - m_j(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}} + O_p(n_j h)^{-1}$$

ดังนั้น $$\int \frac{\hat{m}_j(x) - m_j(x)}{\sigma_j(x)/\sqrt{n_j}} f_j(x)dx = \frac{1}{n_j}\sum_{i=1}^{n_j} \frac{Y_{ij} - m_j(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}} + O_p(n_j h)^{-1}$$

บทแทรก 2 ภายใต้ข้อตกลง 1 – 13 สำหรับ $j = 1, \ldots, k$ ใด ๆ จะได้ว่า

$$\int \frac{\hat{m}(x) - m(x)}{\sigma_j(x)/\sqrt{n_j}} f_j(x)dx = \frac{1}{n}\sum_{r=1}^{k}\sum_{i=1}^{n_r} \frac{Y_{ir} - m(X_{ir})}{\sigma_j(X_{ir})/\sqrt{n_j}} \frac{f_j(X_{ir})}{f_{mix}(X_{ir})} + O_p(nh)^{-1}$$

จาก ฟังก์ชันความหนาแน่นน่าจะเป็นของ $f_j(x)$ คือ

$$\hat{f}(x) = \frac{1}{n_r h}\left[\sum_{i=1}^{n_j} K\left(\frac{x - X_{ir}}{h}\right)\right]$$

ดังนั้นจะได้ว่าฟังก์ชันความหนาแน่นน่าจะเป็นของ $f_{mix}(x)$ คือ

$$\hat{f}_{mix}(x) = \frac{1}{nh}\left[\sum_{r=1}^{k}\sum_{i=1}^{n_r} K\left(\frac{x - X_{ir}}{h}\right)\right]$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

ดังนั้นตัวประมาณของ m(x) คือ

$$\hat{m}(x) = \frac{1}{nh\hat{f}_{mix}(x)} \sum_{r=1}^{k} \sum_{i=1}^{n_r} K\left(\frac{x - X_{ir}}{h}\right) Y_{ir}$$

ดังนั้น $$\hat{m}(x) - m(x) = \frac{1}{nh\hat{f}_{mix}(x)} \sum_{r=1}^{k} \sum_{i=1}^{n_r} K\left(\frac{x - X_{ir}}{h}\right)\left(Y_{ir} - m(x)\right)$$

Wand and Jones (1995) ได้แสดงให้เห็นว่า $\hat{f}_{mix}(x) - f_{mix}(x) = O_p(nh)^{-1/2}$ และ

$\hat{f}_{mix}(x)\,(f_{mix}(x))^{-1} - 1 = O_p(nh)^{-1/2}$ ดังนั้น

$$\hat{m}(x) - m(x) = \frac{1}{nhf_{mix}(x)} \sum_{r=1}^{k} \sum_{i=1}^{n_r} K\left(\frac{x - X_{ir}}{h}\right)\left(Y_{ir} - m(x)\right) + O_p(nh)^{-1}$$ ดังนั้น

$$\int \frac{\hat{m}(x) - m(x)}{\sigma_j(x)/\sqrt{n_j}} f_j(x)dx = \frac{1}{nh} \sum_{j=1}^{k} \sum_{i=1}^{n_j} \int \frac{K\left(\frac{x - X_{ir}}{h}\right)(Y_{ir} - m(x))}{\sigma_j(x)/\sqrt{n_j}} \frac{f_j(X_{ir})}{f_{mix}(X_{ir})} dx$$

$$+ O_p(nh)^{-1}$$

$$\int \frac{K\left(\frac{x - X_{ir}}{h}\right)(Y_{ir} - m(x))}{\sigma_j(x)/\sqrt{n_j}} \frac{f_j(X_{ir})}{f_{mix}(X_{ir})} dx$$ ทำได้โดย

ขั้นที่ 1 ใช้เทคนิคการอินทิเกรตเปลี่ยนตัวแปร

ให้ $S_1(x) = (Y_{ir} - m(x))\sqrt{n_j} f_j(x)(f_{mix}(x)\sigma_j(x))^{-1}$

$$u = \left(\frac{x - X_{ij}}{h}\right)$$

$$hu = x - X_{ir}$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

$x = hu + X_{ir}$

$h(du) = dx$

ดังนั้น $\int \frac{K\left(\frac{x - X_{ir}}{h}\right)(Y_{ir} - m(x))}{\sigma_j(x)/\sqrt{n_j}} \frac{f_j(X_{ir})}{f_{mix}(X_{ir})} dx = h\int K(u)S_1(hu + X_{ir})du$

ขั้นที่ 2 กระจายอนุกรมเทเลอร์อันดับ 2 ของ S_1 รอบ X_{ij} ดังนั้น

$$h\int K(u)S_1(hu + X_{ir})du = hS_1(X_{ir})\int K(u)du + h^2S_1'(X_{ir})\int uK(u)du + O_p(h)^3$$

เนื่องจาก $\int K(u)du = 1$ และ $\int uK(u)du = 0$

$$h\int K(u)S_1(hu + X_{ir})du = hS_1(X_{ir})(1) + h^2S_1'(X_{ir})(0) + O_p(h^3)$$

เมื่อ $$S_1(X_{ir}) = \left(\frac{Y_{ir} - m(X_{ir})}{\sigma_j(X_{ir})\sqrt{n_j}}\right)\left[\frac{f_j(X_{ir})}{f_{mix}(X_{ir})}\right]$$

ดังนั้น $$h\int K(u)S_1(hu + X_{ir})du = h\frac{Y_{ir} - m(X_{ir})}{\sigma_j(X_{ir})/\sqrt{n_j}}\left[\frac{f_j(X_{ir})}{f_{mix}(X_{ir})}\right]$$

$$\int \frac{\hat{m}(x) - m(x)}{\sigma_j(x)/\sqrt{n_j}} f_j(x)dx = \frac{h}{nh}\sum_{r=1}^{k}\sum_{i=1}^{n_r}\frac{Y_{ir} - m(X_{ir})}{\sigma_j(X_{ir})/\sqrt{n_j}}\left[\frac{f_j(X_{ir})}{f_{mix}(X_{ir})}\right] + O_p(nh)^{-1}$$

$$\int \frac{\hat{m}(x) - m(x)}{\sigma_j(x)/\sqrt{n_j}} f_j(x)dx = \frac{1}{n}\sum_{r=1}^{k}\sum_{i=1}^{n_r}\frac{Y_{ir} - m(X_{ir})}{\sigma_j(X_{ir})/\sqrt{n_j}}\left[\frac{f_j(X_{ir})}{f_{mix}(X_{ir})}\right] + O_p(nh)^{-1}$$

ดังนั้น $$\int \frac{\hat{m}(x) - m(x)}{\sigma_j(x)/\sqrt{n_j}} f_j(x)dx = \frac{1}{n}\sum_{r=1}^{k}\sum_{i=1}^{n_r}\frac{Y_{ir} - m(X_{ir})}{\sigma_j(X_{ir})/\sqrt{n_j}}\left[\frac{f_j(X_{ir})}{f_{mix}(X_{ir})}\right] + O_p(nh)^{-1}$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

บทแทรก 3 ภายใต้ข้อตกลง 1 – 13 สำหรับ $j = 1, \ldots, k$ ใด ๆ

$$\int \frac{\hat{\sigma}_j(x) - \sigma_j(x)}{\sigma_j(x)/\sqrt{n_j}} f_j(x)dx = \frac{1}{n_j}\sum_{i=1}^{n_j} \frac{(Y_{ij} - m_j(X_{ij}))^2 - \sigma_j^2(X_{ij})}{2\sigma_j^2(X_{ij})/\sqrt{n_j}} + O_p(n_j h)^{-1}$$

พิสูจน์ จัดรูป $\hat{\sigma}_j(x) - \sigma_j(x)$ โดย

$$\hat{\sigma}_j(x) - \sigma_j(x) = \frac{\hat{\sigma}_j^2(x) - \sigma_j^2(x)}{2\sigma_j(x)} - \frac{\left(\hat{\sigma}_j(x) - \sigma_j(x)\right)^2}{2\sigma_j(x)} \tag{7}$$

จาก Proposition 3 ของ Akritas and Van Keilegom (2001) ได้แสดงว่า เมื่อ $\inf_{x\in R_x} \sigma_j(x) > 0$

เทอมที่ 2 ของสมการ (7) เท่ากับ $O_p((n_j h)^{-1})$ ดังนั้นจะได้ว่า

$$\hat{\sigma}_j(x) - \sigma_j(x) = \frac{\hat{\sigma}_j^2(x) - \sigma_j^2(x)}{2\sigma_j(x)} + O_p((n_j h)^{-1}) \tag{8}$$

จาก $\hat{\sigma}_j^2(x) = \sum_{i=1}^{n_j} W_{ij}(x,h)Y_{ij}^2 - \hat{m}_j^2(x)$ จะได้ว่า

$$\hat{\sigma}_j^2(x) = \sum_{i=1}^{n_j} W_{ij}(x,h)\left(Y_{ij}^2 - \hat{m}_j^2(x)\right) \quad (\text{เนื่องจาก } \sum_{i=1}^{n_j} W_{ij}(x,h)\,\hat{m}_j^2(x) = \hat{m}_j^2(x)) \tag{9}$$

ดังนั้น

$$\int \frac{\hat{\sigma}_j(x) - \sigma_j(x)}{\sigma_j(x)/\sqrt{n_j}} f_j(x)dx = \frac{\frac{1}{n_j h}\sum_{i=1}^{n_j}\int K\frac{(x - X_{ij})}{h}\left[(Y_{ij} - m_j(x))^2 - \sigma_j^2(x)\right]}{2\sigma_j^2(x)/\sqrt{n_j}} \frac{f_j(x)}{f_j(x)}dx$$

$$+ O_p(n_j h)^{-1}$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

$$= \frac{\frac{1}{n_j h}\sum_{i=1}^{n_j}\int K\frac{(x - X_{ij})}{h}\left[(Y_{ij} - m_j(x))^2 - \sigma_j^2(x)\right]}{2\sigma_j^2(x)/\sqrt{n_j}}dx + O_p(n_j h)^{-1}$$

ต้องการ $\int K\frac{(x - X_{ij})}{h}\frac{\left[(Y_{ij} - m_j(x))^2 - \sigma_j^2(x)\right]}{2\sigma_j^2(x)/\sqrt{n_j}}dx$

ขั้นที่ 1 ใช้เทคนิคการอินทิเกรตเปลี่ยนตัวแปร

ให้ $S_1(x) = \left[\frac{(Y_{ij} - m_j(x))^2 - \sigma_j^2(x)}{2\sigma_j^2(x)/\sqrt{n_j}}\right]$

$u = \left(\frac{x - X_{ij}}{h}\right)$

$hu = x - X_{ij}$

$x = hu + X_{ij}$

$h(du) = dx$

ดังนั้น $\int K\frac{(x - X_{ij})}{h}\left[\frac{(Y_{ij} - m_j(x))^2 - \sigma_j^2(x)}{2\sigma_j^2(x)/\sqrt{n_j}}\right]dx = h\int K(u)\, S(hu + X_{ij})du$

ขั้นที่ 2 กระจายอนุกรมเทเลอร์อันดับ 2 ของ S_1 รอบ X_{ij} ดังนั้น

$$h\int K(u)S_1(hu + X_{ij})du = hS_1(X_{ij})\int K(u)du + h^2S_1'(X_{ij})\int uK(u)du + O_p(h)^3$$

จากข้อตกลง 13.2 และ 13.3

$$h\int K(u)S_1(hu + X_{ij})du = hS_1(X_{ij})(1) + h^2S_1'(X_{ij})(0) + O_p(h^3)$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

เมื่อ $$S_1(x) = \left[\frac{(Y_{ij} - m_j(x))^2 - \sigma_j^2(x)}{2\sigma_j^2(x)/\sqrt{n_j}} \right]$$

ดังนั้น $$\int \frac{\hat{\sigma}_j(x) - \sigma_j(x)}{\sigma_j(x)/\sqrt{n_j}} f_j(x)dx = \frac{h}{n_j h} \frac{\sum_{i=1}^{n_j} (Y_{ij} - m_j(X_{ij}))^2 - \sigma_j^2(X_{ij})}{2\sigma_j^2(X_{ij})/\sqrt{n_j}} + O_p(n_j h)^{-1}$$

ดังนั้น $$\int \frac{\hat{\sigma}_j(x) - \sigma_j(x)}{\sigma_j(x)/\sqrt{n_j}} f_X(x)dx = \frac{1}{n_j} \frac{\sum_{i=1}^{n_j} (Y_{ij} - m_j(X_{ij}))^2 - \sigma_j^2(X_{ij})}{2\sigma_j^2(X_{ij})/\sqrt{n_j}} + O_p(n_j h)^{-1}$$

ทฤษฎีบท 2 ภายใต้ข้อตกลง 1 -13 และสมมติฐานหลัก (H_0) สำหรับ $j = 1, \ldots, k$

$$\hat{F}_{\varepsilon j0}(y) - \hat{F}_{\varepsilon j}(y) = f_{\varepsilon j}(y) \sum_{r=1}^{k} p_r \left\{ \frac{1}{n_r} \sum_{i=1}^{n_r} \frac{Y_{ir} - m(X_{ir})}{\sigma_j(X_{ir})/\sqrt{n_j}} \left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r = j)}{p_j} \right) \right\}$$

$$+ o_p(n)^{-1/2}$$

พิสูจน์ $$\hat{F}_{\varepsilon j0}(y) - \hat{F}_{\varepsilon j}(y) = \left[\hat{F}_{\varepsilon j0}(y) - F_{\varepsilon j}(y) \right] - \left[\hat{F}_{\varepsilon j}(y) - F_{\varepsilon j}(y) \right] \quad (10)$$

ประยุกต์จากการพิสูจน์ทฤษฎีบทที่ 1 ของ Akritas and Van Keilegom (2001) เมื่อ

$F_e(y) = P(\varepsilon \le y) = P\left(\frac{(\overline{Y} - m(x))}{\sigma(x)/\sqrt{n}} \le y \right)$ จะได้ว่า

$$\hat{F}_e(y) - F_e(y) = f_e(y) \int \frac{y(\hat{\sigma}(x) - \sigma(x)) + (\hat{m}(x) - m(x)) + (\bar{Y} - Y)}{\sigma(x)\sqrt{n}} dF_X(x)$$

$$+ \frac{1}{n} \sum_{i=1}^{n_j} \left\{ I(\varepsilon_j \le y) \right\} - F_e(y) + o_p(n)^{-1/2} \quad (11)$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

จากสมการ (11) ภายใต้สมมุติฐานหลัก $m(x) = m_1(x) = ... = m_k(x)$

เมื่อ $\varepsilon_j = \dfrac{\bar{Y}_j - m(x)}{\sigma_j(x)/\sqrt{n_j}}$ ดังนั้น

$$\hat{F}_{\varepsilon j0}(y) - F_{\varepsilon j}(y) = f_{\varepsilon j}(y)\int \frac{y(\hat{\sigma}_j(x) - \sigma_j(x)) + (\hat{m}(x) - m(x)) + (\bar{Y}_j - Y_j)}{\sigma_j(x)/\sqrt{n_j}} f_j(x)dx$$

$$+ \frac{1}{n_j}\sum_{i=1}^{n_j} I\left(\frac{\bar{Y}_j - m(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}} \le y\right) - F_{\varepsilon j}(y) + o_p(n)^{-1/2} \qquad (12)$$

จากสมการ (12) ดังนั้น

$$\int \frac{y(\hat{\sigma}_j(x) - \sigma_j(x)) + (\hat{m}(x) - m(x)) + (\bar{Y} - Y)}{\sigma_j(x)/\sqrt{n_j}} f_j(x)dx = y\int\left(\frac{\hat{\sigma}_j(x) - \sigma_j(x)}{\sigma_j(x)/\sqrt{n_j}}\right) f_j(x)dx$$

$$+ \int \frac{\hat{m}(x) - m(x)}{\sigma_j(x)/\sqrt{n_j}} f_j(x)dx$$

$$+ \int \frac{\bar{Y}_j - Y_j}{\sigma_j(x)/\sqrt{n_j}} f_j(x)dx$$

จากบทแทรก 2

$$\int \frac{\hat{m}(x) - m(x)}{\sigma_j(x)} f_j(x)dx = \frac{1}{n}\sum_{r=1}^{k}\sum_{i=1}^{n_r} \frac{Y_{ir} - m(X_{ir})}{\sigma_j(X_{ir})/\sqrt{n_j}} \frac{f_j(X_{ir})}{f_{mix}(X_{ir})} + O_p(nh)^{-1}$$

และบทแทรก 3

$$\int \frac{\hat{\sigma}_j(x) - \sigma_j(x)}{\sigma_j(x)} f_j(x)dx = \frac{1}{n_j}\sum_{i=1}^{n_j} \frac{(Y_{ij} - m_j(X_{ij}))^2 - \sigma_j^2(X_{ij})}{2\sigma_j^2(X_{ij})/\sqrt{n_j}} + O_p(n_j h)^{-1}$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

จะได้ว่า ภายใต้สมมุติฐานหลัก $m(x) = m_1(x) = ... = m_k(x)$

$$\hat{F}_{\varepsilon j0}(y) - F_{\varepsilon j}(y) = yf_{\varepsilon j}(y)\frac{1}{n_j}\sum_{i=1}^{n_j}\frac{(Y_{ij} - m(X_{ij}))^2 - \sigma_j^2(X_{ij})}{2\sigma_j^2(X_{ij})/\sqrt{n_j}}$$

$$+ f_{\varepsilon j}(y)\frac{1}{n}\sum_{r=1}^{k}\sum_{i=1}^{n_r}\frac{Y_{ir} - m(X_{ir})}{\sigma_j(X_{ir})/\sqrt{n_j}}\frac{f_j(X_{ir})}{f_{mix}(X_{ir})}$$

$$+\int\frac{\bar{Y}_j - Y_j}{\sigma_j(x)/\sqrt{n_j}}f_j(x)dx$$

$$+\frac{1}{n_j}\sum_{i=1}^{n_j}\left(\frac{\bar{Y}_j - m(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}} \le y\right) - F_{\varepsilon j}(y) + o_p(n)^{-1/2} \qquad (13)$$

จากบทแทรก 1

$$\int\frac{\hat{m}_j(x) - m_j(x)}{\sigma_j(x)/\sqrt{n_j}}f_j(x)dx = \frac{1}{n_j}\sum_{i=1}^{n_j}\frac{Y_{ij} - m_j(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}} + O_p(n_jh)^{-1}$$ จะได้ว่า

$$\hat{F}_{\varepsilon j}(y) - F_{\varepsilon j}(y) = f_{\varepsilon j}(y)\int\frac{y(\hat{\sigma}_j(x) - \sigma_j(x)) + (\hat{m}_j(x) - m_j(x)) + (\bar{Y}_j - Y_j)}{\sigma_j(x)/\sqrt{n_j}}f_j(x)dx$$

$$+ \frac{1}{n_j}\sum_{i=1}^{n_j}\left\{I\left(\frac{\bar{Y}_j - m(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}} \le y\right)\right\} - F_{\varepsilon j}(y) + o_p(n_j)^{-1/2}$$

$$\hat{F}_{\varepsilon j}(y) - F_{\varepsilon j}(y) = yf_{\varepsilon j}(y)\frac{1}{n_j}\sum_{i=1}^{n_j}\frac{(Y_{ij} - m(X_{ij}))^2 - \sigma_j^2(X_{ij})}{2\sigma_j^2(X_{ij})/\sqrt{n_j}}$$

$$+ f_{\varepsilon j}(y)\frac{1}{n_j}\sum_{i=1}^{n_j}\frac{Y_{ij} - m(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}}$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

$$+\int \frac{\bar{Y}_j - Y_j}{\sigma_j(x)/\sqrt{n_j}} f_j(x)dx$$

$$+ \frac{1}{n_j}\sum_{i=1}^{n_j} I\left(\frac{\bar{Y}_j - m(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}} \le y\right) - F_{\varepsilon j}(y) + o_p(n_j)^{-1/2} \qquad (14)$$

$\hat{F}_{\varepsilon j0}(y) - \hat{F}_{\varepsilon j}(y)$ คำนวณได้จากสมการ (13) – (14) ดังนั้น

$$\hat{F}_{\varepsilon j0}(y) - F_{\varepsilon j}(y) = f_{\varepsilon j}(y)\frac{1}{n}\sum_{r=1}^{k}\sum_{i=1}^{n_r}\frac{Y_{ir} - m(X_{ir})}{\sigma_j(X_{ir})/\sqrt{n_j}}\frac{f_j(X_{ir})}{f_{mix}(X_{ir})}$$

$$- f_{\varepsilon j}(y)\frac{1}{n_j}\sum_{i=1}^{n_j}\frac{Y_{ij} - m(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}} + o_p(n)^{-1/2}$$

จัดรูปใหม่จะได้ว่า

$$\hat{F}_{\varepsilon j0}(y) - \hat{F}_{\varepsilon j}(y) = f_{\varepsilon j}(y)\sum_{r=1}^{k} p_r \left\{\frac{1}{n_r}\sum_{i=1}^{n_r}\frac{Y_{ir} - m(X_{ir})}{\sigma_j(X_{ir})/\sqrt{n_j}}\left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j)}{p_j}\right)\right\}$$

$$+ o_p(n)^{-1/2}$$

ทฤษฎีบท 3 สมมุติข้อตกลง 1 – 13 เป็นจริงภายใต้สมมติฐานหลัก สำหรับ $j = 1, \dots, k$ k – Dimensions ของ $\hat{U}(y) = (\hat{U}_1(y), \dots, \hat{U}_j(y))^t$ Converges Weakly เข้าสู่ $U(y) = (U_1(y), \dots, U_j(y))^t$ เมื่อ

$$U_j(y) = f_{\varepsilon j}(y)p_j\sum_{r=1}^{k}\sum_{i=1}^{n_r}\frac{Y_{ir} - m(X_{ir})}{\sigma_j(X_{ir})}\left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j)}{p_j}\right)$$ คือ ตัวแปรสุ่มปกติ

ที่มีค่าเฉลี่ยเท่ากับ 0 และมีโครงสร้างความแปรปรวนร่วมคือ

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

$$\mathrm{Cov}(U_j(y),U_{j'}(y)) = f_{\varepsilon j}(y)f_{\varepsilon j'}(y)p_j p_{j'}\sum_{r=1}^{k}E\left[\frac{\sigma_r^2(X_{ir})}{\sigma_j(X_{ir})\sigma_{j'}(X_{ir})}\left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})}-\frac{I(r=j)}{p_j}\right)\right.$$

$$\left.\left(\frac{f_{j'}(X_{ir})}{f_{mix}(X_{ir})}-\frac{I(r=j')}{p_{j'}}\right)\right]$$

พิสูจน์ โดย Cramer – Wold device แสดงให้เห็นว่า Weak Convergence ของ Multidimensional process ก็คือ Weak Convergence ของ Linear Combination ใด ๆ ของแต่ละส่วนประกอบ (Serfling, 1980)

กำหนดให้ b_j เป็นจำนวนจริง ใด ๆ $\hat{U}(y) = (\hat{U}_1(y), \ldots, \hat{U}_j(y))^t$

$\hat{U}_j(y) = n_j^{1/2}(\hat{F}_{\varepsilon j0}(y) - \hat{F}_{\varepsilon j}(y))$ โดย Cramer – Wold device Linear Combination ของ $\hat{U}(y)$ คือ

$$\hat{Z}(y) = \sum_{j=1}^{k} b_j n_j^{1/2}(\hat{F}_{\varepsilon j0}(y) - \hat{F}_{\varepsilon j}(y))$$

จากทฤษฎีบท 2

$$\hat{F}_{\varepsilon j0}(y) - F_{\varepsilon j}(y) = f_{\varepsilon j}(y)\frac{1}{n}\sum_{r=1}^{k}\sum_{i=1}^{n_r}\frac{Y_{ir} - m(X_{ir})}{\sigma_j(X_{ir})/\sqrt{n_j}}\frac{f_j(X_{ir})}{f_{mix}(X_{ir})}$$

$$- f_{\varepsilon j}(y)\frac{1}{n_j}\sum_{i=1}^{n_j}\frac{Y_{ij} - m_j(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}} + o_p(n)^{-1/2}$$

ดังนั้น

$$\hat{Z}(y) = \sum_{j=1}^{k} b_j n_j^{1/2}\left[\frac{1}{n}f_{\varepsilon j}(y)\sum_{r=1}^{k}\sum_{i=1}^{n_r}\frac{Y_{ir} - m(X_{ir})}{\sigma_j(X_{ir})/\sqrt{n_j}}\frac{f_j(X_{ir})}{f_{mix}(X_{ir})}\right.$$

$$\left. - \frac{1}{n_j}f_{\varepsilon j}(y)\sum_{i=1}^{n_j}\frac{Y_{ij} - m_j(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}}\right] + o_p(n)^{-1/2}$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

$$\hat{Z}(y) = \sum_{j=1}^{k} b_j n_j \left[\frac{1}{n} f_{\varepsilon j}(y) \sum_{r=1}^{k} \sum_{i=1}^{n_r} \frac{Y_{ir} - m(X_{ir})}{\sigma_j(X_{ir})} \frac{f_j(X_{ir})}{f_{mix}(X_{ir})} \right.$$

$$\left. - f_{\varepsilon j}(y) \sum_{i=1}^{n_j} \frac{Y_{ij} - m_j(X_{ij})}{\sigma_j(X_{ij})} \right] + o_p(n)^{-1/2}$$

จัดรูปใหม่จะได้ว่า $\hat{Z}(y) = \frac{1}{n} \sum_{r=1}^{k} \sum_{i=1}^{n_r} \frac{Y_{ir} - m(X_{ir})}{\sigma_r(X_{ir})} \sum_{j=1}^{k} b_j n_j f_{\varepsilon j}(y) \frac{\sigma_r(X_{ir})}{\sigma_j(X_{ir})} \frac{f_j(X_{ir})}{f_{mix}(X_{ir})}$

$$- \frac{1}{n} \sum_{r=1}^{k} \sum_{i=1}^{n_r} \frac{b_r n_r}{p_r} f_{\varepsilon r}(y) \sum_{i=1}^{n_j} \frac{Y_{ir} - m(X_{ir})}{\sigma_r(X_{ir})} + o_p(n)^{-1/2}$$

ดึงตัวร่วม $\frac{1}{n} \sum_{r=1}^{k} \sum_{i=1}^{n_r} \frac{Y_{ir} - m(X_{ir})}{\sigma_r(X_{ir})}$ จะได้ว่า

$$\hat{Z}(y) = \frac{1}{n} \sum_{r=1}^{k} \sum_{i=1}^{n_r} \frac{Y_{ir} - m(X_{ir})}{\sigma_r(X_{ir})} \left(\sum_{j=1}^{k} b_j n_j f_{\varepsilon j}(y) \frac{\sigma_r(X_{ir})}{\sigma_j(X_{ir})} \frac{f_j(X_{ir})}{f_{mix}(X_{ir})} - \frac{b_r n_r}{p_r} f_{\varepsilon r}(y) \right)$$

$$+ o_p(n)^{-1/2}$$

$$\hat{Z}(y) = \sum_{r=1}^{k} \sum_{i=1}^{n_r} \frac{Y_{ir} - m(X_{ir})}{\sigma_r(X_{ir})} \left(\sum_{j=1}^{k} b_j p_j f_{\varepsilon j}(y) \frac{\sigma_r(X_{ir})}{\sigma_j(X_{ir})} \frac{f_j(X_{ir})}{f_{mix}(X_{ir})} - b_r f_{\varepsilon r}(y) \right)$$

$$+ o_p(n)^{-1/2}$$

$$\hat{Z}(y) = \sum_{r=1}^{k} \sum_{i=1}^{n_r} \Psi_r(X_{ir}, Y_{ir}, y) + o_p(n)^{-1/2}$$

เมื่อ $\Psi_r(u,v,y) = \frac{v - m(u)}{\sigma_r(u)} \left(\sum_{j=1}^{k} b_j p_j f_{\varepsilon j}(y) \frac{\sigma_r(u)}{\sigma_j(u)} \frac{f_j(u)}{f_{mix}(u)} - b_r f_{\varepsilon r}(y) \right)$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

สำหรับ $r = 1, \ldots, k, \quad \hat{Z}_r(y) = \sum_{i=1}^{n_r} \Psi_r(X_{ir}, Y_{ir}, y)$

พิจารณา Class ของ Function

$$\mathcal{F}_r = \{(u,v) \rightarrow \Psi_r(u,v,y), \ -\infty < y < \infty\} \quad \mathcal{F} \text{ คือ } \mathcal{F} - \text{Indexed Process}$$

ดังนั้น $\hat{Z}_r(y)$ คือ $\mathcal{F}_r - \text{Indexed Process}$ (Van der Vaart and Wellner, 1996)

โดยทั่วไปสำหรับ Class ของฟังก์ชัน G_1 และ G_2 ได้นิยาม

$G_1 + G_2 = \{g_1 + g_2, \ g_1 \in G_1, \ g_2 \in G_2\}$ ดังนั้นสัญลักษณ์ของ Class $\mathcal{F}_r$ สามารถเขียนได้ในรูปของ

$$\mathcal{F}_r = \sum_{j=1}^{k+1} F_{rj} \quad \text{สำหรับ } j = 1, \ldots, k$$

ดังนั้น $$\mathcal{F}_{rj} = \left\{(u,v) \rightarrow b_j p_j f_{\varepsilon j}(y) \frac{\sigma_r(u)}{\sigma_j(u)} \frac{f_j(u)}{f_{mix}(u)} \frac{v - m(u)}{\sigma_r(u)}, \ -\infty < y < \infty \right\}$$

$$\mathcal{F}_{r,k+1} = \left\{(u,v) \rightarrow -b_r f_{\varepsilon r}(y) \frac{v - m(u)}{\sigma_r(u)}, \ -\infty < y < \infty \right\}$$

กำหนดให้ M อยู่ในลักษณะของ $\sup_{y,\, j=1,\ldots,k} \left| f_{\varepsilon j}(y) \right| < M$

ดังนั้น $N_{[\ ]}(\delta, \mathcal{F}_{rj}, L_2(P)) \leq 2M\delta^{-1}$ ถ้า $\delta < 2M$ และ

$N_{[\ ]}(\delta, \mathcal{F}_{rj}, L_2(P)) = 1$ ถ้า $\delta > 2M$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

เมื่อ $N_{[\ ]}$ คือ Bracketing Number

P คือ Probability Measure ที่สอดคล้องกับ Joint Distribution ของ (X_r , Y_r)

$L_2(P)$ คือ L_2 - Norm (Van der Vaart and Wellner , 1996)

นอกจากนั้นโดยทฤษฎีบท 2.10.16 ของ Van der Vaart and Wellner (1996) พบว่า

$$N_{[\ ]}(\delta, F_r, L_2(P)) \leq \prod_{j=1}^{k+1} N_{[\ ]}(\delta, F_{rj}, L(p)) \qquad (15)$$

Take log สมการ (15) จะได้ว่า

$$\log N_{[\ ]}(\delta, F_r, L_2(P)) \leq \sum_{j=1}^{k+1} \log N_{[\ ]}(\delta, F_{rj}, L_2(P)) \text{ และ}$$

$$\int_0^{\infty} \sqrt{\log N_{[\ ]}(\delta, F_r, L_2(P))}\, d\delta \leq \sum_{j=1}^{k+1} \int_0^{2M} \sqrt{\log N_{[\ ]}(\delta, F_{rj}, L_2(P))}\, d\delta \qquad (16)$$

จาก (16) สามารถสรุปได้ว่า $\int_0^{\infty} \sqrt{\log N_{[\ ]}(\delta, F_r, L_2(P))}\, d\delta$ Finite และจากทฤษฎี 2.5.6 ใน Van der Vaart and Wellner (1996) Class ของ Function คือ Donsker ดังนั้น Weak Convergence ของ $\hat{Z}_r(y)$ จะลู่เข้าสู่ $Z_r(y)$ ซึ่งคือ Gaussian Process ที่มีค่าเฉลี่ยเท่ากับ 0 และ Covariance Function

$$Cov(Z_r(y), Z_r(y')) = E\left[\Psi(X_r, Y_r, y)\Psi(X_r, Y_r, y')\right]$$

จาก Cramer – Wold Device สามารถประยุกต์ใช้กับ Weak Convergence ของ k – Dimensional ของ $\hat{U}(y)$ กล่าวคือ $\hat{U}_r(y)$ คือ $\hat{Z}_r(y)$ เมื่อ $b_r = 1$ จาก

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

$$\hat{Z}_r(y) = \sum_{i=1}^{n_r} \Psi_r(X_{ir}, Y_{ir}, y)$$

$$\hat{Z}_r(y) = \sum_{i=1}^{n_r} \frac{Y_{ir} - m(X_{ir})}{\sigma_r(X_{ir})} \left(\sum_{j=1}^{k} b_j p_j f_{\varepsilon j}(y) \frac{\sigma_r(X_{ir})}{\sigma_j(X_{ir})} \frac{f_j(X_{ir})}{f_{mix}(X_{ir})} - b_r f_{\varepsilon r}(y) \right) + o_p(n)^{-1/2}$$

ดังนั้น

$$\hat{U}_r(y) = \sum_{i=1}^{n_r} \frac{Y_{ir} - m(X_{ir})}{\sigma_r(X_{ir})} \left(\sum_{j=1}^{k} p_j f_{\varepsilon j}(y) \frac{\sigma_r(X_{ir})}{\sigma_j(X_{ir})} \frac{f_j(X_{ir})}{f_{mix}(X_{ir})} - f_{\varepsilon r}(y) \right) + o_p(n)^{-1/2}$$

$$= f_{\varepsilon j}(y) p_j \sum_{r=1}^{k} \sum_{i=1}^{n_r} \frac{Y_r - m(X_{ir})}{\sigma_j(X_{ir})} \left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j)}{p_j} \right) + o_p(n)^{-1/2} \qquad (17)$$

จากสมการ (17) $\hat{U}(y) = (\hat{U}_1(y), \ldots, \hat{U}_j(y))^t$ จะ Converges Weakly เข้าสู่

$U(y) = (U_1(y), \ldots, U_j(y))^t$ เมื่อ

$$U_j(y) = f_{\varepsilon j}(y) p_j \sum_{r=1}^{k} \sum_{i=1}^{n_r} \frac{Y_{ir} - m(X_{ir})}{\sigma_j(X_{ir})} \left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j)}{p_j} \right)$$

คือตัวแปรสุ่มปกติ ที่มีค่าเฉลี่ยเท่ากับ 0 และมีโครงสร้างความแปรปรวนร่วมคือ

$$Cov(U_j(y), U_{j'}(y)) = f_{\varepsilon j}(y) f_{\varepsilon j'}(y) p_j p_{j'} \sum_{r=1}^{k} E\left[\sum_{i=1}^{n_r} \frac{(Y_{ir} - m(X_{ir}))^2}{\sigma_j(X_{ir}) \sigma_{j'}(X_{ir})} \left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j)}{p_j} \right) \right.$$

$$\left. \left(\frac{f_{j'}(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j')}{p_{j'}} \right) \right]$$

เมื่อ $\sum_{i=1}^{n_r} (Y_{ir} - m(X_{ir}))^2 = \sigma_{ir}^2(X)$ ดังนั้น

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

$$\mathrm{Cov}(U_j(y),U_{j'}(y)) = f_{\varepsilon j}(y)f_{\varepsilon j'}(y)p_j p_{j'}\sum_{r=1}^{k}E\left[\frac{\sigma_r^2(X_{ir})}{\sigma_j(X_{ir})\sigma_{j'}(X_{ir})}\left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j)}{p_j}\right)\right.$$

$$\left.\left(\frac{f_{j'}(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j')}{p_{j'}}\right)\right]$$

ทฤษฎีบท 4.1 สมมติข้อตกลง 1–13 เป็นจริง ภายใต้สมมติฐานหลัก สำหรับ $j = 1, \ldots, k$

$$U_{KS}^1 \xrightarrow{d} \sum_{j=1}^{k}\sup_y\left|U_j(y)\right| \qquad \text{เมื่อ} \quad U_{KS}^1 = \sum_{j=1}^{k}\sup_y\left|\hat{U}_j(y)\right|$$

พิสูจน์ เนื่องจาก $U_{KS}^1 = \sum_{j=1}^{k}\sup_y\left|\hat{U}_j(y)\right|$ จาก ทฤษฎีบท 3 และ Continuous Mapping Theorem จะได้ว่า

$$U_{KS}^1 = \sum_{j=1}^{k}\sup_y\left|\hat{U}_j(y)\right| \xrightarrow{d} \sum_{j=1}^{k}\sup_y\left|U_j(y)\right|$$

เมื่อ $U_j(y) = f_{\varepsilon j}(y)p_j\sum_{r=1}^{k}\sum_{i=1}^{n_r}\frac{Y_{ir} - m(X_{ir})}{\sigma_j(X_{ir})}\left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j)}{p_j}\right)$ คือ ตัวแปรสุ่มปกติ ที่มีค่าเฉลี่ยเท่ากับ 0 และมีโครงสร้างความแปรปรวนร่วมคือ

$$\mathrm{Cov}(U_j(y),U_{j'}(y)) = f_{\varepsilon j}(y)f_{\varepsilon j'}(y)p_j p_{j'}\sum_{r=1}^{k}E\left[\frac{\sigma_r^2(X_{ir})}{\sigma_j(X_{ir})\sigma_{j'}(X_{ir})}\left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j)}{p_j}\right)\right.$$

$$\left.\left(\frac{f_{j'}(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j')}{p_{j'}}\right)\right]$$

ดังนั้น $\left|U_j(y)\right|$ คือ ตัวแปรสุ่ม Half – Normal ทีมีค่าเฉลี่ย

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

$$\sqrt{f_{\varepsilon j}(y)f_{\varepsilon j'}(y)p_j p_{j'}\sum_{r=1}^{k}E\left[\frac{\sigma_r^2(X_{ir})}{\sigma_j(X_{ir})\sigma_{j'}(X_{ir})}\left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})}-\frac{I(r=j)}{p_j}\right)\right.}$$

$$\left.\sqrt{\left(\frac{f_{j'}(X_{ir})}{f_{mix}(X_{ir})}-\frac{I(r=j')}{p_{j'}}\right)}\right]\sqrt{\frac{2}{\pi}}$$

และมีโครงสร้างความแปรปรวนร่วมคือ

$$Cov(U_j(y),U_{j'}(y)) = f_{\varepsilon j}(y)f_{\varepsilon j'}(y)p_j p_{j'}\sum_{r=1}^{k}E\left[\frac{\sigma_r^2(X_{ir})}{\sigma_j(X_{ir})\sigma_{j'}(X_{ir})}\left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})}-\frac{I(r=j)}{p_j}\right)\right.$$

$$\left.\left(\frac{f_{j'}(X_{ir})}{f_{mix}(X_{ir})}-\frac{I(r=j')}{p_{j'}}\right)\right]\left[1-\frac{2}{\pi}\right]$$

ทฤษฎีบท 4.2 สมมติข้อตกลง 1 – 13 เป็นจริง ภายใต้สมมติฐานหลัก สำหรับ $j = 1, \ldots, k$

$$U_{KU}^{1} \xrightarrow{d} \sum_{j=1}^{k}\left[\sup_{y}\left|U_j(y)\right| - \inf_{y}\left|U_j(y)\right|\right]$$

เมื่อ $U_{ku}^{1} = \sum_{j=1}^{k}\left[\sup_{y}\left|\hat{U}_j(y)\right| - \inf_{y}\left|\hat{U}_j(y)\right|\right]$

พิสูจน์ เนื่องจาก $U_{ku}^{1} = \sum_{j=1}^{k}\left[\sup_{y}\left|\hat{U}_j(y)\right| - \inf_{y}\left|\hat{U}_j(y)\right|\right]$

จากทฤษฎีบท 3 และ Continuous Mapping Theorem จะได้ว่า

$$U_{ku}^{1} = \sum_{j=1}^{k}\left[\sup_{y}\left|\hat{U}_j y\right| - \inf_{y}\left|\hat{U}_j y\right|\right] \xrightarrow{d} \sum_{j=1}^{k}\left[\sup_{y}\left|U_j(y)\right| - \inf_{y}\left|U_j(y)\right|\right]$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

เมื่อ $U_j(y) = f_{\varepsilon j}(y)p_j\sum_{r=1}^{k}\sum_{i=1}^{n_r}\frac{Y_{ir} - m(X_{ir})}{\sigma_j(X_{ir})}\left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j)}{p_j}\right)$

คือ ตัวแปรสุ่มปกติที่มีค่าเฉลี่ยเท่ากับ 0 และมีโครงสร้างความแปรปรวนร่วมคือ

$$Cov(U_j(y),U_{j'}(y)) = f_{\varepsilon j}(y)f_{\varepsilon j'}(y)p_j p_{j'}\sum_{r=1}^{k}E\left[\frac{\sigma_r^2(X_{ir})}{\sigma_j(X_{ir})\sigma_{j'}(X_{ir})}\left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j)}{p_j}\right)\right.$$

$$\left.\left(\frac{f_{j'}(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j')}{p_{j'}}\right)\right]$$

ดังนั้น $|U_j(y)|$ คือ ตัวแปรสุ่ม Half – Normal ที่มีค่าเฉลี่ย

$$\sqrt{f_{\varepsilon j}(y)f_{\varepsilon j'}(y)p_j p_{j'}\sum_{r=1}^{k}E\left[\frac{\sigma_r^2(X_{ir})}{\sigma_j(X_{ir})\sigma_{j'}(X_{ir})}\left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j)}{p_j}\right)\right.}$$

$$\left.\sqrt{\left(\frac{f_{j'}(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j')}{p_{j'}}\right)}\right]\sqrt{\frac{2}{\pi}}$$

และ มีโครงสร้างความแปรปรวนร่วมคือ

$$Cov(U_j(y),U_{j'}(y)) = f_{\varepsilon j}(y)f_{\varepsilon j'}(y)p_j p_{j'}\sum_{r=1}^{k}E\left[\frac{\sigma_r^2(X_{ir})}{\sigma_j(X_{ir})\sigma_{j'}(X_{ir})}\left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j)}{p_j}\right)\right.$$

$$\left.\left(\frac{f_{j'}(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j')}{p_{j'}}\right)\right]\left[1 - \frac{2}{\pi}\right]$$

ทฤษฎีบท 4.3 สมมติข้อตกลง 1– 13 เป็นจริง ภายใต้สมมติฐานหลัก สำหรับ $j = 1, \ldots, k$

$$U_{KS}^2 \xrightarrow{d} \sup_y |U(y)|$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

เมื่อ $U_{KS}^2 = \sup_y \left|\hat{U}(y)\right|$

$$U(y) = \sum_{j=1}^{k} p_j^{1/2} U_j(y)$$

พิสูจน์ เนื่องจาก $U_{KS}^2 = \sup_y \left|\hat{U}(y)\right|$ โดย $\hat{U}(y) = n^{1/2}(\hat{F}_{\varepsilon 0}(y) - \hat{F}_{\varepsilon}(y))$ ดังนั้น

$$\hat{U}(y) = n^{1/2}\left(\frac{1}{n}\sum_{j=1}^{k}\sum_{i=1}^{n_j} I\left(\frac{\overline{Y}_j - \hat{m}(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}} \le y\right) - \frac{1}{n}\sum_{j=1}^{k}\sum_{i=1}^{n_j} I\left(\frac{\overline{Y}_j - \hat{m}_j(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}} \le y\right)\right)$$

คูณทั้งเศษและส่วนด้วย n_j

$$\hat{U}(y) = n^{1/2}\sum_{j=1}^{k}\left(\frac{n_j}{n}\frac{1}{n_j}\sum_{i=1}^{n_j} I\left(\frac{\overline{Y}_j - \hat{m}(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}} \le y\right) - \frac{n_j}{n}\frac{1}{n_j}\sum_{i=1}^{n_j} I\left(\frac{\overline{Y}_j - \hat{m}_j(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}} \le y\right)\right)$$

$$\hat{U}(y) = n^{-1/2}\sum_{j=1}^{k}\left(\frac{n_j}{n_j}\sum_{i=1}^{n_j} I\left(\frac{\overline{Y}_j - \hat{m}(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}} \le y\right) - \frac{n_j}{n_j}\sum_{i=1}^{n_j} I\left(\frac{\overline{Y}_j - \hat{m}_j(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}} \le y\right)\right)$$

$$\hat{U}(y) = n^{-1/2}\sum_{j=1}^{k} n_j\left(\hat{F}_{\varepsilon j0}(y) - \hat{F}_{\varepsilon j}(y)\right)$$

คูณทั้งเศษและส่วนด้วย $n_j^{-1/2}$

$$\hat{U}(y) = n^{-1/2}\sum_{j=1}^{k}\left(\frac{n_j^{-1/2}}{n_j^{-1/2}}\right) n_j\left(\hat{F}_{\varepsilon j0}(y) - \hat{F}_{\varepsilon j}(y)\right)$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

$$\hat{U}(y) = \sum_{j=1}^{k}\left(\frac{n_j^{1/2}}{n^{1/2}}\right) n_j^{1/2}\left(\hat{F}_{\varepsilon j0}(y) - \hat{F}_{\varepsilon j}(y)\right)$$

$$\hat{U}(y) = \sum_{j=1}^{k} p_j^{1/2}\hat{U}_j(y)$$

จาก $\hat{Z}(y) = \sum_{j=1}^{k} b_j n_j^{1/2}(\hat{F}_{\varepsilon j0}(y) - \hat{F}_{\varepsilon j}(y))$ เป็น Linear Combination ของ $\hat{U}_j(y)$

เมื่อ $\hat{U}_j(y) = n_j^{1/2}(\hat{F}_{\varepsilon j0}(y) - \hat{F}_{\varepsilon j}(y))$ และเมื่อให้ $b_j = p_j^{1/2}$

จะได้ว่า $\hat{U}(y) = \sum_{j=1}^{k} p_j^{1/2}(\hat{U}_j(y)) = \sum_{j=1}^{k} p_j^{1/2} n_j^{1/2}(\hat{F}_{\varepsilon j0}(y) - \hat{F}_{\varepsilon j}(y))$

ดังนั้นสามารถกล่าวได้ว่า $\hat{Z}(y)$ คือ กรณีเฉพาะสำหรับ Linear Combination ของ $\hat{U}(y)$ เมื่อ

$b_j = p_j^{1/2}$ ดังนั้นจากทฤษฎีบท 3 และ Continuous Mapping Theorem จะได้ว่า

$$U_{KS}^2 = \sup_y\left|\hat{U}(y)\right| \xrightarrow{d} \sup_y\left|U(y)\right|$$

ดังนั้น $U_{KS}^2 \xrightarrow{d} \sup_y\left|U(y)\right|$, $U(y) = \sum_{j=1}^{k} p_j^{1/2} U_j(y)$

เนื่องจาก $U_j(y) = f_{\varepsilon j}(y) p_j \sum_{r=1}^{k}\sum_{i=1}^{n_r} \frac{Y_{ir} - m(X_{ir})}{\sigma_j(X_{ir})}\left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j)}{p_j}\right)$

คือตัวแปรสุ่มปกติ ที่มีค่าเฉลี่ยเท่ากับ 0 และโครงสร้างความแปรปรวนร่วมคือ

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

$$\mathrm{Cov}(U_j(y),U_{j'}(y)) = f_{\varepsilon j}(y)f_{\varepsilon j'}(y)p_j p_{j'}\sum_{r=1}^{k}E\left[\frac{\sigma_r^2(X_{ir})}{\sigma_j(X_{ir})\sigma_{j'}(X_{ir})}\left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})}-\frac{I(r=j)}{p_j}\right)\right.$$

$$\left.\left(\frac{f_{j'}(X_{ir})}{f_{mix}(X_{ir})}-\frac{I(r=j')}{p_{j'}}\right)\right]$$

ดังนั้น $\left|U_j(y)\right|$ คือ ตัวแปรสุ่ม half – normal ที่มีค่าเฉลี่ย

$$\sqrt{f_{\varepsilon j}(y)f_{\varepsilon j'}(y)p_j p_{j'}\sum_{r=1}^{k}E\left[\frac{\sigma_r^2(X_{ir})}{\sigma_j(X_{ir})\sigma_{j'}(X_{ir})}\left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})}-\frac{I(r=j)}{p_j}\right)\right.}$$

$$\left.\sqrt{\left(\frac{f_{j'}(X_{ir})}{f_{mix}(X_{ir})}-\frac{I(r=j')}{p_{j'}}\right)}\right]\sqrt{\frac{2}{\pi}}$$

และมีโครงสร้างความแปรปรวนร่วมคือ

$$\mathrm{Cov}(U_j(y),U_{j'}(y)) = f_{\varepsilon j}(y)f_{\varepsilon j'}(y)p_j p_{j'}\sum_{r=1}^{k}E\left[\frac{\sigma_r^2(X_{ir})}{\sigma_j(X_{ir})\sigma_{j'}(X_{ir})}\left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})}-\frac{I(r=j)}{p_j}\right)\right.$$

$$\left.\left(\frac{f_{j'}(X_{ir})}{f_{mix}(X_{ir})}-\frac{I(r=j')}{p_{j'}}\right)\right]\left[1-\frac{2}{\pi}\right]$$

ทฤษฎีบท 4.4 สมมติข้อตกลง 1– 13 เป็นจริง ภายใต้สมมติฐานหลัก สำหรับ $j = 1, \ldots, k$

$$U_{ku}^2 \xrightarrow{d} \left[\sup_y\left|U(y)\right| - \inf_y\left|U(y)\right|\right]$$

เมื่อ $U_{ku}^2 = \left[\sup_y\left|\hat{U}(y)\right| - \inf_y\left|\hat{U}(y)\right|\right]$, $U(y) = \sum_{j=1}^{k}p_j^{1/2}U_j(y)$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

พิสูจน์ เนื่องจาก $U_{ku}^2 = \left[\sup_y \left|\hat{U}(y)\right| - \inf_y \left|\hat{U}(y)\right| \right]$

โดย $\hat{U}(y) = n^{1/2}(\hat{F}_{\varepsilon 0}(y) - \hat{F}_{\varepsilon}(y))$ ดังนั้น

$$\hat{U}(y) = n^{1/2}\left(\frac{1}{n}\sum_{j=1}^{k}\sum_{i=1}^{n_j} I\left(\frac{\overline{Y}_j - \hat{m}(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}} \le y \right) - \frac{1}{n}\sum_{j=1}^{k}\sum_{i=1}^{n_j} I\left(\frac{\overline{Y}_j - \hat{m}_j(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}} \le y \right) \right)$$

คูณทั้งเศษและส่วนด้วย n_j

$$\hat{U}(y) = n^{1/2}\sum_{j=1}^{k}\left(\frac{n_j}{n}\frac{1}{n_j}\sum_{i=1}^{n_j} I\left(\frac{\overline{Y}_j - \hat{m}(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}} \le y \right) - \frac{n_j}{n}\frac{1}{n_j}\sum_{i=1}^{n_j} I\left(\frac{\overline{Y}_j - \hat{m}_j(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}} \le y \right) \right)$$

$$\hat{U}(y) = n^{-1/2}\sum_{j=1}^{k}\left(\frac{n_j}{n_j}\sum_{i=1}^{n_j} I\left(\frac{\overline{Y}_j - \hat{m}(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}} \le y \right) - \frac{n_j}{n_j}\sum_{i=1}^{n_j} I\left(\frac{\overline{Y}_j - \hat{m}_j(X_{ij})}{\hat{\sigma}_j(X_{ij})/\sqrt{n_j}} \le y \right) \right)$$

$$\hat{U}(y) = n^{-1/2}\sum_{j=1}^{k} n_j\left(\hat{F}_{\varepsilon j0}(y) - \hat{F}_{\varepsilon j}(y) \right)$$

คูณทั้งเศษและส่วนด้วย $n_j^{-1/2}$

$$\hat{U}(y) = n^{-1/2}\sum_{j=1}^{k}\left(\frac{n_j^{-1/2}}{n_j^{-1/2}} \right) n_j\left(\hat{F}_{\varepsilon j0}(y) - \hat{F}_{\varepsilon j}(y) \right)$$

$$\hat{U}(y) = \sum_{j=1}^{k}\left(\frac{n_j^{1/2}}{n^{1/2}} \right) n_j^{1/2}\left(\hat{F}_{\varepsilon j0}(y) - \hat{F}_{\varepsilon j}(y) \right),\ \hat{U}(y) = \sum_{j=1}^{k} p_j^{1/2}\hat{U}_j(y)$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

จาก $\hat{Z}(y) = \sum_{j=1}^{k} b_j n_j^{1/2}(\hat{F}_{\varepsilon j0}(y) - \hat{F}_{\varepsilon j}(y))$ เป็น Linear Combination ของ $\hat{U}_j(y)$

เมื่อ $\hat{U}_j(y) = n_j^{1/2}(\hat{F}_{\varepsilon j0}(y) - \hat{F}_{\varepsilon j}(y))$ และเมื่อให้ $b_j = p_j^{1/2}$

จะได้ว่า $\hat{U}(y) = \sum_{j=1}^{k} p_j^{1/2}(\hat{U}_j(y)) = \sum_{j=1}^{k} p_j^{1/2} n_j^{1/2}(\hat{F}_{\varepsilon j0}(y) - \hat{F}_{\varepsilon j}(y))$

ดังนั้นสามารถกล่าวได้ว่า $\hat{Z}(y)$ คือ กรณีเฉพาะสำหรับ Linear Combination ของ $\hat{U}(y)$ เมื่อ

$b_j = p_j^{1/2}$ ดังนั้นจากทฤษฎีบท 3 และ Continuous Mapping Theorem จะได้ว่า

$$U_{ku}^2 = \left[\sup_y \left|\hat{U}(y)\right| - \inf_y \left|\hat{U}(y)\right| \right] \xrightarrow{d} \left[\sup_y \left|U(y)\right| - \inf_y \left|U(y)\right| \right]$$

$$U(y) = \sum_{j=1}^{k} p_j^{1/2} U_j(y)$$

เนื่องจาก $U_j(y) = f_{\varepsilon j}(y) p_j \sum_{r=1}^{k} \sum_{i=1}^{n_r} \frac{Y_{ir} - m(X_{ir})}{\sigma_j(X_{ir})} \left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j)}{p_j} \right)$ คือ

ตัวแปรสุ่มปกติ ที่มีค่าเฉลี่ยเท่ากับ 0 และมีโครงสร้างความแปรปรวนร่วมคือ

$$Cov(U_j(y), U_{j'}(y)) = f_{\varepsilon j}(y) f_{\varepsilon j'}(y) p_j p_{j'} \sum_{r=1}^{k} E\left[\frac{\sigma_r^2(X_{ir})}{\sigma_j(X_{ir})\sigma_{j'}(X_{ir})} \left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j)}{p_j} \right) \right.$$

$$\left. \left(\frac{f_{j'}(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j')}{p_{j'}} \right) \right]$$

ดังนั้น $\left|U_j(y)\right|$ คือ ตัวแปรสุ่ม half – normal ที่มีค่าเฉลี่ย

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

$$\sqrt{f_{\varepsilon j}(y)f_{\varepsilon j'}(y)p_j p_{j'}\sum_{r=1}^{k}E\left[\frac{\sigma_r^2(X_{ir})}{\sigma_j(X_{ir})\sigma_{j'}(X_{ir})}\left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})}-\frac{I(r=j)}{p_j}\right)\right.}$$

$$\left.\sqrt{\left(\frac{f_{j'}(X_{ir})}{f_{mix}(X_{ir})}-\frac{I(r=j')}{p_{j'}}\right)}\right]\sqrt{\frac{2}{\pi}}$$

และมีโครงสร้างความแปรปรวนร่วมคือ

$$Cov(U_j(y),U_{j'}(y)) = f_{\varepsilon j}(y)f_{\varepsilon j'}(y)p_j p_{j'}\sum_{r=1}^{k}E\left[\frac{\sigma_r^2(X_{ir})}{\sigma_j(X_{ir})\sigma_{j'}(X_{ir})}\left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})}-\frac{I(r=j)}{p_j}\right)\right.$$

$$\left.\left(\frac{f_{j'}(X_{ir})}{f_{mix}(X_{ir})}-\frac{I(r=j')}{p_{j'}}\right)\right]\left[1-\frac{2}{\pi}\right]$$

ทฤษฎีบทที่ 5 สมมติข้อตกลง 11 – 14 เป็นจริง ภายใต้สมมติฐานทางเลือกสำหรับ $j = 1, \ldots, k$

$$\hat{F}_{\varepsilon j0}(y) - \hat{F}_{\varepsilon j}(y) = f_{\varepsilon j}(y)\sum_{r=1}^{k}p_r\left\{\frac{1}{n_r}\sum_{i=1}^{n_r}\frac{Y_{ir}-m(X_{ir})}{\sigma_j(X_{ir})/\sqrt{n_j}}\left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})}-\frac{I(r=j)}{p_j}\right)\right\}$$

$$+\ n^{-1/2}f_{\varepsilon j}(y)E\left[\frac{S(X_{ij})-s_j(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}}\right]+o_p(n)^{-1/2}$$

พิสูจน์ ภายใต้สมมติฐานทางเลือก $\hat{m}(x)$ เป็นตัวประมาณของ

$$m_n(x) = m_0(x) + n^{-1/2}S(x) \qquad (18)$$

เมื่อ $S(x) = \sum_{j=1}^{k}p_j\frac{f_j(x)}{f_{mix}(x)}s_j(x)$ และ $\hat{m}_{jn}(x)$ เป็นตัวประมาณของ

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

$$m_{jn}(x) = m_0(x) + n^{-1/2}s_j(x) \quad (19)$$

สมการ (18) – (19) จะได้ว่า

$$m_n(x) - m_{jn}(x) = n^{-1/2}\left(S(x) - s_j(x)\right)$$

ดังนั้น $$m_n(x) = m_{jn}(x) + n^{-1/2}\left(S(x) - s_j(x)\right) \quad (20)$$

จากฟังก์ชันความน่าจะเป็นของ y สามารถประยุกต์ใช้ Taylor Expansion ได้ดังนี้

เนื่องจาก $\hat{F}_{\varepsilon j0}(y)$ เป็นตัวประมาณของ $F_{\varepsilon j0}(y)$ และ $F_{\varepsilon j0}(y) = P\left(\frac{\bar{Y} - m_n(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}} \le y\right)$

แทน $m_n(X_j)$ ด้วย $m_{jn}(X_{ij}) + n^{-1/2}\left(S(X_{ij}) - s_j(X_{ij})\right)$ จะได้

$$F_{\varepsilon j0}(y) = P\left(\frac{\bar{Y}_j - m_{jn}(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}} - n^{-1/2}\left(\frac{S(X_{ij}) - s_j(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}}\right) \le y\right)$$

$$= \int\left(P\left(\frac{\bar{Y}_j - m_{jn}(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}} - n^{-1/2}\left(\frac{S(X_{ij}) - s_j(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}}\right) \le y \middle| X_{ij} = x\right)\right) f_j(x)dx$$

ประยุกต์ใช้ Taylor Expansion จะได้ว่า

$$F_{\varepsilon j0}(y) = \int P\left(\frac{\bar{Y} - m_{jn}(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}} \le y \middle| X_{ij} = x\right) f_j(x)dx$$

$$+ n^{-1/2}P'\left(\frac{\bar{Y} - m_{jn}(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}} \le y \middle| X_{ij} = x\right)\int\left(\frac{S(X_{ij}) - s_j(X_{ij}}{\sigma_j(X_{ij})/\sqrt{n_j}} \le y \middle| X_{ij} = x\right) f_j(x)dx$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

$$+ n^{-1/2}P''\left(\frac{\bar{Y} - m_{jn}(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}} \le y \middle| X_{ij} = x\right)\int\left(\frac{S(X_{ij}) - s_j(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}} \le y \middle| X_{ij} = x\right)f_j(x)dx$$

$$+ o_p(n)^{-1/2}$$

ดังนั้น $$F_{\varepsilon j0}(y) = F_{\varepsilon j}(y) + n^{-1/2}f_{\varepsilon j}(y)E\left(\frac{S(X_{ij}) - s_j(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}}\right) + o_p(n)^{-1/2} \qquad (21)$$

จากสมการ (12) ในการพิสูจน์ทฤษฎีบท 2

$$\hat{F}_{\varepsilon j0}(y) = \frac{1}{n_j}\sum_{i=1}^{n_j} I\left(\frac{\bar{Y}_j - m_n(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}} \le y\right)$$

$$+ f_{\varepsilon j}(y)\int\frac{y(\hat{\sigma}_j(x) - \sigma_j(x))}{\sigma_j(x)/\sqrt{n_j}}f_j(x)dx \qquad (22)$$

$$+ f_{\varepsilon j}(y)\int\frac{\bar{Y}_j - Y_j}{\sigma_j(x)/\sqrt{n_j}}f_j(x)dx$$

$$+ f_{\varepsilon j}(y)\int\frac{\hat{m}(x) - m_n(x)}{\sigma_j(x)/\sqrt{n_j}}f_j(x)dx + o_p(n)^{-1/2}$$

จากสมการ (21) จะได้ว่า

$$\hat{F}_{\varepsilon j0}(y) = \frac{1}{n_j}\sum_{i=1}^{n_j} I\left(\frac{\bar{Y}_j - m_{jn}(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}} \le y\right)$$

$$+ f_{\varepsilon j}(y)\int\frac{y(\hat{\sigma}_j(x) - \sigma_j(x))}{\sigma_j(x)/\sqrt{n_j}}f_j(x)dx$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

$$+ f_{\varepsilon j}(y)\int \frac{\bar{Y}_j - Y_j}{\sigma_j(x)/\sqrt{n_j}} f_j(x)dx$$

$$+ f_{\varepsilon j}(y)\int \frac{\hat{m}(x) - m_n(x)}{\sigma_j(x)/\sqrt{n_j}} f_j(x)dx$$

$$+ n^{-1/2} f_{\varepsilon j}(y) E\left(\frac{S(X_{ij}) - s_j(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}} \right) + o_p(n)^{-1/2} \qquad (23)$$

จากผลการพิสูจน์ในทฤษฎีบทที่ 2 สมการ (14)

$$\hat{F}_{\varepsilon j}(y) = \frac{1}{n_j}\sum_{i=1}^{n_j} I\left(\frac{\bar{Y}_j - m_{jn}(X_{ij})}{\sigma_j(X_{ij}/\sqrt{n_j})} \le y \right)$$

$$+ f_{\varepsilon j}(y)\int \frac{y(\hat{\sigma}(x) - \sigma(x))}{\sigma_j(x)/\sqrt{n_j}} f_j(x)dx$$

$$+ f_{\varepsilon j}(y)\int \frac{\bar{Y}_j - Y_j}{\sigma_j(x)/\sqrt{n_j}} f_j(x)dx$$

$$+ f_{\varepsilon j}(y)\int \left(\frac{\hat{m}_j(x) - m_{jn}(x)}{\sigma_j(x)/\sqrt{n_j}} \right) f_j(x)dx + o_p(n_j)^{-1/2} \qquad (24)$$

สมการ (24) – (23) จะได้ว่า

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

$$\hat{F}_{\varepsilon j0}(y) - \hat{F}_{\varepsilon j}(y) = f_{\varepsilon j}(y)\int\left(\frac{\hat{m}(x) - m_n(x)}{\sigma_j(x)/\sqrt{n_j}}\right)f_j(x)dx$$

$$- f_{\varepsilon j}(y)\int\left(\frac{\hat{m}_j(x) - m_{jn}(x)}{\sigma_j(x)/\sqrt{n_j}}\right)f_j(x)dx$$

$$+ n^{-1/2}f_{\varepsilon j}(y)E\left(\frac{S(X_{ij}) - s_j(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}}\right) + o_p(n)^{-1/2}$$

จากบทแทรกที่ 1

$$\int\left(\frac{\hat{m}(x) - m_n(x)}{\sigma_j(x)/\sqrt{n_j}}\right)f_j(x)dx = \frac{1}{n}\sum_{r=1}^{k}\sum_{i=1}^{n_r}\frac{Y_{ir} - m_n(X_{ir})}{\sigma_j(X_{ir})/\sqrt{n_j}}\frac{f_j(X_{ir})}{f_{mix}(X_{ir})} + O_p(nh)^{-1}$$

จากบทแทรกที่ 2

$$\int\left(\frac{\hat{m}_j(x) - m_{jn}(x)}{\sigma_j(x)/\sqrt{n_j}}\right)f_j(x)dx = \frac{1}{n_j}\sum_{i=1}^{n_j}\frac{Y_{ij} - m_{jn}(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}} + O_p(n_jh)^{-1}$$ ดังนั้น

$$\hat{F}_{\varepsilon j0}(y) - \hat{F}_{\varepsilon j}(y) = f_{\varepsilon j}(y)\left[\frac{1}{n}\sum_{r=1}^{k}\sum_{i=1}^{n_j}\frac{Y_{ir} - m_{rn}(X_{ir})}{\sigma_j(X_{ir})/\sqrt{n_j}}\frac{f_j(X_{ir})}{f_{mix}(X_{ir})}\right]$$

$$- f_{\varepsilon j}(y)\left[\frac{1}{n_j}\sum_{i=1}^{n_j}\frac{Y_{ij} - m_{jn}(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}}\right]$$

$$+ n^{-1/2}f_{\varepsilon j}(y)E\left(\frac{(S(X_{ij}) - s_j(X_{ij}))}{\sigma_j(X_{ij})/\sqrt{n_j}}\right) + o_p(n)^{-1/2} \qquad (25)$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

จากสมการ (25) จัดรูปใหม่จะได้ว่า

$$\hat{F}_{\varepsilon j0}(y) - \hat{F}_{\varepsilon j}(y) = f_{\varepsilon j}(y)\sum_{r=1}^{k} p_r \left\{ \frac{1}{n_r}\sum_{i=1}^{n_r} \frac{Y_{ir} - m_{rn}(X_{ir})}{\sigma_j(X_{ir})/\sqrt{n_j}} \left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j)}{p_j} \right) \right\}$$

$$+ n^{-1/2} f_{\varepsilon j}(y) E\left(\frac{S(X_{ij}) - s_j(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}} \right) + o_p(n)^{-1/2}$$

ทฤษฎีบทที่ 6 สมมติข้อตกลง 1 – 14 เป็นจริง ภายใต้สมมติฐานทางเลือก สำหรับ $j = 1, \ldots, k$ k – Dimensions ของ $\hat{U}(y) = (\hat{U}_1(y), \ldots, \hat{U}_j(y))^t$ Converges Weakly เข้าสู่ $U(y) + E(y)$ เมื่อ $U(y) = (U_1(y), \ldots, U_k(y))^t$ และ

$U_j(y) = f_{\varepsilon j}(y) p_j \sum_{r=1}^{k}\sum_{i=1}^{n_r} \frac{Y_{ir} - m(X_{ir})}{\sigma_j(X_{ir})} \left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j)}{p_j} \right)$ คือ ตัวแปรสุ่มปกติที่มีค่าเฉลี่ยเท่ากับ 0 และมีโครงสร้างความแปรปรวนร่วมคือ

$$Cov(U_j(y), U_{j'}(y)) = f_{\varepsilon j}(y) f_{\varepsilon j'}(y) p_j p_{j'} \sum_{r=1}^{k} E\left[\frac{\sigma_r^2(X_{ir})}{\sigma_j(X_{ir})\sigma_{j'}(X_{ir})} \left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j)}{p_j} \right) \right.$$

$$\left. \left(\frac{f_{j'}(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j')}{p_{j'}} \right) \right]$$

$$E(y) = (p_1^{1/2} f_{\varepsilon 1}(y) e_1, \ldots, p_k^{1/2} f_{\varepsilon k}(y) e_k)$$

$$e_j = E\left[\frac{S(X_{ij}) - s_j(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}} \right]$$

$$S(u) = \sum_{j=1}^{k} p_j \frac{f_j(u)}{f_{mix}(u)} s_j(u)$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

พิสูจน์ จาก Cramer – Wold Device (Serfling, 1980) แสดงให้เห็นว่า Weak Convergence ของ Multidimensional เท่ากับ Weak Convergence ของ Linear Combination ของแต่ละส่วนประกอบ พิจารณา Linear Combination ของ $\hat{Z}(y)$

$$\hat{Z}(y) = \sum_{j=1}^{k} b_j n_j^{1/2} (\hat{F}_{\varepsilon j0}(y) - \hat{F}_{\varepsilon j}(y)) \quad \text{จากทฤษฎีบทที่ 5}$$

$$\hat{F}_{\varepsilon j0}(y) - \hat{F}_{\varepsilon j}(y) = f_{\varepsilon j}(y)\left[\frac{1}{n}\sum_{r=1}^{k}\sum_{i=1}^{n_r}\frac{Y_{ir} - m_{rn}(X_{ir})}{\sigma_j(X_{ir})/\sqrt{n_j}}\frac{f_j(X_{ir})}{f_{mix}(X_{ir})}\right]$$

$$- f_{\varepsilon j}(y)\left[\frac{1}{n_j}\sum_{i=1}^{n_j}\frac{Y_{ij} - m_{jn}(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}}\right]$$

$$+ n^{-1/2} f_{\varepsilon j}(y) E\left(\frac{S(X_{ij}) - s_j(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}}\right) + o_p(n)^{-1/2} \quad \text{ดังนั้น}$$

$$\hat{Z}(y) = \sum_{j=1}^{k} b_j n_j^{1/2}\left[\frac{1}{n} f_{\varepsilon j}(y \sum_{r=1}^{k}\sum_{i=1}^{n_r}\frac{Y_{ir} - m_{rn}(X_{ir})}{\sigma_r(X_{ir})/\sqrt{n_j}}\frac{\sigma_r(X_{ir})}{\sigma_j(X_{ir})}\frac{f_j(X_{ir})}{f_{mix}(X_{ir})}\right.$$

$$\left. - \frac{1}{n_j} f_{\varepsilon j}(y)\frac{1}{n_j}\sum_{i=1}^{n_j}\frac{Y_{ij} - m_{jn}(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}}\right]$$

$$+ \sum_{j=1}^{k} b_j n_j^{1/2} n^{-1/2} f_{\varepsilon j}(y) E\left(\frac{S(X_{ij}) - s_j(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}}\right) + o_p(n)^{-1/2}$$

จัดรูปใหม่จะได้

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

$$\hat{Z}(y) = \sum_{r=1}^{k}\left[\frac{1}{n}\sum_{i=1}^{n_r}\frac{Y_{ir} - m_{rn}(X_{ir})}{\sigma_r(X_{ir})}\left[\sum_{j=1}^{k} b_j n_j f_{\varepsilon j}(y)\frac{\sigma_r(X_{ir})}{\sigma_j(X_{ir})}\frac{f_j(X_{ir})}{f_{mix}(X_{ir})} - \frac{b_r n_r}{p_r} f_{\varepsilon r}(y)\right]\right.$$

$$\left. + b_r n_r^{1/2} n^{-1/2} f_{\varepsilon r}(y) E\left(\frac{S(X_{ir}) - s_r(X_{ir})}{\sigma_r(X_{ir})/\sqrt{n_r}}\right)\right] + o_p(n)^{-1/2} \quad \text{ดังนั้น}$$

$$\hat{Z}(y) = \sum_{r=1}^{k}\left[\sum_{i=1}^{n_r}\frac{Y_{ir} - m_{rn}(X_{ir})}{\sigma_r(X_{ir})}\left[\sum_{j=1}^{k} b_j p_j f_{\varepsilon j}(y)\frac{\sigma_r(X_{ir})}{\sigma_j(X_{ir})}\frac{f_j(X_{ir})}{f_{mix}(X_{ir})} - b_r f_{\varepsilon r}(y)\right]\right.$$

$$\left. + b_r n_r^{1/2} n^{-1/2} f_{\varepsilon r}(y) E\left(\frac{S(X_{ir}) - s_r(X_{ir})}{\sigma_r(X_{ir})/\sqrt{n_r}}\right)\right] + o_p(n)^{-1/2}$$

$$= \sum_{r=1}^{k}\left[\sum_{i=1}^{n_r}\Psi_r(X_{ir}, Y_{ir}, y)\right] + o_p(n)^{-1/2}$$

เมื่อ

$$\Psi_r(u,v,y) = \frac{v - m(u)}{\sigma_r(u)}\left(\sum_{j=1}^{k} b_j p_j f_{\varepsilon j}(y)\frac{\sigma_r(u)}{\sigma_j(u)}\frac{f_j(u)}{f_{mix}(u)} - b_r f_{\varepsilon r}(y)\right)$$

$$+ b_r p_r^{1/2} f_{\varepsilon r}(y) E\left(\frac{S(u) - s_r(u)}{\sigma_r(u)/\sqrt{n_r}}\right) + o_p(n)^{-1/2}$$

สำหรับ $r = 1, \ldots, k$

$$\hat{Z}_r(y) = \sum_{i=1}^{n_r}\Psi_r(X_{ir}, Y_{ir}, y)$$

พิจารณา Class ของ Function

$$F_r = \left\{(u,v) \to \Psi_r(u,v,y)\,, -\infty < y < \infty\right\} \quad F \text{ คือ } F \text{ - Indexed Process}$$

ดังนั้น $\hat{Z}_r(y)$ คือ F - Indexed Process (Van der Vaart and Wellner, 1996)

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

โดยทั่วไปสำหรับ class ของฟังก์ชัน G_1 และ G_2 ได้นิยาม

$G_1 + G_2 = \{g_1 + g_2,\ g_1 \in G_1,\ g_2 \in G_2\}$ ดังนั้นสัญลักษณ์ของ class F_r สามารถเขียนได้ในรูปของ

$$F_r = \sum_{j=1}^{k+1} F_{rj} \quad \text{สำหรับ } j = 1, \ldots, k$$

ดังนั้น $F_{rj} = \left\{ (u,v) \to b_j p_j f_{\varepsilon j}(y) \frac{\sigma_r(u)}{\sigma_j(u)} \frac{f_j(u)}{f_{mix}(u)} \frac{v - m(u)}{\sigma_r(u)},\ -\infty < y < \infty \right\}$

$$F_{r,k+1} = \left\{ (u,v) \to - b_r f_{\varepsilon r}(y) \frac{v - m(u)}{\sigma_r(u)} + b_r p_r^{1/2} f_{\varepsilon r}(y) E\left(\frac{S(u) - s_r(u)}{\sigma_r(u)/\sqrt{n_r}} \right) \right\}$$

กำหนดให้ M อยู่ในลักษณะของ $\sup_{y,\, j = 1, \ldots, k} \left| f_{\varepsilon j}(y) \right| < M$

ดังนั้น $N_{[\ ]}(\delta, F_{rj}, L_2(P)) \le 2M\delta^{-1}$ ถ้า $\delta < 2M$ และ

$N_{[\ ]}(\delta, F_{rj}, L_2(P)) = 1$ ถ้า $\delta > 2M$

เมื่อ $N_{[\ ]}$ คือ Bracketing Number

P คือ Probability Measure ที่สอดคล้องกับ Joint Distribution ของ (X_r, Y_r)

$L_2(P)$ คือ L_2 - norm (Van der Vaart and Wellner, 1996)

นอกจากนั้นโดยทฤษฎีบท 2.10.16 ของ Van der Vaart and Wellner (1996) พบว่า

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

$$N_{[\ ]}(\delta, F_r, L_2(P)) \ \le \ \prod_{j=1}^{k+1} N_{[\ ]}(\delta, F_{rj}, L(p)) \qquad (26)$$

Take log สมการ (26) จะได้ว่า

$$\log N_{[\ ]}(\delta, F_r, L_2(P)) \ \le \ \sum_{j=1}^{k} \log N_{[\ ]}(\delta, F_{rj}, L_2(P)) \ \text{และ}$$

$$\int_0^{\infty} \sqrt{\log N_{[\ ]}(\delta, F_r, L_2(P))}\, d\delta \ \le \ \sum_{j=1}^{k+1} \int_0^{2M} \sqrt{\log N_{[\ ]}(\delta, F_{rj}, L_2(P))}\ d\delta \qquad (27)$$

จาก (27) สามารถสรุปได้ว่า $\int_0^{\infty} \sqrt{\log N_{[\ ]}(\delta, F_r, L_2(P))}\, d\delta$ finite และจากทฤษฎีบท 2.5.6 ใน Van der Vaart and Wellner (1996) Class ของ Function คือ Donsker ดังนั้น Weak Convergence ของ $\hat{Z}_r(y)$ จะลู่เข้าสู่ $Z_r(y)$ ซึ่งคือ Gaussian Process ที่มีค่าเฉลี่ยเท่ากับ 0 โครงสร้างของความแปรปรวน

$$\mathrm{Cov}(Z_r(y), Z_r(y')) \ = \ E\left[\Psi(X_r, Y_r, y)\Psi(X_r, Y_r, y')\right]$$

เช่นเดียวกับการพิสูจน์ในทฤษฎีบทที่ 3 เนื่องจาก $\hat{U}_r(y)$ คือ $\hat{Z}_r(y)$ เมื่อ $b_j = 1$

จาก $\hat{Z}_r(y) \ = \ \sum_{i=1}^{n_r} \dfrac{Y_{ir} - m_{rn}(X_{ir})}{\sigma_r(X_{ir})} \left[\sum_{j=1}^{k} b_j p_j f_{\varepsilon j}(y) \dfrac{\sigma_r(X_{ir})}{\sigma_j(X_{ir})} \dfrac{f_j(X_{ir})}{f_{mix}(X_{ir})} \ - \ b_r f_{\varepsilon r}(y) \right]$

$$+ \ b_r p_r^{1/2} f_{\varepsilon r}(y) E\left(\frac{S(X_{ir}) - s_r(X_{ir})}{\sigma_r(X_{ir}) / \sqrt{n_r}} \right) \ + o_p(n)^{-1/2}$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

$$\text{ดังนั้น } \hat{U}_r(y) = \sum_{i=1}^{n_r} \frac{Y_{ir} - m_{rm}(X_{ir})}{\sigma_r(X_{ir})} \left[\sum_{j=1}^{k} p_j f_{\varepsilon j}(y) \frac{\sigma_r(X_{ir})}{\sigma_j(X_{ir})} \frac{f_j(X_{ir})}{f_{mix}(X_{ir})} - f_{\varepsilon r}(y) \right]$$

$$+ p_r^{1/2} f_{\varepsilon r}(y) E\left(\frac{(S(X_{ir}) - s_r(X_{ir}))}{\sigma_r(X_{ir})/\sqrt{n_r}} \right) + o_p(1)$$

$$\hat{U}_r(y) = f_{\varepsilon j}(y) p_j \sum_{r=1}^{k} \sum_{i=1}^{n_r} \frac{Y_r - m_{rn}(X_r)}{\sigma_r(X_r)} \left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j)}{p_j} \right)$$

$$+ p_r^{1/2} f_{\varepsilon r}(y) E\left(\frac{(S(X_{ir}) - s_r(X_{ir}))}{\sigma_r(X_{ir})/\sqrt{n_r}} \right) + o_p(1) \qquad (28)$$

จากสมการ (28) จะได้ว่า $\hat{U}(y)$ จะ Converges Weakly เข้าสู่ $U(y) + E(y)$

$U(y) = (U_1(y), \ldots, U_k(y))^t$ เมื่อ

$U_j(y) = f_{\varepsilon j}(y) p_j \sum_{r=1}^{k} \sum_{i=1}^{n_r} \frac{Y_{ir} - m(X_{ir})}{\sigma_j(X_{ir})} \left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j)}{p_j} \right)$ คือ ตัวแปรสุ่มปกติ

ที่มีค่าเฉลี่ยเท่ากับ 0 และมีโครงสร้างความแปรปรวนร่วมคือ

$$Cov(U_j(y), U_{j'}(y)) = f_{\varepsilon j}(y) f_{\varepsilon j'}(y) p_j p_{j'} \sum_{r=1}^{k} E\left[\sum_{i=1}^{n_r} \frac{(Y_{ir} - m(X_{ir}))^2}{\sigma_j(X_{ir}) \sigma_{j'}(X_{ir})} \left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j)}{p_j} \right) \right.$$

$$\left. \left(\frac{f_{j'}(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j')}{p_{j'}} \right) \right]$$

เมื่อ $\sum_{i=1}^{n_r} (Y_{ir} - m(X_{ir}))^2 = \sigma_{ir}^2(X)$ ดังนั้น

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

$$\mathrm{Cov}(U_j(y),U_{j'}(y)) = f_{\varepsilon j}(y)f_{\varepsilon j'}(y)p_j p_{j'} \sum_{r=1}^{k} E\left[\frac{\sigma_r^2(X_{ir})}{\sigma_j(X_{ir})\sigma_{j'}(X_{ir})} \left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j)}{p_j} \right) \right.$$

$$\left. \left(\frac{f_{j'}(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j')}{p_{j'}} \right) \right]$$

$$E(y) = (p_1^{1/2} f_{\varepsilon 1}(y)e_1 \, , \ldots , p_k^{1/2} f_{\varepsilon k}(y)e_k)$$

$$e_j = E\left[\frac{S(X_{ij}) - s_j(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}} \right]$$

ทฤษฎีบทที่ 7.1 สมมติข้อตกลง 11– 14 เป็นจริง ภายใต้สมมติฐานทางเลือก สำหรับ $j = 1 \ldots , k$

$$U_{KS}^1 \xrightarrow{d} \sum_{j=1}^{k} \sup_y \left| U_j(y) + p_j^{1/2} e_j f_{\varepsilon j}(y) \right|$$

เมื่อ $U_{KS}^1 = \sum_{j=1}^{k} \sup_y \left| \hat{U}_j(y) \right|$

พิสูจน์ จากทฤษฎีบทที่ 6 และ Continuous Mapping Theorem จะได้ว่า

$$U_{KS}^1 = \sum_{j=1}^{k} \sup_y \left| \hat{U}_j(y) \right| \xrightarrow{d} \sum_{j=1}^{k} \sup_y \left| U_j(y) + p_j^{1/2} e_j f_{\varepsilon j}(y) \right|,$$

$$e_j = E\left[\frac{S(X_{ij}) - s_j(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}} \right]$$

ทฤษฎีบทที่ 7.2 สมมติข้อตกลง 11– 14 เป็นจริง ภายใต้สมมติฐานทางเลือก สำหรับ $j = 1 \ldots , k$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

$$U^1_{KU} \xrightarrow{d} \sum_{j=1}^{k}\left[\sup_y\left|U_j(y) + p_j^{1/2}e_j f_{\varepsilon j}(y)\right| - \inf_y\left|U_j(y) + p_j^{1/2}e_j f_{\varepsilon j}(y)\right|\right]$$

เมื่อ $U^1_{ku} = \sum_{i=1}^{k}\left[\sup_y\left|\hat{U}_j y\right| - \inf_y\left|\hat{U}_j y\right|\right]$

จากทฤษฎีบท 6 และ Continuous Mapping Theorem จะได้ว่า

$$U^1_{KU} \xrightarrow{d} \sum_{j=1}^{k}\left[\sup_y\left|U_j(y) + p_j^{1/2}e_j f_{\varepsilon j}(y)\right| - \inf_y\left|U_j(y) + p_j^{1/2}e_j f_{\varepsilon j}(y)\right|\right]$$

$$e_j = E\left[\frac{S(X_{ij}) - s_j(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}}\right]$$

ทฤษฎีบทที่ 7.3 สมมติข้อตกลง 11 – 14 เป็นจริง ภายใต้สมมติฐานทางเลือก สำหรับ $j = 1 \ldots, k$

$$U^2_{KS} \xrightarrow{d} \sup_y\left|U(y) + e(y)\right|$$

เมื่อ $U^2_{KS} = \sup_y\left|\hat{U}(y)\right|$

$$U(y) = \sum_{j=1}^{k} p_j^{1/2}U_j(y)$$

$$e(y) = \sum_{j=1}^{k} p_j^{1/2}f_{\varepsilon j}(y)e_j$$

พิสูจน์ เช่นเดียวกับการพิสูจน์ทฤษฎีบทที่ 4.3 จากทฤษฎีบทที่ 6 และ Continuous Mapping Theorem จะได้ว่า

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

$$U_{KS}^2 = \sup_y \left|\hat{U}(y)\right| \xrightarrow{d} \sup_y \left|U(y) + e(y)\right| \quad \text{เมื่อ} \quad U(y) = \sum_{j=1}^{k} p_j^{1/2} U_j(y)$$

$$e(y) = \sum_{j=1}^{k} p_j^{1/2} f_{\varepsilon j}(y) e_j$$

ทฤษฎีบทที่ 7.4 สมมติข้อตกลง 11– 14 เป็นจริง ภายใต้สมมติฐานทางเลือก สำหรับ $j = 1 \ldots, k$

$$U_{KU}^2 \xrightarrow{d} \left[\sup_y \left|U(y) + e(y)\right| - \inf_y \left|U(y) + e(y)\right| \right]$$

เมื่อ $U_{ku}^2 = \left[\sup_y \left|\hat{U}(y)\right| - \inf_y \left|\hat{U}(y)\right| \right]$

$$U(y) = \sum_{j=1}^{k} p_j^{1/2} U_j(y)$$

$$e(y) = \sum_{j=1}^{k} p_j^{1/2} f_{\varepsilon j}(y) e_j$$

พิสูจน์ เช่นเดียวกับการพิสูจน์ทฤษฎีบทที่ 4.4 จากทฤษฎีบท 6 และ Continuous Mapping Theorem จะได้ว่า

$$U_{ku}^2 = \left[\sup_y \left|\hat{U}(y)\right| - \inf_y \left|\hat{U}(y)\right| \right] \xrightarrow{d} \left[\sup_y \left|U(y) + E(y)\right| - \inf_y \left|U(y) + E(y)\right| \right]$$

$$e(y) = \sum_{j=1}^{k} p_j^{1/2} f_{\varepsilon j}(y) e_j$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

2. การจำลองแบบ (Simulation)

2.1 การทดสอบการเท่ากันของฟังก์ชันการถดถอย 2 ฟังก์ชัน ดำเนินการดังต่อไปนี้

2.1.1 กำหนดฟังก์ชันการถดถอยใน 4 ลักษณะคือ

1) ฟังก์ชันคงที่ $Y = b$

2) ฟังก์ชันเชิงเส้นตรง $Y = a(x) + b$

3) ฟังก์ชันเอ็กโพเนนเชียล $Y = \exp(ax) + b$

4) ฟังก์ชันตรีโกณมิติ $Y = \sin(a\pi x) + b$

ฟังก์ชันการถดถอยเมื่อสมมติฐานหลักเป็นจริง อยู่ในรูปของ

1) ฟังก์ชันคงที่ $m_1(x) = m_2(x)$ เมื่อ $b_1 = b_2 = 2$

2) ฟังก์ชันเชิงเส้นตรง $m_1(x) = m_2(x)$ เมื่อ $a_1 = a_2 = 2$, $b_1 = b_2 = 0$

3) ฟังก์ชันเอ็กโพเนนเชียล $m_1(x) = m_2(x)$ เมื่อ $a_1 = a_2 = 2$, $b_1 = b_2 = 0$

4) ฟังก์ชันตรีโกณมิติ $m_1(x) = m_2(x)$ เมื่อ $a_1 = a_2 = 2$, $b_1 = b_2 = 0$

ฟังก์ชันการถดถอยเมื่อสมมติฐานทางเลือกเป็นจริง ได้แก่ฟังก์ชันการการถดถอยในรูปของ $m_2(x) = m_1(x) + n^{-1/2}s(x)$ โดยศึกษาใน 3 ลักษณะคือ

1) $S_1(x) = 1$ (Constant Shift)

2) $S_2(x) = 1 + x$ (Affine Shift)

3) $S_3(x) = 1 + x + x^2$ (Quadratic Shift)

2.1.2 ส่วนเบี่ยงเบนมาตรฐานในตัวแบบการถดถอยอยู่ในลักษณะของ $\sigma_1(x) = 0.25$ และ $\sigma_2(x) = 0.50$

2.1.3 กำหนดให้ความคลาดเคลื่อนของแต่ละฟังก์ชันการถดถอยมีการแจกแจงใน 2 ลักษณะดังต่อไปนี้

1) $\varepsilon_1 \sim N(0,1)$, $\varepsilon_2 \sim N(0,1)$

2) $\varepsilon_1 \sim$ Exponential (1) - 1, $\varepsilon_2 \sim$ Exponential (1) - 1

2.1.4 กำหนดให้ตัวแปรอิสระ X_1 และ X_2 มีการแจกแจงแบบยูนิฟอร์ม [0 , 1]

$X_1 \sim U(0,1)$, $X_2 \sim U(0,1)$

2.1.5 ในการประมาณฟังก์ชันการถดถอยกำหนดให้ Kernel Function คือ Epanechnikov โดย $K(u) = 0.75(1 - u^2)I(|u| < 1)$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

2.1.6 พารามิเตอร์ปรับให้เรียบ h กำหนดให้มีค่าดังนี้ $h_j = cn_j^{-3/10}$ เมื่อ c มีค่าเท่ากับ 1 n_j คือ ขนาดของตัวอย่างของข้อมูลชุดที่ j

2.1.7 กำหนดให้ขนาดของตัวอย่างมีลักษณะดังต่อไปนี้

1) $(n_1, n_2) = (20,20)$

2) $(n_1, n_2) = (20,50)$

3) $(n_1, n_2) = (50,50)$

4) $(n_1, n_2) = (50,100)$

5) $(n_1, n_2) = (100,100)$

2.2 การทดสอบการเท่ากันของ 3 ฟังก์ชันการถดถอย ดำเนินการดังต่อไปนี้

2.2.1 ฟังก์ชันการถดถอยเมื่อสมมติฐานหลักเป็นจริงอยู่ในรูปของ

1) ฟังก์ชันคงที่ $m_1(x) = m_2(x) = m_3(x)$ เมื่อ $b_1 = b_2 = b_3 = 2$

2) ฟังก์ชันเชิงเส้นตรง $m_1(x) = m_2(x) = m_3(x)$ เมื่อ $a_1 = a_2 = a_3 = 2,\ b_1 = b_2 = b_3 = 0,$

3) ฟังก์ชันเอ็กโพเนนเชียล $m_1(x) = m_2(x) = m_3(x)$ เมื่อ $a_1 = a_2 = a_3 = 2,\ b_1 = b_2 = b_3 = 0$

4) ฟังก์ชันตรีโกณมิติ $m_1(x) = m_2(x) = m_3(x)$ เมื่อ $a_1 = a_2 = a_3 = 2,\ b_1 = b_2 = b_3 = 0$

ฟังก์ชันการถดถอยเมื่อสมมติฐานทางเลือกเป็นจริงได้แก่ฟังก์ชันการการถดถอยในรูปของ $m_2(x) = m_1(x) + n^{-1/2}s(x) = m_3(x)$ โดยศึกษาใน 3 ลักษณะคือ

1) $S_1(x) = 1$ (Constant Shift)

2) $S_2(x) = 1 + x$ (Affine Shift)

3) $S_3(x) = 1 + x + x^2$ (Quadratic Shift)

2.2.2 ส่วนเบี่ยงเบนมาตรฐานในตัวแบบการถดถอยอยู่ในลักษณะของ $\sigma_1(x) = 0.25,\ \sigma_2(x) = 0.50, \sigma_3(x) = 0.75$

2.2.3 กำหนดให้ความคลาดเคลื่อนของแต่ละฟังก์ชันการถดถอยมีการแจกแจงใน 2 ลักษณะดังต่อไปนี้

1) $\varepsilon_1 \sim N(0,1)$, $\varepsilon_2 \sim N(0,1)$, $\varepsilon_3 \sim N(0,1)$

2) $\varepsilon_1 \sim$ Exponential (1) - 1, $\varepsilon_2 \sim$ Exponential (1) - 1,

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

$$\varepsilon_3 \sim \text{Exponential (1)} - 1$$

2.2.4 กำหนดให้ตัวแปรอิสระ X_1 , X_2 และ X_3 มีการแจกแจงแบบยูนิฟอร์ม [0 , 1]

$$X_1 \sim U[0,1],\ X_2 \sim U[0,1]\ ,\ X_3 \sim U[0,1]$$

2.2.5 ในการประมาณฟังก์ชันการถดถอยกำหนดให้ Kernel Function คือ Epanechnikov โดย $K(u) = 0.75(1 - u^2)I(|u| < 1)$

2.2.6 พารามิเตอร์ปรับให้เรียบ กำหนดให้มีค่าดังนี้ $h_j = cn_j^{-3/10}$ เมื่อ c มีค่าเท่ากับ 1, n_j คือ ขนาดของตัวอย่างของข้อมูลชุดที่ j

2.2.7 กำหนดให้ขนาดของตัวอย่างมีลักษณะดังต่อไปนี้

1) $(n_1 , n_2 , n_3) = (20,20,20)$

2) $(n_1 , n_2 , n_3) = (20,50,20)$

3) $(n_1 , n_2 , n_3) = (50,50,50)$

4) $(n_1 , n_2 , n_3) = (50,100,50)$

5) $(n_1 , n_2 , n_3) = (100,100,100)$

2.3 การคำนวณค่าวิกฤต (Critical Value) ของสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ใช้หลักการของ Bootstrap Mechanism โดยกำหนดให้ Bootstrap Replication B = 200 และทำการทดลองทั้งสิ้น 1,000 ครั้ง โดย Bootstrap Mechanism มีขั้นตอนในการดำเนินการดังต่อไปนี้

2.3.1 กำหนดให้ Bootstrap Replication b= 1, … , B (B = 200) สำหรับ $i = 1, ... ,n_j$, $j = 1, ... ,k$ ให้ $\varepsilon^*_{ij,b}$ เป็นตัวอย่างสุ่มที่มีฟังก์ชันการแจกแจงเป็น $(1 - h_j^2)^{1/2}V_j + h_jZ$ V_j คือ ฟังก์ชันการแจกแจงของ ε_j, $Z \sim N(0,1)$, $h_j = cn_j^{-3/10}$ และ c มีค่าเท่ากับ 1

2.3.2 สำหรับ $i = 1, ... ,n_j$, $j = 1, ... ,k$ กำหนด New Response ภายใต้สมมติฐานหลักคือ

$$Y^*_{ij,b} = \hat{m}_j(X_{ij}) + \hat{\sigma}_j(X_{ij})\varepsilon^*_{ij,b}$$

2.3.3 คำนวณสถิติทดสอบ $U^{1*}_{KS}, U^{1*}_{KU}, U^{2*}_{KS}, U^{2*}_{KU}$ โดย $U^{1*}_{KS}, U^{1*}_{KU}, U^{2*}_{KS}, U^{2*}_{KU}$ เป็นสถิติทดสอบที่ได้มาจาก Bootstrap Sample $\{(X_{ij}, Y^*_{ij,b}), i = 1, ..., n_j, j = 1, ... , k\}$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

2.3.4 จากขั้นตอน 2.3.3 หลังจากทำจนครบจำนวน Bootstrap Sample (B = 200) ให้ $U^{1*}_{KS\,,\,(b)}$ เป็น Order Statistic ของ $U^{1*}_{KS\,,\,1}, \ldots, U^{1*}_{KS\,,\,B}$ ซึ่งสถิติทดสอบตัวอื่น ๆ $U^{1*}_{KU,\,(b)}$, $U^{2*}_{KS,\,(b)}$, $U^{2*}_{KU\,\,(b)}$ ก็ดำเนินการเช่นเดียวกัน ดังนั้น $U^{1*}_{KS,\,(1-\alpha)B}$, $U^{1*}_{KS,\,(1-\alpha)B}$, $U^{1*}_{KS,\,(1-\alpha)B}$, $U^{1*}_{KS,\,(1-\alpha)B}$ จะเป็นค่าประมาณของ $(1-\alpha)$ - Quantiles ของการแจกแจงของสถิติทดสอบ U^{1}_{KS}, U^{1}_{KU}, U^{2}_{KS}, U^{2}_{KU}

2.3.5 จากขั้นตอน 2.3.1 ถึง 2.3.4 ของสถิติทดสอบ U^{1}_{KS}, U^{1}_{KU}, U^{2}_{KS}, U^{2}_{KU} ทำซ้ำ 1,000 ครั้ง และคำนวณหาสัดส่วนของการปฏิเสธภายใต้สมมติฐานหลัก (ความคลาดเคลื่อนประเภทที่ 1) และสัดส่วนของการปฏิเสธภายใต้สมมติฐานทางเลือก (อำนาจการทดสอบ) โดยรายละเอียดของความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 แสดงในภาคผนวก ก

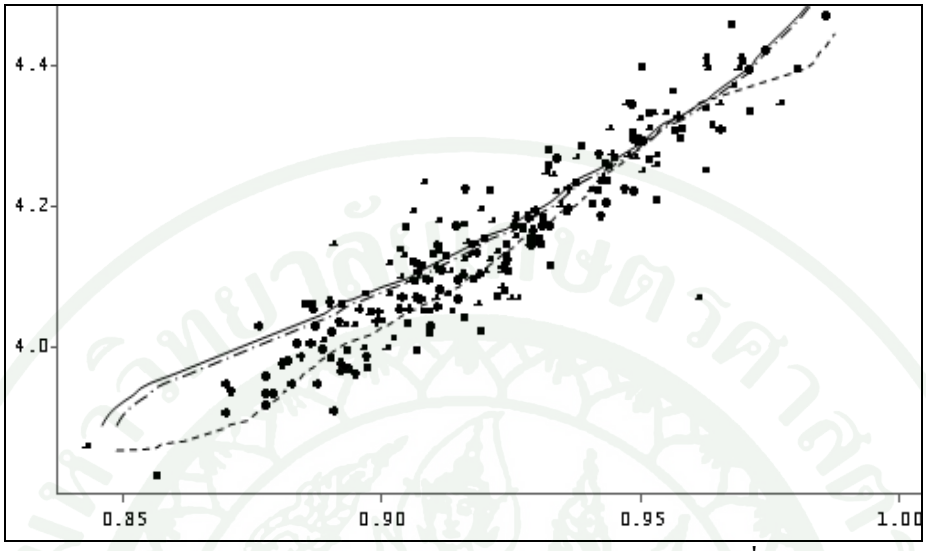
3. การประยุกต์สถิติทดสอบการเท่ากันของฟังก์ชันการถดถอย k ฟังก์ชัน (k = 3) กับข้อมูลจริง

เป็นการประยุกต์ใช้สถิติสำหรับการทดสอบการเท่ากันของฟังก์ชันการถดถอย k ฟังก์ชัน (k = 3) กับข้อมูลจริงซึ่งได้แก่ ค่าใช้จ่ายเฉลี่ยต่อครัวเรือนและรายได้เฉลี่ยต่อครัวเรือนจำแนกตามจังหวัดจำนวน 76 จังหวัด โดยกำหนดให้ Y = log (ค่าใช้จ่ายเฉลี่ยต่อครัวเรือน) X = log (รายได้เฉลี่ยต่อครัวเรือน) และทำการเปรียบเทียบเส้นโค้งการถดถอยของข้อมูล 3 กลุ่ม ซึ่งได้แก่ข้อมูลในปี พ.ศ. 2547 (วงกลม) ปี พ.ศ. 2549 (สี่เหลี่ยม) และปี พ.ศ. 2550 (สามเหลี่ยม) โดยภาพที่ 4 แสดงแผนภาพการกระจายระหว่าง log (รายได้เฉลี่ยต่อครัวเรือน) และ log (ค่าใช้จ่ายเฉลี่ยต่อครัวเรือน) ซึ่งประมาณเส้นโค้งการถดถอยระหว่าง log (รายได้เฉลี่ยต่อครัวเรือน) และ log (ค่าใช้จ่ายเฉลี่ยต่อครัวเรือน) ปี พ.ศ. 2547 ด้วยเส้นประ ประมาณเส้นโค้งการถดถอยระหว่าง log (รายได้เฉลี่ยต่อครัวเรือน) และ log (ค่าใช้จ่ายเฉลี่ยต่อครัวเรือน) ปี พ.ศ. 2549 ด้วยเส้นประจุดประ ประมาณเส้นโค้งการถดถอยระหว่าง log (รายได้เฉลี่ยต่อครัวเรือน) และ log (ค่าใช้จ่ายเฉลี่ยต่อครัวเรือน) ปี พ.ศ. 2550 ด้วยเส้นทึบ นอกจากนั้นในส่วนของตัวแปรอิสระ X = log (รายได้เฉลี่ยต่อครัวเรือน) ได้มีการแปลงให้อยู่ในช่วงของ [0,1]

การทดสอบการเท่ากันของ 3 ฟังก์ชันการถดถอย สำหรับสถิติทดสอบ U^{1}_{KS} และ U^{1}_{KU} ดำเนินการโดยกำหนดให้พารามิเตอร์ปรับให้เรียบ มีค่าเท่ากับ 0.10, 0.15, 0.20, 0.25, 0.30, 0.35 และ 0.40 หลังจากนั้นหาค่า P – value ของ สถิติทดสอบ U^{1}_{KS} และ U^{1}_{KU} จาก 1,000 Bootstrap Replications

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

log (ค่าใช้จ่ายเฉลี่ยต่อครัวเรือน)



Log (รายได้เฉลี่ยต่อครัวเรือน)

ภาพที่ 4 แผนภาพการกระจายระหว่าง log (รายได้เฉลี่ยต่อครัวเรือน) และ log (ค่าใช้จ่ายเฉลี่ยต่อครัวเรือน) ปี พ.ศ. 2547 (วงกลม) ปี พ.ศ. 2549 (สี่เหลี่ยม) และปี พ.ศ. 2550 (สามเหลี่ยม) โดยประมาณเส้นโค้งการถดถอยระหว่าง log (รายได้เฉลี่ยต่อครัวเรือน) และ log (ค่าใช้จ่ายเฉลี่ยต่อครัวเรือน) ปี พ.ศ. 2547 ด้วยเส้นประ ปี พ.ศ. 2549 ด้วยเส้นประจุดประ และ ปี พ.ศ. 2550 ด้วยเส้นทึบ

ที่มา: การสำรวจภาวะเศรษฐกิจและสังคมของครัวเรือน พ.ศ. 2545 2547 2549 และ 2550 สำนักงานสถิติแห่งชาติ (2551)

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

ผลและวิจารณ์

ผล

ผลการวิเคราะห์ข้อมูลแบ่งเป็น 3 ส่วนดังนี้

ส่วนที่ 1 การทดสอบการเท่ากันของฟังก์ชันการถดถอย 2 ฟังก์ชัน

กำหนดให้ฟังก์ชันการถดถอยเมื่อสมมติฐานหลักเป็นจริงอยู่ในรูปของ

1) ฟังก์ชันคงที่ $m_1(x) = m_2(x)$ เมื่อ $b_1 = b_2 = 2$
2) ฟังก์ชันเชิงเส้นตรง $m_1(x) = m_2(x)$ เมื่อ $a_1 = a_2 = 2$, $b_1 = b_2 = 0$
3) ฟังก์ชันเอ็กโพเนนเชียล $m_1(x) = m_2(x)$ เมื่อ $a_1 = a_2 = 2$, $b_1 = b_2 = 0$
4) ฟังก์ชันตรีโกณมิติ $m_1(x) = m_2(x)$ เมื่อ $a_1 = a_2 = 2$, $b_1 = b_2 = 0$

กำหนดฟังก์ชันการถดถอยเมื่อสมมติฐานทางเลือกเป็นจริงได้แก่ฟังก์ชันการการถดถอยในรูปของ $m_2(x) = m_1(x) + n^{-1/2}s(x)$ โดยศึกษาใน 3 ลักษณะคือ

1) $S_1(x) = 1$ (Constant Shift)
2) $S_2(x) = 1 + x$ (Affine Shift)
3) $S_3(x) = 1 + x + x^2$ (Quadratic Shift)

กำหนดให้ความคลาดเคลื่อนของแต่ละฟังก์ชันการถดถอยมีการแจกแจง ใน 2 ลักษณะดังต่อไปนี้

1) $\varepsilon_1 \sim N(0,1)$, $\varepsilon_2 \sim N(0,1)$
2) $\varepsilon_1 \sim$ Exponential (1) - 1, $\varepsilon_2 \sim$ Exponential (1) - 1

ผลการวิเคราะห์ข้อมูลแสดงดัง ตารางที่ 1 – 8

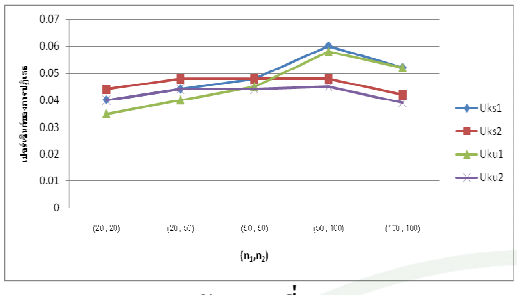
ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

ตารางที่ 1 เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานหลักของตัวแบบที่ 1) - 4) สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \sigma_2(x) = 0.50$ เมื่อการแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim N(0,1), \varepsilon_2 \sim N(0,1)$

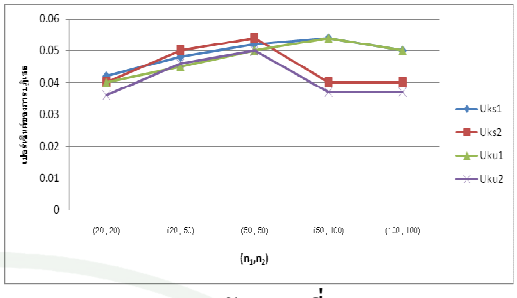
(n_1, n_2)	ตัวแบบที่	U^1_{KS} ระดับนัยสำคัญ		U^2_{KS} ระดับนัยสำคัญ		U^1_{KU} ระดับนัยสำคัญ		U^2_{KU} ระดับนัยสำคัญ	
		0.05	0.10	0.05	0.10	0.05	0.10	0.05	0.10
(20 , 20)	1)	0.040	0.074*	0.044	0.080*	0.035*	0.074*	0.040	0.077*
	2)	0.042	0.080*	0.040	0.078*	0.040	0.078*	0.036	0.075*
	3)	0.030*	0.074*	0.038	0.074*	0.028*	0.070*	0.034*	0.071*
	4)	0.034*	0.076*	0.040	0.080*	0.030*	0.074*	0.040	0.077*
(20 , 50)	1)	0.044	0.090	0.048	0.084	0.040	0.088	0.044	0.081*
	2)	0.048	0.092	0.050	0.086	0.045	0.090	0.046	0.084
	3)	0.034*	0.084	0.040	0.082*	0.032*	0.080*	0.036	0.080*
	4)	0.035	0.086	0.040	0.080*	0.032*	0.082*	0.036	0.078*
(50 , 50)	1)	0.048	0.080*	0.048	0.082*	0.045	0.080*	0.044	0.080*
	2)	0.052	0.090	0.054	0.094	0.050	0.088	0.050	0.092
	3)	0.054	0.084	0.052	0.090	0.052	0.080*	0.049	0.088
	4)	0.060	0.080*	0.058	0.098	0.058	0.082*	0.055	0.096
(50,100)	1)	0.060	0.114	0.048	0.084	0.058	0.112	0.045	0.082*
	2)	0.054	0.104	0.040	0.082*	0.054	0.100	0.037	0.080*
	3)	0.060	0.114	0.044	0.090	0.058	0.112	0.041	0.089
	4)	0.056	0.104	0.042	0.088	0.054	0.102	0.039	0.087
(100,100)	1)	0.052	0.100	0.042	0.089	0.052	0.100	0.039	0.087
	2)	0.050	0.100	0.040	0.087	0.050	0.100	0.037	0.085
	3)	0.052	0.102	0.040	0.086	0.052	0.100	0.038	0.084
	4)	0.054	0.106	0.042	0.088	0.052	0.102	0.040	0.085

หมายเหตุ * แทนความคลาดเคลื่อนประเภทที่ 1 อยู่นอกช่วงของการควบคุม

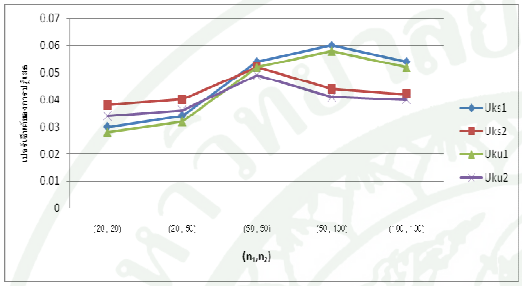
ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



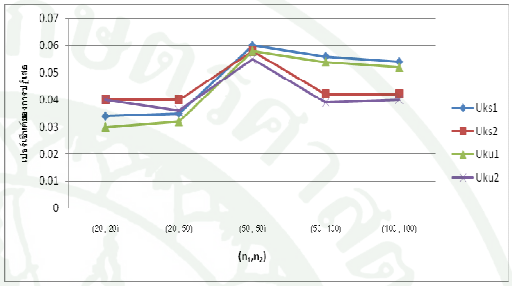
ตัวแบบที่ 1



ตัวแบบที่ 2

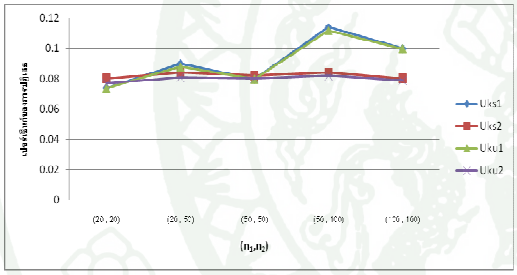


ตัวแบบที่ 3

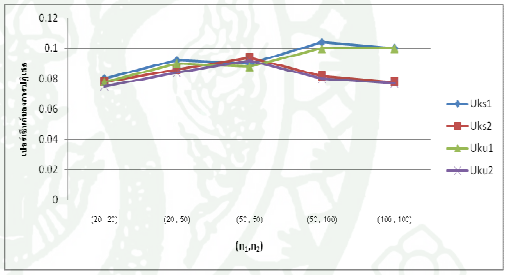


ตัวแบบที่ 4

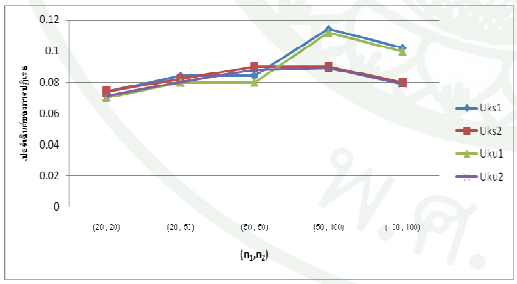
ระดับนัยสำคัญ 0.05



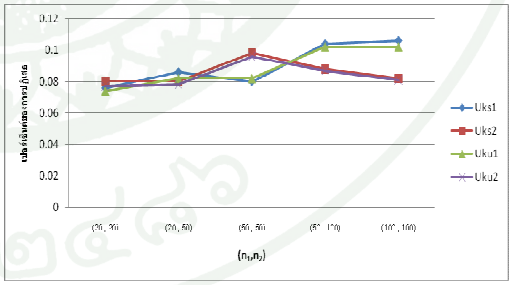
ตัวแบบที่ 1



ตัวแบบที่ 2



ตัวแบบที่ 3



ตัวแบบที่ 4

ระดับนัยสำคัญ 0.10

ภาพที่ 5 เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานหลักของตัวแบบที่ 1) - 4) สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \sigma_2(x) = 0.50$ การแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim N(0,1), \varepsilon_2 \sim N(0,1)$ เมื่อระดับนัยสำคัญเท่ากับ 0.05 และ 0.10

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

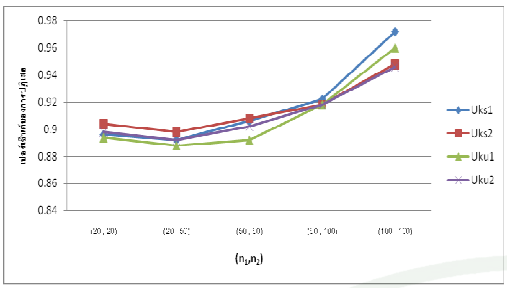
จากตารางที่ 1 และภาพที่ 5 พบว่าการควบคุมความคลาดเคลื่อนประเภทที่ 1 สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ จะทำได้ดีเมื่อตัวอย่างมีขนาดใหญ่ขึ้น โดยเมื่อกำหนดระดับนัยสำคัญเท่ากับ 0.05 ขนาดตัวอย่าง $(n_1 , n_2) = (100,100)$ ความคลาดเคลื่อนประเภทที่ 1 จะอยู่ภายในช่วง $[0.036 , 0.063]$ และเมื่อกำหนดระดับนัยสำคัญเท่ากับ 0.10 ขนาดตัวอย่าง $(n_1 , n_2) = (100,100)$ ความคลาดเคลื่อนประเภทที่ 1 จะอยู่ในช่วง $[0.084 , 0.115]$ ในทุกตัวแบบการถดถอย ซึ่งแสดงให้เห็นว่าสถิติทดสอบสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดีเมื่อตัวอย่างมีขนาดใหญ่

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

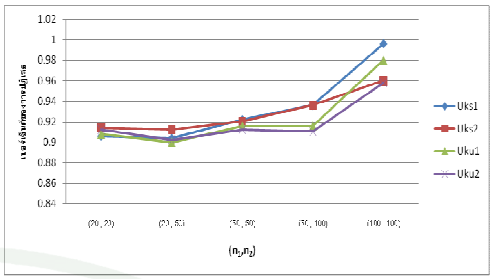
ตารางที่ 2 เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Constant Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25,\ \sigma_2(x) = 0.50$ เมื่อการแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim N(0,1), \varepsilon_2 \sim N(0,1)$

(n_1, n_2)	ตัวแบบที่	U^1_{KS}		U^2_{KS}		U^1_{KU}		U^2_{KU}	
		ระดับนัยสำคัญ		ระดับนัยสำคัญ		ระดับนัยสำคัญ		ระดับนัยสำคัญ	
		0.05	0.10	0.05	0.10	0.05	0.10	0.05	0.10
(20 , 20)	1)	0.896	0.920	0.904	0.924	0.894	0.910	0.898	0.914
	2)	0.906	0.926	0.914	0.930	0.908	0.920	0.912	0.920
	3)	0.882	0.922	0.902	0.926	0.888	0.912	0.892	0.916
	4)	0.380	0.442	0.396	0.452	0.376	0.432	0.380	0.436
(20 , 50)	1)	0.892	0.916	0.898	0.922	0.888	0.902	0.892	0.906
	2)	0.904	0.950	0.912	0.954	0.900	0.940	0.902	0.944
	3)	0.890	0.896	0.896	0.900	0.876	0.886	0.896	0.896
	4)	0.408	0.500	0.428	0.506	0.388	0.496	0.392	0.498
(50 , 50)	1)	0.906	0.926	0.908	0.924	0.892	0.912	0.902	0.916
	2)	0.922	0.960	0.920	0.958	0.916	0.946	0.912	0.942
	3)	0.914	0.952	0.908	0.942	0.898	0.938	0.898	0.932
	4)	0.392	0.520	0.392	0.510	0.332	0.500	0.336	0.504
(50,100)	1)	0.922	0.946	0.918	0.936	0.918	0.936	0.918	0.934
	2)	0.936	0.960	0.936	0.946	0.916	0.940	0.910	0.926
	3)	0.918	0.960	0.898	0.940	0.912	0.948	0.896	0.937
	4)	0.468	0.552	0.408	0.512	0.448	0.532	0.406	0.509
(100,100)	1)	0.972	0.980	0.948	0.958	0.960	0.970	0.946	0.955
	2)	0.996	0.996	0.960	0.940	0.980	0.980	0.958	0.937
	3)	0.966	0.976	0.926	0.936	0.956	0.966	0.924	0.933
	4)	0.546	0.640	0.444	0.526	0.538	0.596	0.442	0.523

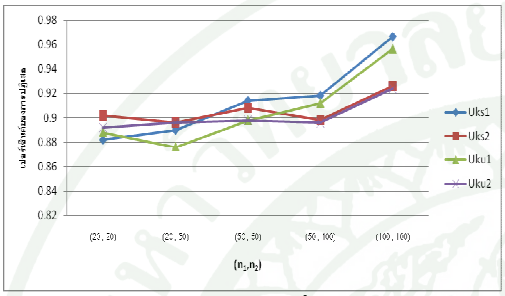
ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



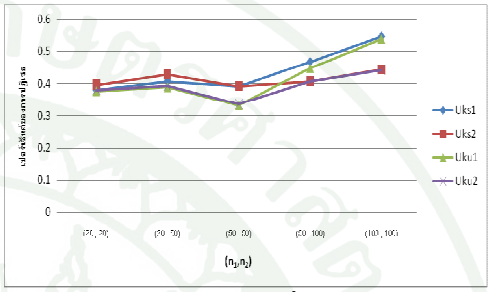
ตัวแบบที่ 1



ตัวแบบที่ 2

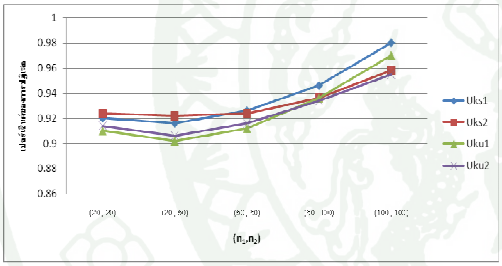


ตัวแบบที่ 3

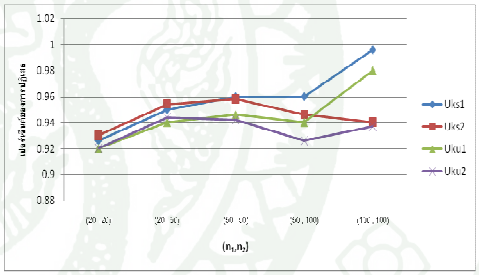


ตัวแบบที่ 4

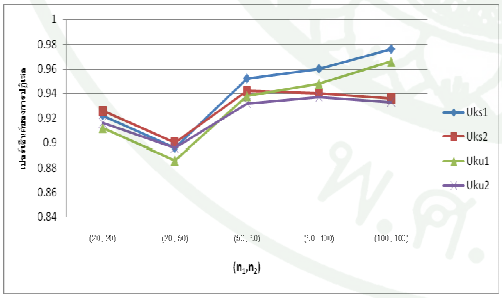
ระดับนัยสำคัญ 0.05



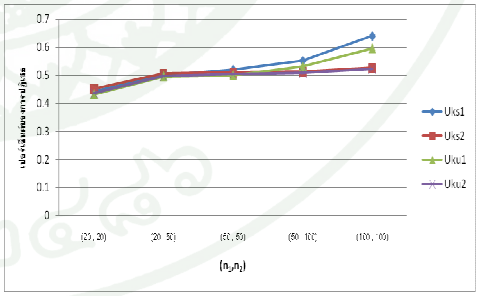
ตัวแบบที่ 1



ตัวแบบที่ 2



ตัวแบบที่ 3



ตัวแบบที่ 4

ระดับนัยสำคัญ 0.10

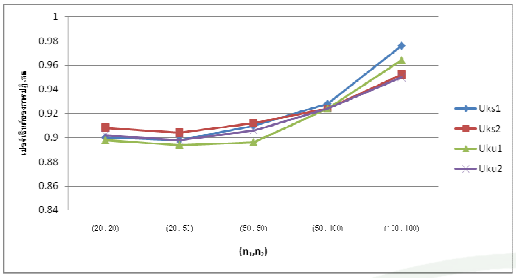
ภาพที่ 6 เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Constant Shift สำหรับสถิติทดสอบ U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU} ในกรณีที่ $\sigma_1(x) = 0.25$, $\sigma_2(x) = 0.50$ การแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim N(0,1)$, $\varepsilon_2 \sim N(0,1)$ เมื่อระดับนัยสำคัญเท่ากับ 0.05 และ 0.10

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

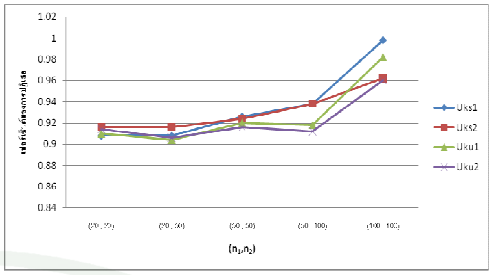
ตารางที่ 3 เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Affine Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25,\ \sigma_2(x) = 0.50$ เมื่อการแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim N(0,1), \varepsilon_2 \sim N(0,1)$

(n_1, n_2)	ตัวแบบที่	U^1_{KS}		U^2_{KS}		U^1_{KU}		U^2_{KU}	
		ระดับนัยสำคัญ		ระดับนัยสำคัญ		ระดับนัยสำคัญ		ระดับนัยสำคัญ	
		0.05	0.10	0.05	0.10	0.05	0.10	0.05	0.10
(20 , 20)	1)	0.900	0.920	0.908	0.924	0.898	0.910	0.902	0.914
	2)	0.908	0.928	0.916	0.932	0.910	0.922	0.914	0.922
	3)	0.886	0.926	0.906	0.930	0.892	0.916	0.896	0.920
	4)	0.382	0.446	0.398	0.456	0.378	0.436	0.382	0.440
(20 , 50)	1)	0.898	0.920	0.904	0.926	0.894	0.906	0.898	0.910
	2)	0.908	0.950	0.916	0.954	0.904	0.940	0.906	0.944
	3)	0.892	0.898	0.898	0.902	0.878	0.888	0.898	0.898
	4)	0.414	0.502	0.434	0.508	0.394	0.498	0.398	0.500
(50 , 50)	1)	0.910	0.928	0.912	0.926	0.896	0.914	0.906	0.918
	2)	0.926	0.962	0.924	0.960	0.920	0.948	0.916	0.944
	3)	0.920	0.956	0.914	0.946	0.904	0.942	0.904	0.936
	4)	0.396	0.520	0.396	0.510	0.336	0.500	0.340	0.504
(50,100)	1)	0.928	0.948	0.924	0.938	0.924	0.938	0.924	0.936
	2)	0.938	0.962	0.938	0.948	0.918	0.942	0.912	0.928
	3)	0.924	0.960	0.904	0.940	0.918	0.948	0.902	0.937
	4)	0.474	0.556	0.414	0.516	0.454	0.536	0.412	0.513
(100,100)	1)	0.976	0.980	0.952	0.958	0.964	0.970	0.950	0.955
	2)	0.998	0.998	0.962	0.942	0.982	0.982	0.960	0.939
	3)	0.968	0.978	0.928	0.938	0.958	0.968	0.926	0.935
	4)	0.546	0.642	0.444	0.528	0.538	0.598	0.442	0.525

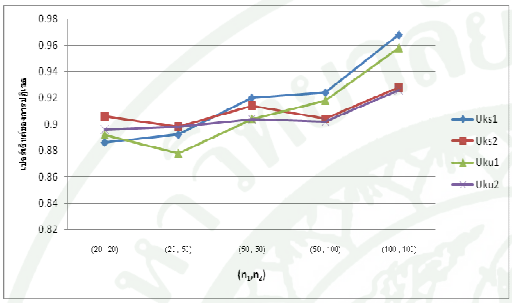
ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



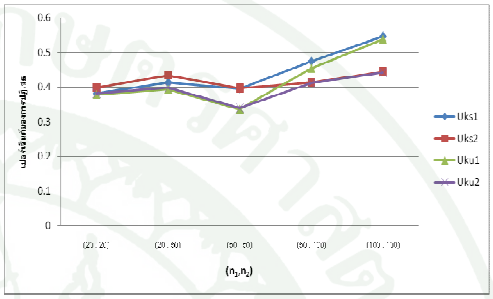
ตัวแบบที่ 1



ตัวแบบที่ 2

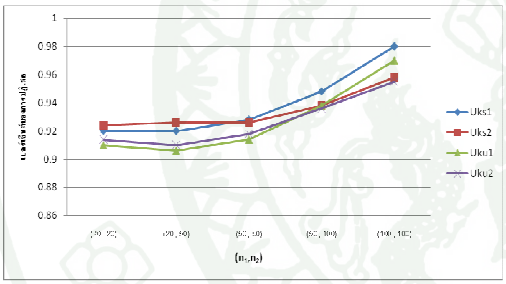


ตัวแบบที่ 3

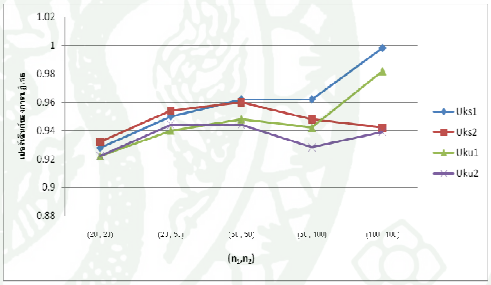


ตัวแบบที่ 4

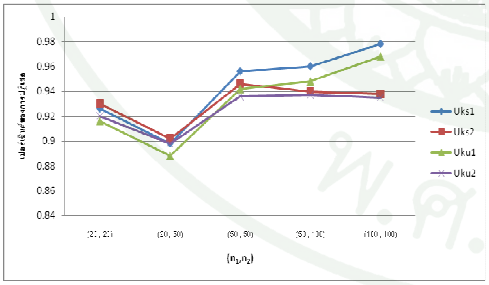
ระดับนัยสำคัญ 0.05



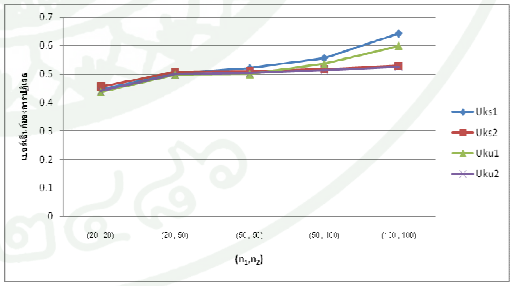
ตัวแบบที่ 1



ตัวแบบที่ 2



ตัวแบบที่ 3



ตัวแบบที่ 4

ระดับนัยสำคัญ 0.10

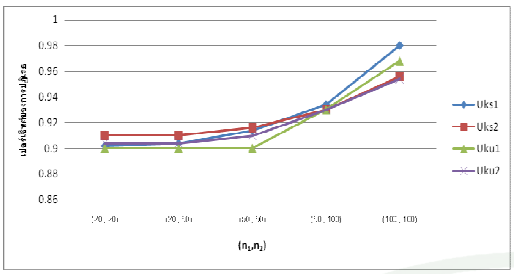
ภาพที่ 7 เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Affine Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25$, $\sigma_2(x) = 0.50$ การแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim N(0,1)$, $\varepsilon_2 \sim N(0,1)$ เมื่อระดับนัยสำคัญเท่ากับ 0.05 และ 0.10

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

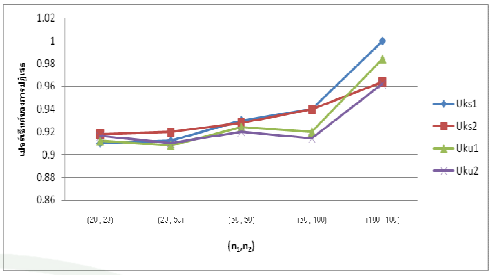
ตารางที่ 4 เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Quadratic Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25,\ \sigma_2(x) = 0.50$ เมื่อการแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim N(0,1), \varepsilon_2 \sim N(0,1)$

(n_1, n_2)	ตัวแบบที่	U^1_{KS}		U^2_{KS}		U^1_{KU}		U^2_{KU}	
		ระดับนัยสำคัญ		ระดับนัยสำคัญ		ระดับนัยสำคัญ		ระดับนัยสำคัญ	
		0.05	0.10	0.05	0.10	0.05	0.10	0.05	0.10
(20 , 20)	1)	0.902	0.920	0.910	0.924	0.900	0.910	0.904	0.914
	2)	0.910	0.930	0.918	0.934	0.912	0.924	0.916	0.924
	3)	0.890	0.930	0.910	0.934	0.896	0.920	0.900	0.924
	4)	0.384	0.450	0.400	0.460	0.380	0.440	0.384	0.444
(20 , 50)	1)	0.904	0.924	0.910	0.930	0.900	0.910	0.904	0.914
	2)	0.912	0.950	0.920	0.954	0.908	0.940	0.910	0.944
	3)	0.894	0.900	0.900	0.904	0.880	0.890	0.900	0.900
	4)	0.420	0.504	0.440	0.510	0.400	0.500	0.404	0.502
(50 , 50)	1)	0.914	0.930	0.916	0.928	0.900	0.916	0.910	0.920
	2)	0.930	0.964	0.928	0.962	0.924	0.950	0.920	0.946
	3)	0.926	0.960	0.920	0.950	0.910	0.946	0.910	0.940
	4)	0.400	0.520	0.400	0.510	0.340	0.500	0.344	0.504
(50,100)	1)	0.934	0.950	0.930	0.940	0.930	0.940	0.930	0.938
	2)	0.940	0.964	0.940	0.950	0.920	0.944	0.914	0.930
	3)	0.930	0.960	0.910	0.940	0.924	0.948	0.908	0.937
	4)	0.480	0.560	0.420	0.520	0.460	0.540	0.418	0.517
(100,100)	1)	0.980	0.980	0.956	0.958	0.968	0.970	0.954	0.955
	2)	1.000	1.000	0.964	0.944	0.984	0.984	0.962	0.941
	3)	0.970	0.980	0.930	0.940	0.960	0.970	0.928	0.937
	4)	0.546	0.644	0.444	0.530	0.538	0.600	0.442	0.527

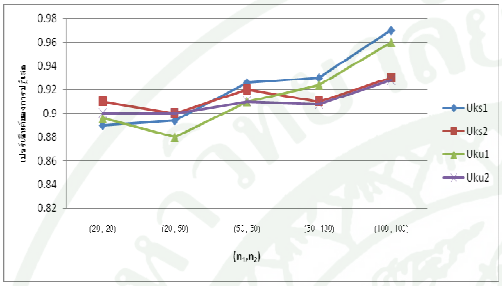
ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



ตัวแบบที่ 1



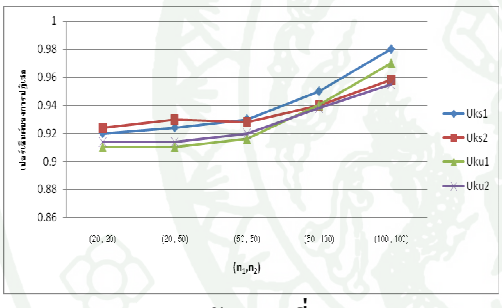
ตัวแบบที่ 2



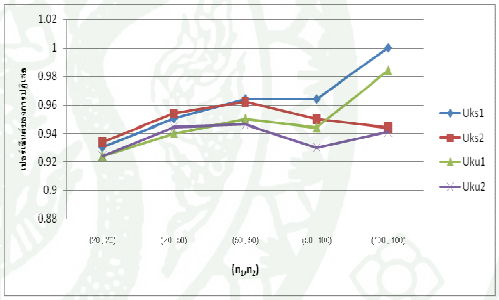
ตัวแบบที่ 3

ตัวแบบที่ 4

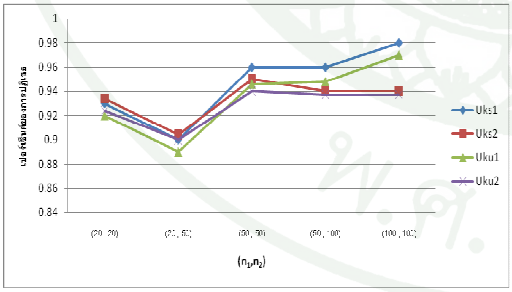
ระดับนัยสำคัญ 0.05



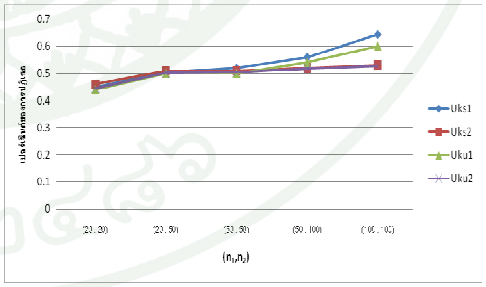
ตัวแบบที่ 1



ตัวแบบที่ 2



ตัวแบบที่ 3



ตัวแบบที่ 4

ระดับนัยสำคัญ 0.10

ภาพที่ 8 เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Quadratic Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \sigma_2(x) = 0.50$ การแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim N(0,1)$, $\varepsilon_2 \sim N(0,1)$ เมื่อระดับนัยสำคัญเท่ากับ 0.05 และ 0.10

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

จากตารางที่ 2 – 4 และภาพที่ 6 - 8 พบว่าที่ระดับนัยสำคัญเท่ากับ 0.05 และ 0.10 เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \sigma_2(x) = 0.50$ เมื่อการแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim N(0,1), \varepsilon_2 \sim N(0,1)$ จะมีค่าอำนาจการทดสอบสูงขึ้นเมื่อขนาดของตัวอย่างเพิ่มขึ้น โดยเฉพาะในตัวแบบที่ 1 ตัวแบบที่ 2 และตัวแบบที่ 3 ค่าอำนาจการทดสอบสูงขึ้นจนเกือบถึง 1 ในขณะที่ตัวแบบที่ 4 ซึ่งได้แก่ตัวแบบตรีโกณมิติจะให้ค่าอำนาจการทดสอบต่ำกว่าตัวแบบอื่น ๆ เมื่อพิจารณาที่การ Shift พบว่ากรณี Quadratic Shift ให้ค่าอำนาจการทดสอบสูงที่สุด รองลงมาได้แก่ กรณี Affine Shift และ Constant Shift ตามลำดับ

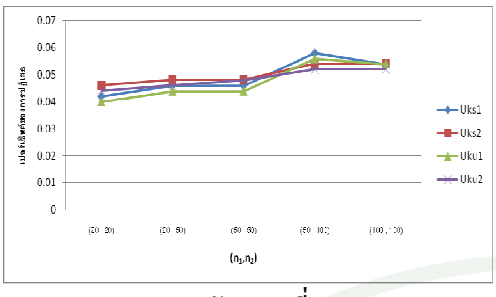
ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

ตารางที่ 5 เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานหลักของตัวแบบที่ 1) - 4) สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \sigma_2(x) = 0.50$ เมื่อการแจกแจงของความคลาดเคลื่อน ε_1 ~ Exponential (1) - 1, ε_2 ~ Exponential (1) - 1

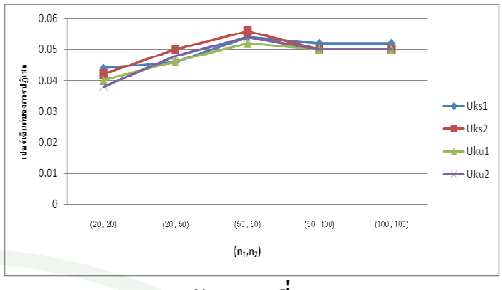
(n_1, n_2)	ตัวแบบที่	U^1_{KS}		U^2_{KS}		U^1_{KU}		U^2_{KU}	
		ระดับนัยสำคัญ		ระดับนัยสำคัญ		ระดับนัยสำคัญ		ระดับนัยสำคัญ	
		0.05	0.10	0.05	0.10	0.05	0.10	0.05	0.10
(20 , 20)	1)	0.042	0.078*	0.046	0.082*	0.040	0.074*	0.044	0.080*
	2)	0.044	0.084	0.042	0.078*	0.040	0.080*	0.038	0.076*
	3)	0.028*	0.076*	0.040	0.076*	0.028*	0.074*	0.040	0.074*
	4)	0.034*	0.080*	0.040	0.082*	0.030*	0.076*	0.040	0.080*
(20 , 50)	1)	0.046	0.092	0.048	0.090	0.044	0.090	0.046	0.090
	2)	0.046	0.092	0.050	0.086	0.046	0.090	0.048	0.084
	3)	0.036	0.086	0.042	0.084	0.034*	0.080*	0.042	0.084
	4)	0.038	0.088	0.042	0.082*	0.034*	0.084	0.040	0.080*
(50 , 50)	1)	0.046	0.080*	0.048	0.084	0.044	0.080*	0.048	0.080*
	2)	0.054	0.094*	0.056	0.096	0.052	0.090	0.054	0.094
	3)	0.054	0.086	0.054	0.090	0.054	0.084	0.054	0.090
	4)	0.058	0.080*	0.056	0.098	0.054	0.080	0.054	0.094
(50,100)	1)	0.058	0.110	0.054	0.108	0.056	0.104	0.052	0.104
	2)	0.052	0.102	0.050	0.102	0.050	0.100	0.050	0.102
	3)	0.058	0.112	0.056	0.114	0.054	0.110	0.054	0.110
	4)	0.056	0.104	0.058	0.108	0.054	0.102	0.054	0.104
(100,100)	1)	0.054	0.100	0.054	0.102	0.054	0.098	0.052	0.100
	2)	0.052	0.100	0.050	0.102	0.050	0.100	0.050	0.100
	3)	0.052	0.104	0.054	0.108	0.052	0.102	0.052	0.104
	4)	0.054	0.108	0.054	0.110	0.054	0.106	0.054	0.106

หมายเหตุ * แทนความคลาดเคลื่อนประเภทที่ 1 อยู่นอกช่วงของการควบคุม

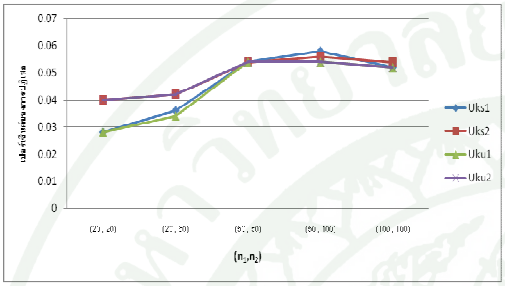
ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



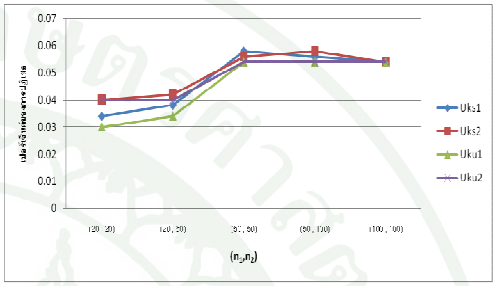
ตัวแบบที่ 1



ตัวแบบที่ 2

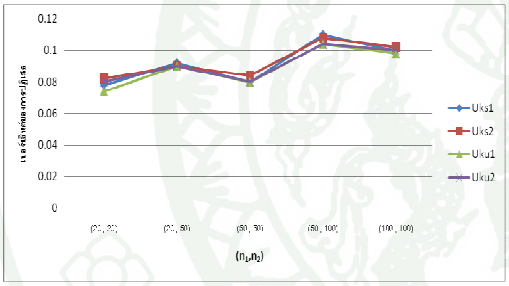


ตัวแบบที่ 3

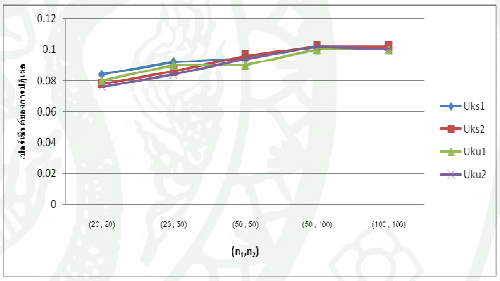


ตัวแบบที่ 4

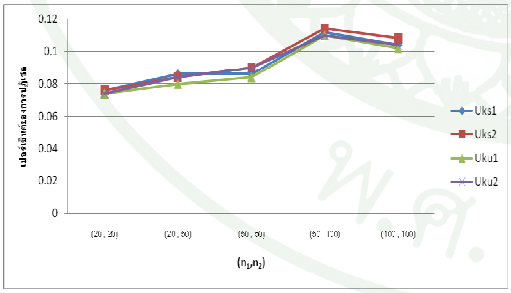
ระดับนัยสำคัญ 0.05



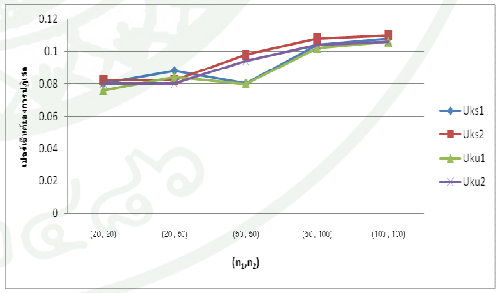
ตัวแบบที่ 1



ตัวแบบที่ 2



ตัวแบบที่ 3



ตัวแบบที่ 4

ระดับนัยสำคัญ 0.10

ภาพที่ 9 เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานหลักของตัวแบบที่ 1) - 4) สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \sigma_2(x) = 0.50$ การแจกแจงของความคลาดเคลื่อน ε_1 ~ Exponential (1) - 1, ε_2 ~ Exponential (1) - 1 เมื่อระดับนัยสำคัญเท่ากับ 0.05 และ 0.10

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

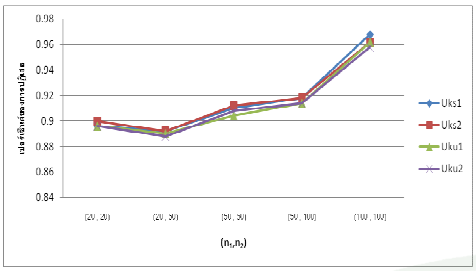
จากตารางที่ 5 และภาพที่ 9 พบว่าการควบคุมความคลาดเคลื่อนประเภทที่ 1 สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ จะทำได้ดีเมื่อตัวอย่างมีขนาดใหญ่ขึ้น โดยเมื่อกำหนดระดับนัยสำคัญเท่ากับ 0.05 ขนาดตัวอย่าง $(n_1 , n_2) = (50,100)$ และ $(n_1 , n_2) = (100,100)$ ความคลาดเคลื่อนประเภทที่ 1 จะอยู่ในช่วง [0.036 , 0.063] ในทุกตัวแบบการถดถอย และเมื่อกำหนดระดับนัยสำคัญเท่ากับ 0.10 ขนาดตัวอย่าง $(n_1 , n_2) = (50,100)$ และ $(n_1 , n_2) = (100,100)$ ความคลาดเคลื่อนประเภทที่ 1 จะอยู่ในช่วง [0.084 , 0.115] ในทุกตัวแบบการถดถอย ซึ่งแสดงให้เห็นว่าสถิติทดสอบสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดีเมื่อตัวอย่างมีขนาดใหญ่

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

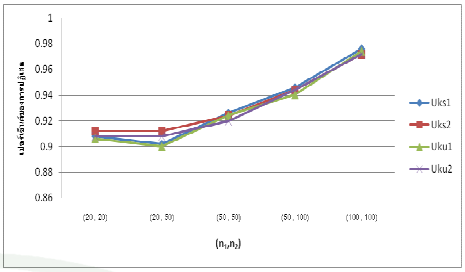
ตารางที่ 6 เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Constant Shift สำหรับสถิติทดสอบ U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU} ในกรณีที่ $\sigma_1(x) = 0.25$, $\sigma_2(x) = 0.50$ เมื่อการแจกแจงของความคลาดเคลื่อน ε_1 ~ Exponential (1) - 1, ε_2 ~ Exponential (1) - 1

(n_1, n_2)	ตัวแบบที่	U^1_{KS}		U^2_{KS}		U^1_{KU}		U^2_{KU}	
		ระดับนัยสำคัญ		ระดับนัยสำคัญ		ระดับนัยสำคัญ		ระดับนัยสำคัญ	
		0.05	0.10	0.05	0.10	0.05	0.10	0.05	0.10
(20 , 20)	1)	0.896	0.920	0.900	0.922	0.896	0.918	0.896	0.920
	2)	0.908	0.924	0.912	0.928	0.906	0.920	0.908	0.926
	3)	0.872	0.910	0.892	0.912	0.872	0.906	0.892	0.912
	4)	0.376	0.422	0.396	0.426	0.370	0.416	0.390	0.422
(20 , 50)	1)	0.892	0.912	0.892	0.914	0.890	0.908	0.888	0.914
	2)	0.902	0.940	0.912	0.946	0.900	0.934	0.908	0.940
	3)	0.896	0.906	0.906	0.924	0.896	0.912	0.902	0.920
	4)	0.388	0.496	0.398	0.506	0.388	0.496	0.394	0.504
(50 , 50)	1)	0.910	0.920	0.912	0.926	0.904	0.920	0.908	0.922
	2)	0.926	0.954	0.924	0.954	0.924	0.950	0.920	0.948
	3)	0.918	0.946	0.916	0.946	0.914	0.942	0.912	0.942
	4)	0.402	0.510	0.402	0.512	0.398	0.508	0.398	0.510
(50,100)	1)	0.918	0.942	0.918	0.940	0.914	0.932	0.914	0.936
	2)	0.946	0.950	0.944	0.948	0.940	0.946	0.944	0.946
	3)	0.940	0.944	0.930	0.940	0.934	0.944	0.926	0.938
	4)	0.438	0.516	0.432	0.512	0.432	0.512	0.428	0.510
(100,100)	1)	0.968	0.970	0.962	0.968	0.962	0.978	0.958	0.964
	2)	0.976	0.980	0.972	0.976	0.974	0.976	0.972	0.976
	3)	0.960	0.972	0.958	0.972	0.956	0.968	0.956	0.966
	4)	0.520	0.536	0.504	0.530	0.518	0.532	0.500	0.526

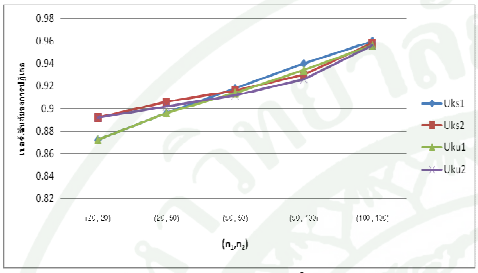
ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



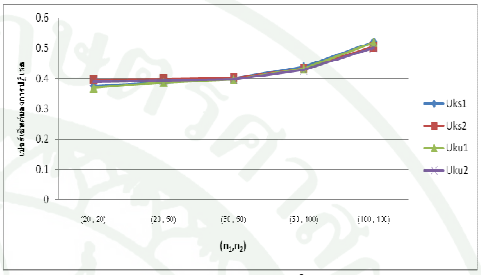
ตัวแบบที่ 1



ตัวแบบที่ 2

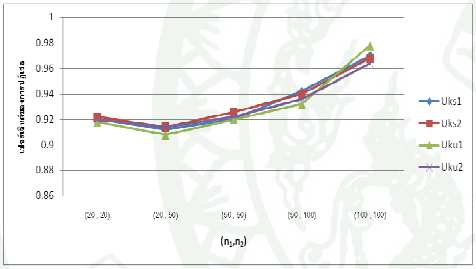


ตัวแบบที่ 3

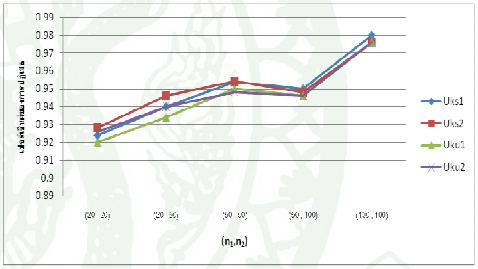


ตัวแบบที่ 4

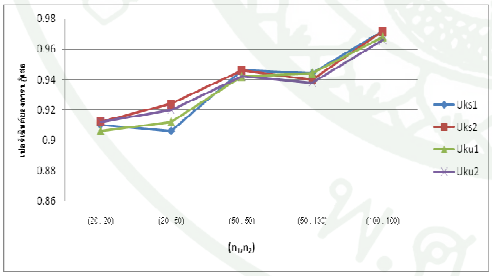
ระดับนัยสำคัญ 0.05



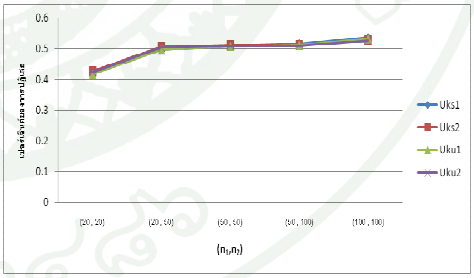
ตัวแบบที่ 1



ตัวแบบที่ 2



ตัวแบบที่ 3



ตัวแบบที่ 4

ระดับนัยสำคัญ 0.10

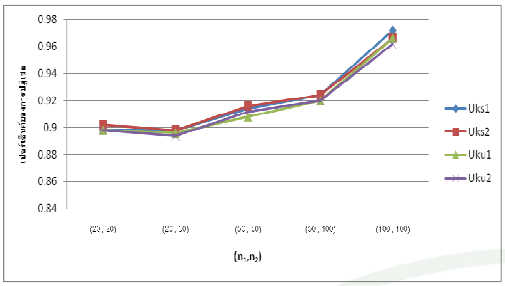
ภาพที่ 10 เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Constant Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \sigma_2(x) = 0.50$ การแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim$ Exponential (1) - 1, $\varepsilon_2 \sim$ Exponential (1) - 1 เมื่อระดับนัยสำคัญเท่ากับ 0.05 และ 0.10

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

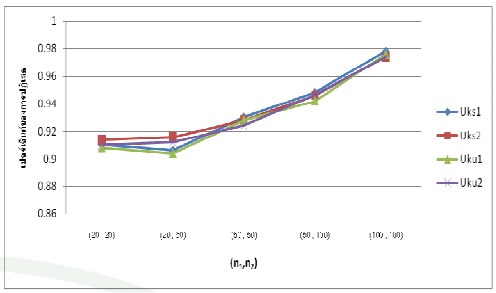
ตารางที่ 7 เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Affine Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25,\ \sigma_2(x) = 0.50$ เมื่อการแจกแจงของความคลาดเคลื่อน ε_1 ~ Exponential (1) - 1, ε_2 ~ Exponential (1) - 1

(n_1, n_2)	ตัวแบบที่	U^1_{KS}		U^2_{KS}		U^1_{KU}		U^2_{KU}	
		ระดับนัยสำคัญ		ระดับนัยสำคัญ		ระดับนัยสำคัญ		ระดับนัยสำคัญ	
		0.05	0.10	0.05	0.10	0.05	0.10	0.05	0.10
(20 , 20)	1)	0.898	0.920	0.902	0.922	0.898	0.918	0.898	0.920
	2)	0.910	0.926	0.914	0.930	0.908	0.922	0.910	0.928
	3)	0.876	0.914	0.896	0.916	0.876	0.910	0.896	0.916
	4)	0.378	0.426	0.398	0.430	0.372	0.420	0.378	0.426
(20 , 50)	1)	0.898	0.916	0.898	0.918	0.896	0.912	0.894	0.918
	2)	0.906	0.940	0.916	0.946	0.904	0.934	0.912	0.940
	3)	0.898	0.908	0.908	0.926	0.898	0.914	0.904	0.922
	4)	0.394	0.498	0.404	0.508	0.394	0.498	0.400	0.506
(50 , 50)	1)	0.914	0.922	0.916	0.928	0.908	0.922	0.912	0.924
	2)	0.930	0.956	0.928	0.956	0.928	0.952	0.924	0.950
	3)	0.924	0.950	0.922	0.950	0.920	0.946	0.918	0.946
	4)	0.406	0.510	0.406	0.512	0.402	0.508	0.402	0.510
(50,100)	1)	0.924	0.944	0.924	0.942	0.920	0.934	0.920	0.938
	2)	0.948	0.952	0.946	0.950	0.942	0.948	0.946	0.948
	3)	0.944	0.948	0.936	0.940	0.940	0.944	0.932	0.938
	4)	0.444	0.520	0.438	0.516	0.438	0.516	0.434	0.514
(100,100)	1)	0.972	0.970	0.966	0.968	0.966	0.978	0.962	0.964
	2)	0.978	0.982	0.974	0.978	0.976	0.978	0.974	0.978
	3)	0.962	0.974	0.960	0.974	0.958	0.970	0.958	0.968
	4)	0.520	0.538	0.504	0.532	0.518	0.534	0.500	0.528

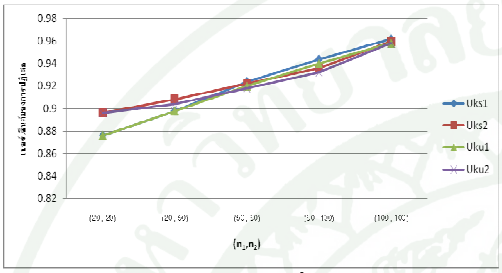
ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



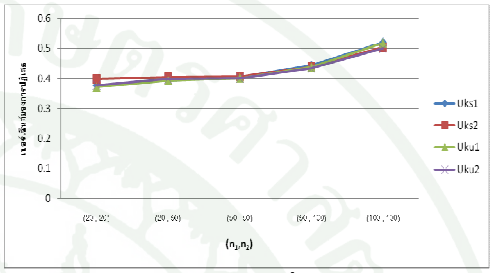
ตัวแบบที่ 1



ตัวแบบที่ 2

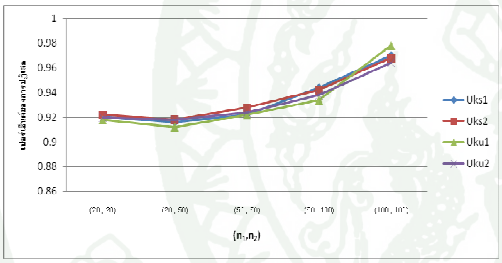


ตัวแบบที่ 3

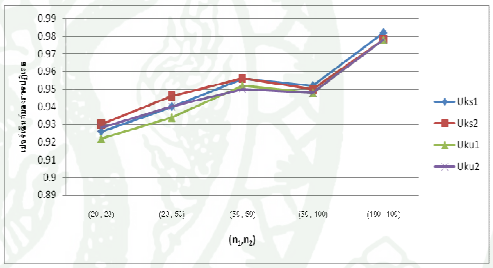


ตัวแบบที่ 4

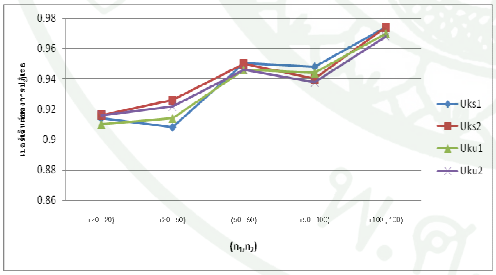
ระดับนัยสำคัญ 0.05



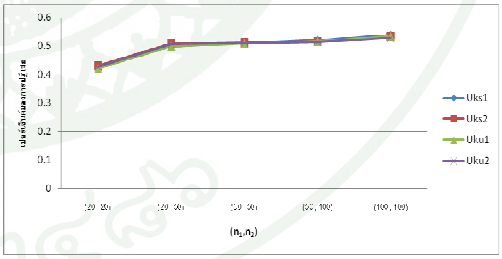
ตัวแบบที่ 1



ตัวแบบที่ 2



ตัวแบบที่ 3



ตัวแบบที่ 4

ระดับนัยสำคัญ 0.10

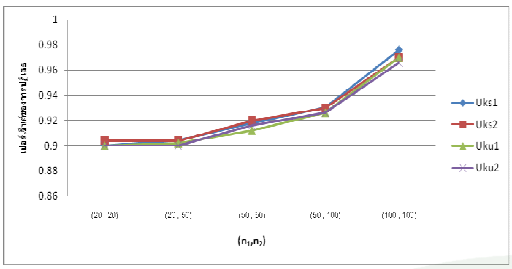
ภาพที่ 11 เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Affine Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25$, $\sigma_2(x) = 0.50$ การแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim$ Exponential (1) - 1, $\varepsilon_2 \sim$ Exponential (1) - 1 เมื่อระดับนัยสำคัญเท่ากับ 0.05 และ 0.10

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

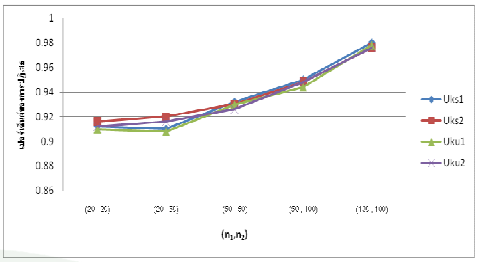
ตารางที่ 8 เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Quadratic Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25,\ \sigma_2(x) = 0.50$ เมื่อการแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim$ Exponential (1) - 1, $\varepsilon_2 \sim$ Exponential (1) - 1

(n_1, n_2)	ตัวแบบที่	U^1_{KS}		U^2_{KS}		U^1_{KU}		U^2_{KU}	
		ระดับนัยสำคัญ		ระดับนัยสำคัญ		ระดับนัยสำคัญ		ระดับนัยสำคัญ	
		0.05	0.10	0.05	0.10	0.05	0.10	0.05	0.10
(20 , 20)	1)	0.900	0.920	0.904	0.922	0.900	0.918	0.900	0.920
	2)	0.912	0.928	0.916	0.932	0.910	0.924	0.912	0.930
	3)	0.880	0.918	0.900	0.920	0.880	0.914	0.900	0.920
	4)	0.380	0.430	0.400	0.434	0.374	0.424	0.340	0.430
(20 , 50)	1)	0.904	0.920	0.904	0.922	0.902	0.916	0.900	0.922
	2)	0.910	0.940	0.920	0.946	0.908	0.934	0.916	0.940
	3)	0.900	0.910	0.910	0.928	0.900	0.916	0.906	0.924
	4)	0.400	0.500	0.410	0.510	0.400	0.500	0.406	0.508
(50 , 50)	1)	0.918	0.924	0.920	0.930	0.912	0.924	0.916	0.926
	2)	0.932	0.960	0.930	0.960	0.930	0.956	0.926	0.954
	3)	0.930	0.954	0.928	0.954	0.926	0.950	0.924	0.950
	4)	0.410	0.510	0.410	0.512	0.406	0.508	0.406	0.510
(50,100)	1)	0.930	0.946	0.930	0.944	0.926	0.936	0.926	0.940
	2)	0.950	0.954	0.948	0.952	0.944	0.950	0.948	0.950
	3)	0.948	0.950	0.938	0.946	0.942	0.950	0.934	0.944
	4)	0.450	0.524	0.444	0.520	0.444	0.520	0.440	0.518
(100,100)	1)	0.976	0.970	0.970	0.968	0.970	0.978	0.966	0.964
	2)	0.980	0.984	0.976	0.980	0.978	0.980	0.976	0.980
	3)	0.964	0.976	0.962	0.976	0.960	0.972	0.960	0.970
	4)	0.520	0.540	0.504	0.534	0.518	0.536	0.500	0.530

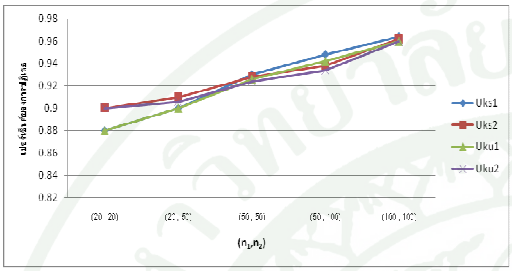
ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



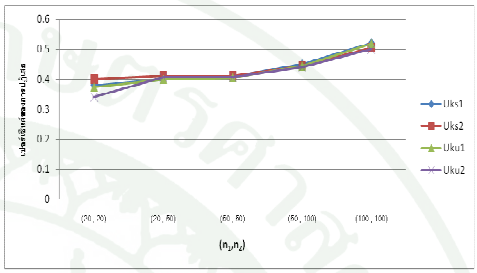
ตัวแบบที่ 1



ตัวแบบที่ 2

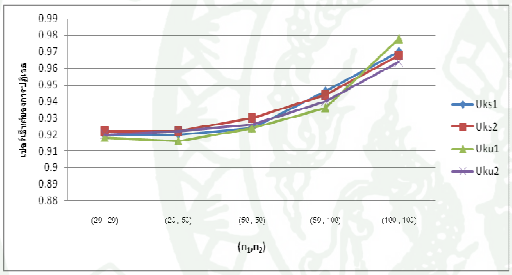


ตัวแบบที่ 3

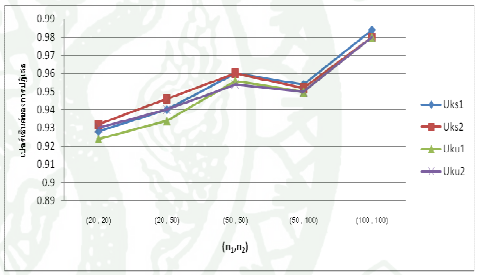


ตัวแบบที่ 4

ระดับนัยสำคัญ 0.05

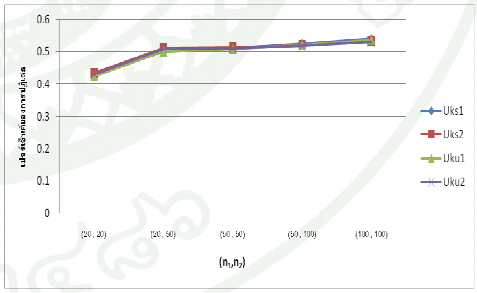


ตัวแบบที่ 1



ตัวแบบที่ 2

ตัวแบบที่ 1



ตัวแบบที่ 2

ระดับนัยสำคัญ 0.10

ภาพที่ 12 เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Quadratic Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \sigma_2(x) = 0.50$ เมื่อการแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim$ Exponential (1) - 1, $\varepsilon_2 \sim$ Exponential (1) - 1 เมื่อระดับนัยสำคัญเท่ากับ 0.05 และ 0.10

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

จากตารางที่ 6 – 8 และภาพที่ 10 - 12 พบว่าที่ระดับนัยสำคัญเท่ากับ 0.05 และ 0.10 เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือก ของตัวแบบที่ 1) - 4) สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \sigma_2(x) = 0.50$ เมื่อการแจกแจงของความคลาดเคลื่อน ε_1 ~ Exponential (1) -1 , ε_2 ~ Exponential (1) - 1 จะมีค่าอำนาจการทดสอบสูงขึ้นเมื่อขนาดของตัวอย่างเพิ่มขึ้น โดยเฉพาะในตัวแบบที่ 1 ตัวแบบที่ 2 และตัวแบบที่ 3 ค่าอำนาจการทดสอบสูงขึ้นจนเกือบถึง 1 ในขณะที่ตัวแบบที่ 4 ซึ่งได้แก่ ตัวแบบตรีโกณมิติจะให้ค่าอำนาจการทดสอบต่ำกว่าตัวแบบอื่น ๆ เมื่อพิจารณาที่การ Shift พบว่ากรณี Quadratic Shift ให้ค่าอำนาจการทดสอบสูงที่สุด รองลงมาได้แก่ กรณี Affine Shift และ Constant Shift ตามลำดับ

ส่วนที่ 2 การทดสอบการเท่ากันของฟังก์ชันการถดถอย 3 ฟังก์ชัน

กำหนดฟังก์ชันการถดถอยเมื่อสมมติฐานหลักเป็นจริงอยู่ในรูปของ

1) ฟังก์ชันคงที่ $m_1(x) = m_2(x) = m_3(x)$ เมื่อ $b_1 = b_2 = b_3 = 2$
2) ฟังก์ชันเชิงเส้นตรง $m_1(x) = m_2(x) = m_3(x)$ เมื่อ
   $a_1 = a_2 = a_3 = 2,\ b_1 = b_2 = b_3 = 0$
3) ฟังก์ชันเอ็กโพเนนเชียล $m_1(x) = m_2(x) = m_3(x)$ เมื่อ
   $a_1 = a_2 = a_3 = 2,\ b_1 = b_2 = b_3 = 0,$
4) ฟังก์ชันตรีโกณมิติ $m_1(x) = m_2(x) = m_3(x)$ เมื่อ
   $a_1 = a_2 = a_3 = 2,\ b_1 = b_2 = b_3 = 0$

กำหนดฟังก์ชันการถดถอยเมื่อสมมติฐานทางเลือกเป็นจริงได้แก่ฟังก์ชันการถดถอยในรูปของ $m_2(x) = m_1(x) + n^{-1/2}s(x) = m_3(x)$ โดยศึกษาใน 3 ลักษณะคือ

1) $S_1(x) = 1$ (Constant Shift)
2) $S_2(x) = 1 + x$ (Affine Shift)
3) $S_3(x) = 1 + x + x^2$ (Quadratic Shift)

กำหนดให้ความคลาดเคลื่อนของแต่ละฟังก์ชันการถดถอยมีการแจกแจงใน 2 ลักษณะต่อไปนี้ 1) ε_1 ~ N(0,1) , ε_2 ~ N(0,1), ε_3 ~ N(0,1)

2) ε_1 ~ Exponential (1) - 1, ε_2 ~ Exponential (1) - 1, ε_3 ~ Exponential(1) - 1

กำหนดให้ส่วนเบี่ยงเบนมาตรฐานในตัวแบบการถดถอยอยู่ในลักษณะของ $\sigma_1(x) = 0.25, \sigma_2(x) = 0.50, \sigma_3(x) = 0.75$ ผลการวิเคราะห์ข้อมูลแสดงดังตารางที่ 9 – 16

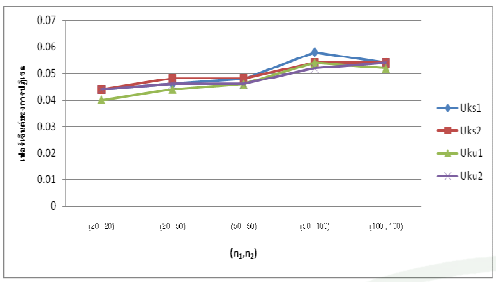
ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

ตารางที่ 9 เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานหลักของตัวแบบที่ 1) - 4) สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \sigma_2(x) = 0.50, \sigma_3(x) = 0.75$ เมื่อการแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim N(0,1)$, $\varepsilon_2 \sim N(0,1)$, $\varepsilon_3 \sim N(0,1)$

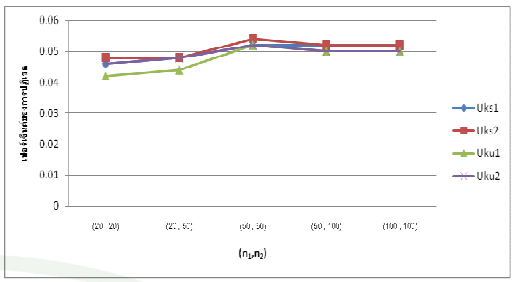
(n_1, n_2, n_3)	ตัวแบบที่	U^1_{KS}		U^2_{KS}		U^1_{KU}		U^2_{KU}	
		ระดับนัยสำคัญ		ระดับนัยสำคัญ		ระดับนัยสำคัญ		ระดับนัยสำคัญ	
		0.05	0.10	0.05	0.10	0.05	0.10	0.05	0.10
(20 , 20, 20)	1)	0.044	0.080*	0.044	0.082*	0.040	0.076*	0.044	0.080*
	2)	0.046	0.084	0.048	0.088	0.042	0.080*	0.046	0.084
	3)	0.040	0.080	0.044	0.080*	0.036	0.076*	0.042	0.080*
	4)	0.040	0.082*	0.040	0.080*	0.040	0.080*	0.040	0.080*
(20 , 50 , 20)	1)	0.046	0.094	0.048	0.096	0.044	0.090	0.046	0.092
	2)	0.048	0.090	0.048	0.092	0.044	0.090	0.048	0.090
	3)	0.042	0.088	0.046	0.090	0.040	0.084	0.044	0.088
	4)	0.042	0.086	0.046	0.092	0.040	0.084	0.046	0.090
(50 , 50 ,50)	1)	0.048	0.096	0.048	0.098	0.046	0.094	0.046	0.094
	2)	0.052	0.092	0.054	0.094	0.052	0.090	0.052	0.094
	3)	0.056	0.092	0.056	0.096	0.054	0.090	0.054	0.094
	4)	0.058	0.088	0.058	0.090	0.054	0.084	0.054	0.090
(50 ,100 ,50)	1)	0.058	0.116	0.054	0.112	0.054	0.112	0.052	0.110
	2)	0.052	0.102	0.052	0.110	0.050	0.102	0.050	0.110
	3)	0.060	0.112	0.056	0.108	0.056	0.110	0.054	0.104
	4)	0.058	0.090	0.054	0.102	0.054	0.084	0.052	0.100
(100,100,100)	1)	0.054	0.112	0.054	0.110	0.052	0.110	0.054	0.110
	2)	0.052	0.100	0.052	0.100	0.050	0.098	0.050	0.100
	3)	0.058	0.102	0.054	0.098	0.054	0.090	0.052	0.098
	4)	0.054	0.092	0.052	0.098	0.052	0.090	0.052	0.098

หมายเหตุ * แทนความคลาดเคลื่อนประเภทที่ 1 อยู่นอกช่วงของการควบคุม

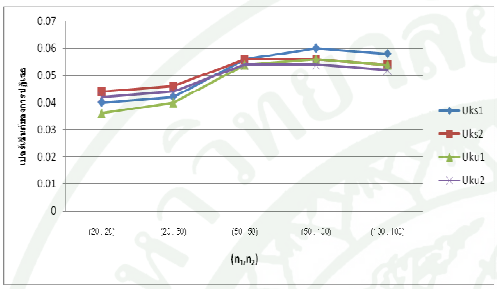
ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



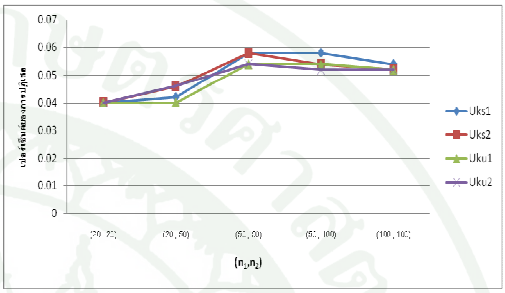
ตัวแบบที่ 1



ตัวแบบที่ 2

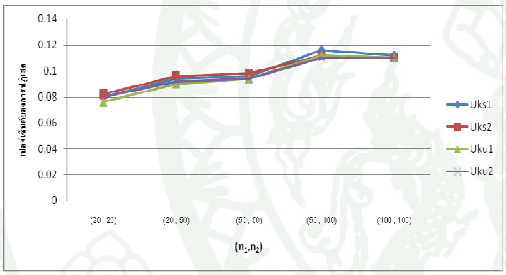


ตัวแบบที่ 3

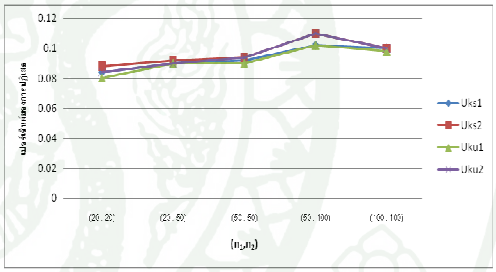


ตัวแบบที่ 4

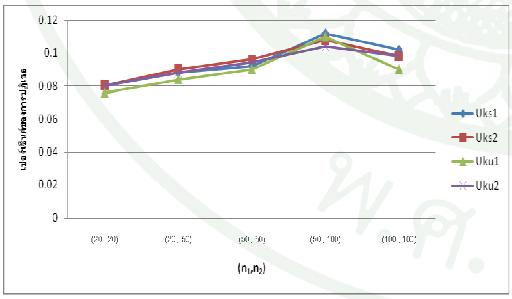
ระดับนัยสำคัญ 0.05



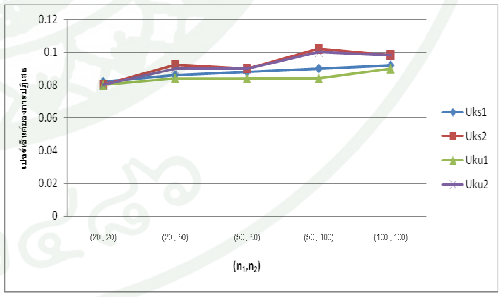
ตัวแบบที่ 1



ตัวแบบที่ 2



ตัวแบบที่ 3



ตัวแบบที่ 4

ระดับนัยสำคัญ 0.10

ภาพที่ 13 เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานหลักของตัวแบบที่ 1) - 4) สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25$, $\sigma_2(x) = 0.50$, $\sigma_3(x) = 0.75$ เมื่อการแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim N(0,1)$, $\varepsilon_2 \sim N(0,1)$, $\varepsilon_3 \sim N(0,1)$ ที่ระดับนัยสำคัญเท่ากับ 0.05 และ 0.10

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

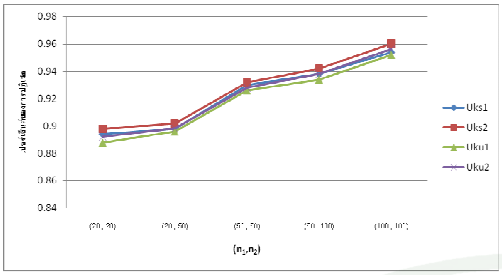
จากตารางที่ 9 และภาพที่ 13 พบว่าการควบคุมความคลาดเคลื่อนประเภทที่ 1 สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ จะทำได้ดีเมื่อตัวอย่างมีขนาดใหญ่ขึ้น เนื่องจากเมื่อกำหนดระดับนัยสำคัญเท่ากับ 0.05 ความคลาดเคลื่อนประเภทที่ 1 จะอยู่ในช่วง [0.036 , 0.063] ในทุกตัวแบบการถดถอย และเมื่อกำหนดระดับนัยสำคัญเท่ากับ 0.10 ความคลาดเคลื่อนประเภทที่ 1 จะอยู่ในช่วง [0.084 , 0.115] ในทุกตัวแบบการถดถอย ซึ่งแสดงให้เห็นว่าสถิติทดสอบสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้เมื่อตัวอย่างมีขนาดใหญ่

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

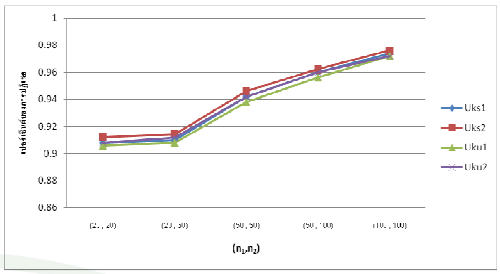
ตารางที่ 10 เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Constant Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25,\ \sigma_2(x) = 0.50, \sigma_3(x) = 0.75$ เมื่อการแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim N(0,1),\ \varepsilon_2 \sim N(0,1),\ \varepsilon_3 \sim N(0,1)$

(n_1, n_2, n_3)	ตัวแบบที่	U^1_{KS}		U^2_{KS}		U^1_{KU}		U^2_{KU}	
		ระดับนัยสำคัญ		ระดับนัยสำคัญ		ระดับนัยสำคัญ		ระดับนัยสำคัญ	
		0.05	0.10	0.05	0.10	0.05	0.10	0.05	0.10
(20 , 20, 20)	1)	0.894	0.916	0.898	0.920	0.888	0.912	0.892	0.916
	2)	0.908	0.924	0.912	0.928	0.906	0.920	0.908	0.926
	3)	0.880	0.926	0.892	0.932	0.876	0.922	0.892	0.930
	4)	0.406	0.432	0.412	0.436	0.400	0.428	0.410	0.432
(20 , 50 , 20)	1)	0.898	0.914	0.902	0.916	0.896	0.912	0.898	0.912
	2)	0.910	0.930	0.914	0.932	0.908	0.924	0.912	0.930
	3)	0.896	0.932	0.902	0.936	0.896	0.930	0.900	0.932
	4)	0.402	0.440	0.410	0.444	0.398	0.436	0.408	0.440
(50 , 50 , 50)	1)	0.930	0.936	0.932	0.938	0.926	0.932	0.928	0.936
	2)	0.942	0.954	0.946	0.956	0.938	0.948	0.942	0.950
	3)	0.906	0.932	0.910	0.936	0.902	0.926	0.908	0.932
	4)	0.412	0.450	0.416	0.454	0.408	0.448	0.412	0.450
(50 ,100 , 50)	1)	0.938	0.954	0.942	0.956	0.934	0.950	0.938	0.950
	2)	0.960	0.964	0.962	0.966	0.956	0.960	0.960	0.958
	3)	0.912	0.940	0.918	0.944	0.908	0.936	0.918	0.940
	4)	0.432	0.456	0.488	0.462	0.428	0.452	0.434	0.456
(100,100,100)	1)	0.954	0.964	0.960	0.970	0.952	0.962	0.956	0.962
	2)	0.974	0.976	0.976	0.980	0.972	0.974	0.972	0.976
	3)	0.928	0.942	0.932	0.946	0.926	0.940	0.930	0.942
	4)	0.454	0.470	0.458	0.472	0.450	0.466	0.450	0.466

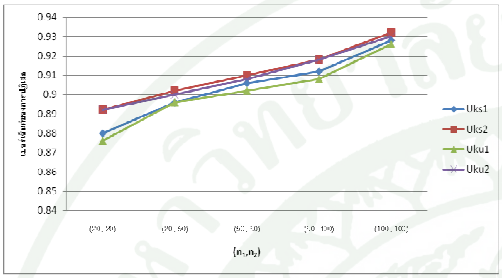
ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



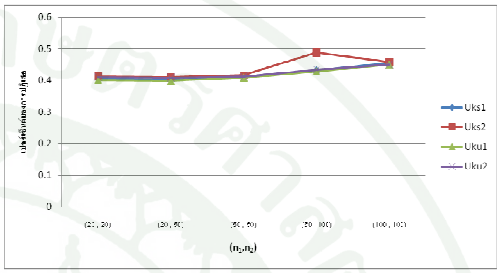
ตัวแบบที่ 1



ตัวแบบที่ 2

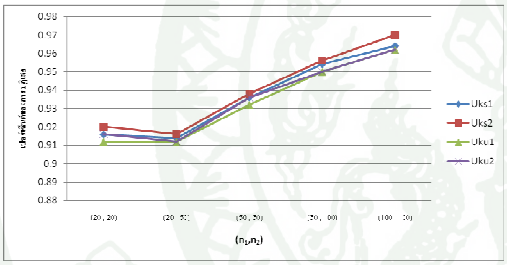


ตัวแบบที่ 3

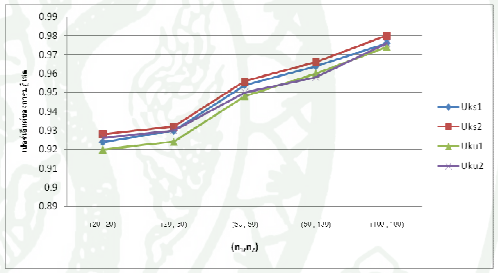


ตัวแบบที่ 4

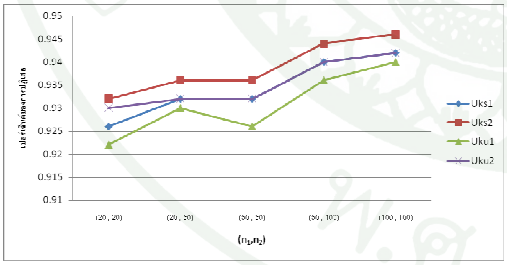
ระดับนัยสำคัญ 0.05



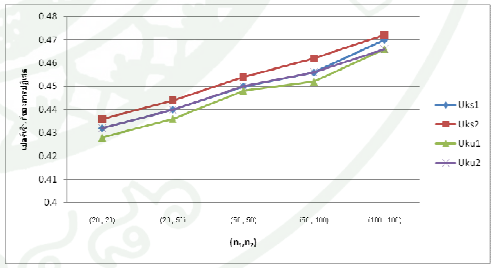
ตัวแบบที่ 1



ตัวแบบที่ 2



ตัวแบบที่ 3



ตัวแบบที่ 4

ระดับนัยสำคัญ 0.10

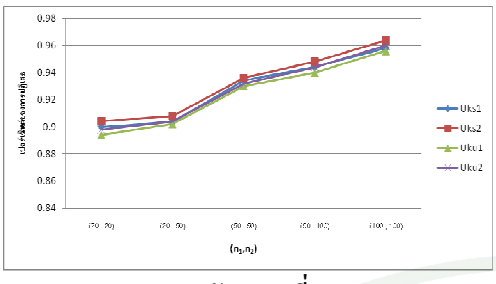
ภาพที่ 14 เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Constant Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \sigma_2(x) = 0.50, \sigma_3(x) = 0.75$ เมื่อการแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim N(0,1)$, $\varepsilon_2 \sim N(0,1), \varepsilon_3 \sim N(0,1)$ ที่ระดับนัยสำคัญเท่ากับ 0.05 และ 0.10

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

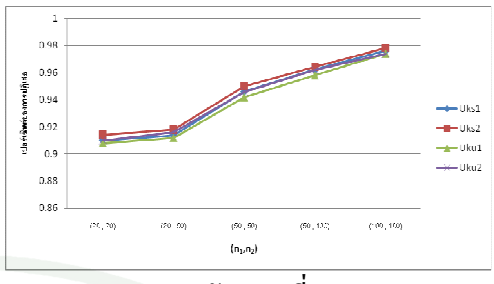
ตารางที่ 11 เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Affine Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25,\ \sigma_2(x) = 0.50, \sigma_3(x) = 0.75$ เมื่อการแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim N(0,1),\ \varepsilon_2 \sim N(0,1),\ \varepsilon_3 \sim N(0,1)$

(n_1, n_2, n_3)	ตัวแบบที่	U^1_{KS}		U^2_{KS}		U^1_{KU}		U^2_{KU}	
		ระดับนัยสำคัญ		ระดับนัยสำคัญ		ระดับนัยสำคัญ		ระดับนัยสำคัญ	
		0.05	0.10	0.05	0.10	0.05	0.10	0.05	0.10
(20 , 20, 20)	1)	0.900	0.918	0.904	0.922	0.894	0.914	0.898	0.918
	2)	0.910	0.926	0.914	0.930	0.908	0.922	0.910	0.928
	3)	0.884	0.930	0.896	0.936	0.880	0.926	0.896	0.934
	4)	0.408	0.436	0.414	0.440	0.402	0.432	0.412	0.436
(20 , 50 , 20)	1)	0.904	0.918	0.908	0.920	0.902	0.916	0.904	0.916
	2)	0.914	0.930	0.918	0.932	0.912	0.924	0.916	0.930
	3)	0.898	0.934	0.904	0.938	0.898	0.932	0.902	0.934
	4)	0.408	0.442	0.416	0.446	0.404	0.438	0.414	0.442
(50 , 50 , 50)	1)	0.934	0.938	0.936	0.940	0.930	0.934	0.932	0.938
	2)	0.946	0.956	0.950	0.958	0.942	0.950	0.946	0.952
	3)	0.912	0.936	0.916	0.940	0.908	0.930	0.914	0.936
	4)	0.416	0.450	0.420	0.454	0.412	0.448	0.416	0.450
(50 ,100 , 50)	1)	0.944	0.956	0.948	0.958	0.940	0.952	0.944	0.952
	2)	0.962	0.966	0.964	0.968	0.958	0.962	0.962	0.960
	3)	0.918	0.940	0.924	0.944	0.914	0.936	0.924	0.940
	4)	0.438	0.460	0.494	0.466	0.434	0.456	0.440	0.460
(100,100,100)	1)	0.958	0.964	0.964	0.970	0.956	0.962	0.960	0.962
	2)	0.976	0.978	0.978	0.982	0.974	0.976	0.974	0.978
	3)	0.930	0.944	0.934	0.948	0.928	0.942	0.932	0.944
	4)	0.454	0.472	0.458	0.474	0.450	0.468	0.450	0.468

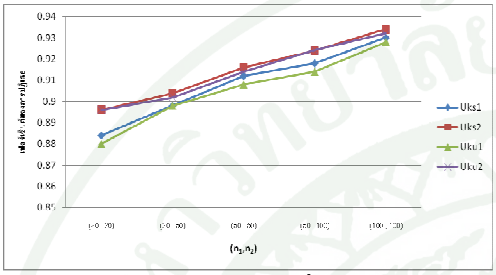
ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



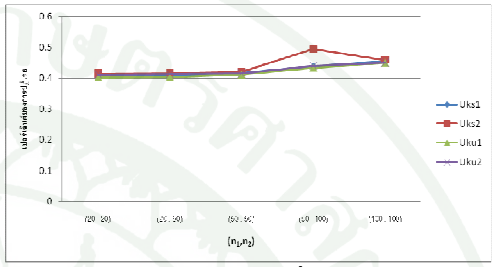
ตัวแบบที่ 1



ตัวแบบที่ 2

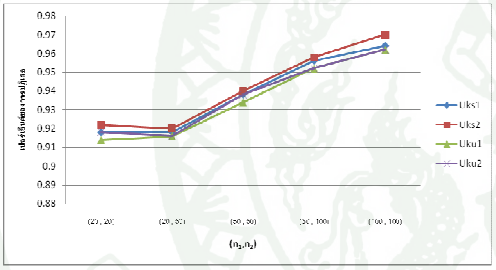


ตัวแบบที่ 3

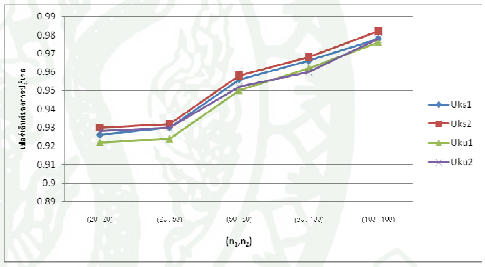


ตัวแบบที่ 4

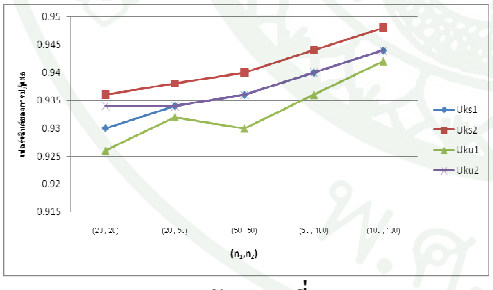
ระดับนัยสำคัญ 0.05



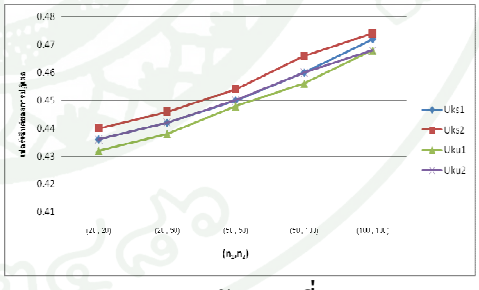
ตัวแบบที่ 1



ตัวแบบที่ 2



ตัวแบบที่ 3



ตัวแบบที่ 4

ระดับนัยสำคัญ 0.10

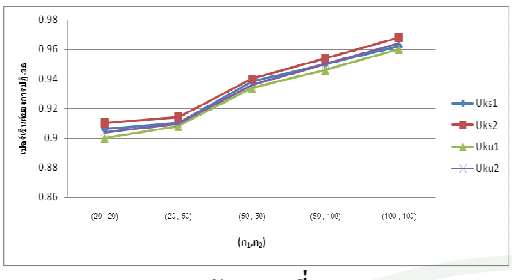
ภาพที่ 15 เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Affine Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \sigma_2(x) = 0.50, \sigma_3(x) = 0.75$ เมื่อการแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim N(0,1), \varepsilon_2 \sim N(0,1), \varepsilon_3 \sim N(0,1)$ ที่ระดับนัยสำคัญเท่ากับ 0.05 และ 0.10

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

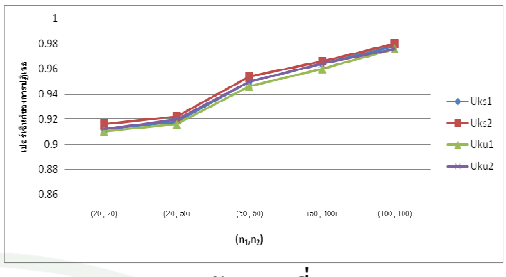
ตารางที่ 12 เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Quadratic Shift สำหรับสถิติทดสอบ U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU} ในกรณีที่ $\sigma_1(x) = 0.25,\ \sigma_2(x) = 0.50, \sigma_3(x) = 0.75$ เมื่อการแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim N(0,1),\ \varepsilon_2 \sim N(0,1),\ \varepsilon_3 \sim N(0,1)$

(n_1, n_2, n_3)	ตัวแบบที่	U^1_{KS}		U^2_{KS}		U^1_{KU}		U^2_{KU}	
		ระดับนัยสำคัญ		ระดับนัยสำคัญ		ระดับนัยสำคัญ		ระดับนัยสำคัญ	
		0.05	0.10	0.05	0.10	0.05	0.10	0.05	0.10
(20 , 20, 20)	1)	0.906	0.920	0.910	0.924	0.900	0.916	0.904	0.920
	2)	0.912	0.928	0.916	0.932	0.910	0.924	0.912	0.930
	3)	0.888	0.934	0.900	0.940	0.884	0.930	0.900	0.938
	4)	0.410	0.440	0.416	0.444	0.404	0.436	0.414	0.440
(20 , 50 , 20)	1)	0.910	0.922	0.914	0.924	0.908	0.920	0.910	0.920
	2)	0.918	0.930	0.922	0.932	0.916	0.924	0.920	0.930
	3)	0.900	0.936	0.906	0.940	0.900	0.934	0.904	0.936
	4)	0.414	0.444	0.422	0.448	0.410	0.440	0.420	0.444
(50 , 50 , 50)	1)	0.938	0.940	0.940	0.942	0.934	0.936	0.936	0.940
	2)	0.950	0.958	0.954	0.960	0.946	0.952	0.950	0.954
	3)	0.918	0.940	0.922	0.944	0.914	0.934	0.920	0.940
	4)	0.420	0.450	0.424	0.454	0.416	0.448	0.420	0.450
(50 ,100 , 50)	1)	0.950	0.958	0.954	0.960	0.946	0.954	0.950	0.954
	2)	0.964	0.968	0.966	0.970	0.960	0.964	0.964	0.962
	3)	0.924	0.940	0.930	0.944	0.920	0.936	0.930	0.940
	4)	0.444	0.464	0.500	0.470	0.440	0.460	0.446	0.464
(100,100,100)	1)	0.962	0.964	0.968	0.970	0.960	0.962	0.964	0.962
	2)	0.978	0.980	0.980	0.984	0.976	0.978	0.976	0.980
	3)	0.932	0.946	0.936	0.950	0.930	0.944	0.934	0.946
	4)	0.454	0.474	0.458	0.476	0.450	0.470	0.450	0.470

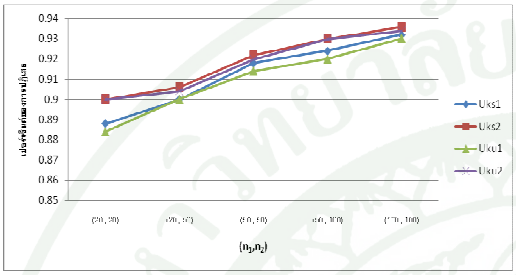
ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



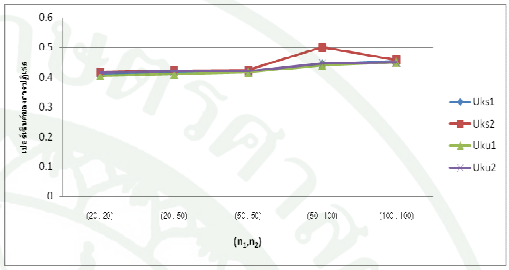
ตัวแบบที่ 1



ตัวแบบที่ 2

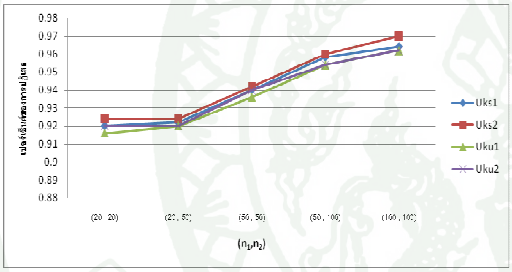


ตัวแบบที่ 3

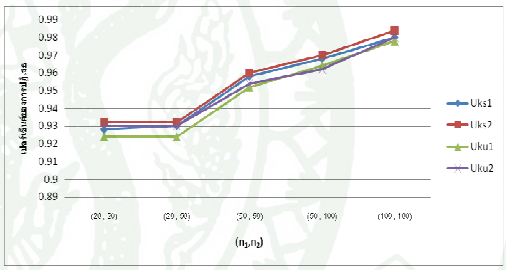


ตัวแบบที่ 4

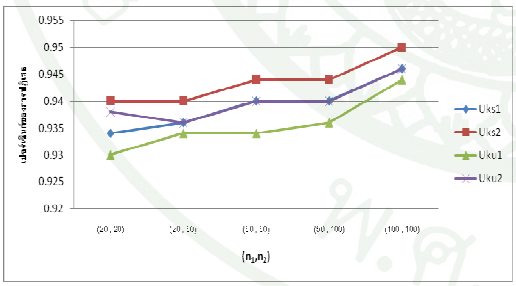
ระดับนัยสำคัญ 0.05



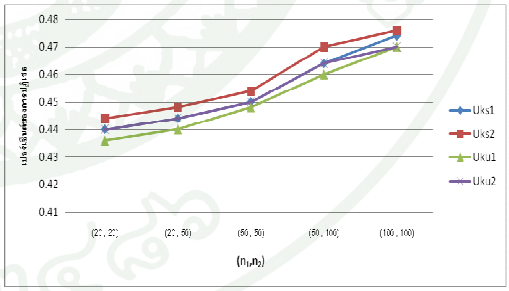
ตัวแบบที่ 1



ตัวแบบที่ 2



ตัวแบบที่ 3



ตัวแบบที่ 4

ระดับนัยสำคัญ 0.10

ภาพที่ 16 เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Quadratic Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \sigma_2(x) = 0.50, \sigma_3(x) = 0.75$ เมื่อการแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim N(0,1)$, $\varepsilon_2 \sim N(0,1)$, $\varepsilon_3 \sim N(0,1)$ ที่ระดับนัยสำคัญเท่ากับ 0.05 และ 0.10

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

จากตารางที่ 10 - 12 และภาพที่ 14 - 16 พบว่าที่ระดับนัยสำคัญเท่ากับ 0.05 และ 0.10 เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25,\ \sigma_2(x) = 0.50,\ \sigma_3(x) = 0.75$ เมื่อการแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim N(0,1),\ \varepsilon_2 \sim N(0,1),\ \varepsilon_3 \sim N(0,1)$ จะมีค่าอำนาจการทดสอบสูงขึ้นเมื่อขนาดของตัวอย่างเพิ่มขึ้น โดยเฉพาะในตัวแบบที่ 1 ตัวแบบที่ 2 และตัวแบบที่ 3 ค่าอำนาจการทดสอบสูงขึ้นจนเกือบถึง 1 ในขณะที่ตัวแบบที่ 4 ซึ่งได้แก่ ตัวแบบตรีโกณมิติจะให้ค่าอำนาจการทดสอบต่ำกว่าตัวแบบอื่น ๆ เมื่อพิจารณาที่การ Shift พบว่ากรณี Quadratic Shift ให้ค่าอำนาจการทดสอบสูงที่สุด รองลงมาได้แก่กรณี Affine Shift และ Constant Shift ตามลำดับ

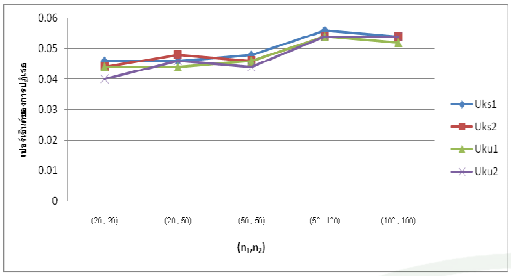
ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

ตารางที่ 13 เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานหลักของตัวแบบที่ 1) - 4) สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \sigma_2(x) = 0.50, \sigma_3(x) = 0.75$ เมื่อการแจกแจงของความคลาดเคลื่อน ε_1 ~ Exponential (1) - 1, ε_2 ~ Exponential (1) - 1, ε_3 ~ Exponential (1) - 1

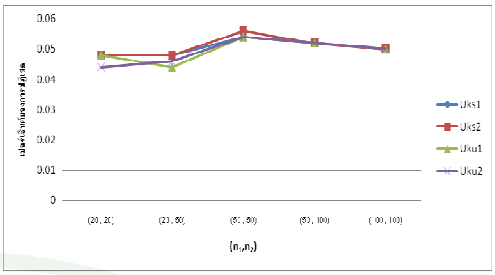
(n_1, n_2, n_3)	ตัวแบบที่	U^1_{KS}		U^2_{KS}		U^1_{KU}		U^2_{KU}	
		ระดับนัยสำคัญ		ระดับนัยสำคัญ		ระดับนัยสำคัญ		ระดับนัยสำคัญ	
		0.05	0.10	0.05	0.10	0.05	0.10	0.05	0.10
(20 , 20, 20)	1)	0.046	0.080*	0.044	0.082*	0.044	0.078*	0.040	0.080*
	2)	0.048	0.088	0.048	0.090	0.048	0.086	0.044	0.086
	3)	0.030*	0.080*	0.032*	0.084	0.030*	0.078*	0.030*	0.080*
	4)	0.038	0.082*	0.040	0.084	0.034*	0.080*	0.040	0.082*
(20 , 50 , 20)	1)	0.046	0.094	0.048	0.096	0.044	0.090	0.046	0.094
	2)	0.048	0.094	0.048	0.094	0.044	0.092	0.046	0.090
	3)	0.038	0.086	0.040	0.090	0.034*	0.084	0.038	0.088
	4)	0.040	0.090	0.042	0.094	0.040	0.090	0.040	0.092
(50 , 50 , 50)	1)	0.048	0.084	0.046	0.086	0.046	0.080*	0.044	0.084
	2)	0.054	0.098	0.056	0.096	0.054	0.096	0.054	0.094
	3)	0.054	0.096	0.054	0.098	0.052	0.092	0.054	0.096
	4)	0.056	0.084	0.056	0.088	0.054	0.080*	0.054	0.084
(50 ,100 , 50)	1)	0.056	0.112	0.054	0.110	0.054	0.110	0.054	0.110
	2)	0.052	0.104	0.052	0.102	0.052	0.102	0.052	0.102
	3)	0.056	0.110	0.056	0.112	0.056	0.108	0.054	0.110
	4)	0.054	0.106	0.054	0.104	0.054	0.104	0.052	0.104
(100,100,100)	1)	0.054	0.110	0.054	0.112	0.052	0.108	0.054	0.110
	2)	0.050	0.102	0.050	0.104	0.050	0.102	0.050	0.102
	3)	0.052	0.104	0.052	0.106	0.052	0.102	0.052	0.104
	4)	0.054	0.110	0.054	0.112	0.054	0.108	0.054	0.110

หมายเหตุ * แทนความคลาดเคลื่อนประเภทที่ 1 อยู่นอกช่วงของการควบคุม

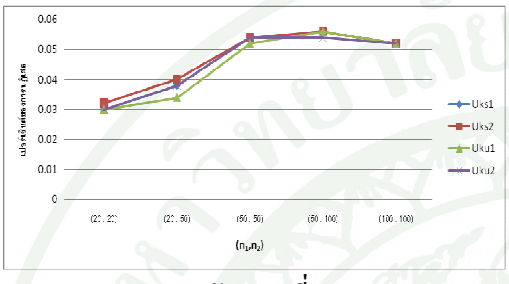
ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



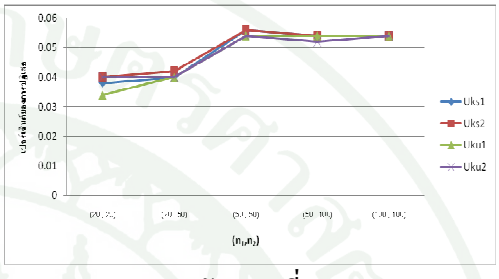
ตัวแบบที่ 1



ตัวแบบที่ 2

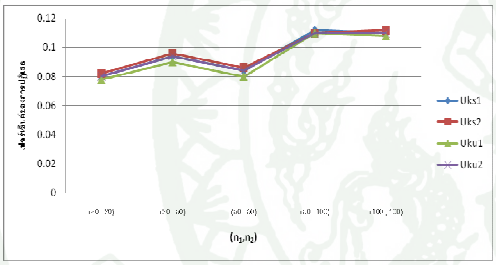


ตัวแบบที่ 3

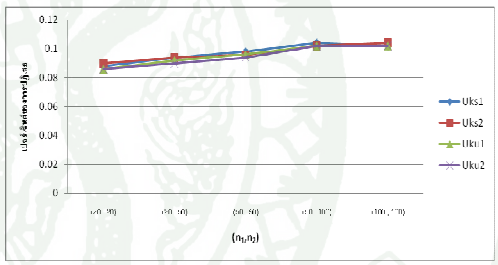


ตัวแบบที่ 4

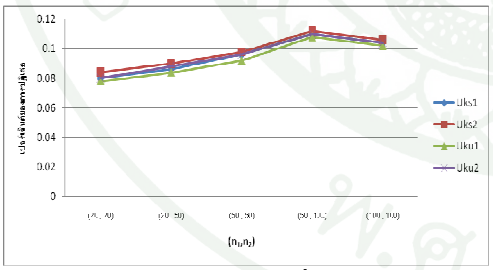
ระดับนัยสำคัญ 0.05



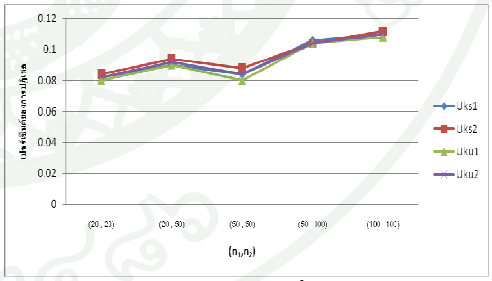
ตัวแบบที่ 1



ตัวแบบที่ 2



ตัวแบบที่ 3



ตัวแบบที่ 4

ระดับนัยสำคัญ 0.10

ภาพที่ 17 เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานหลักของตัวแบบที่ 1) - 4) สำหรับสถิติทดสอบ U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU} ในกรณีที่ $\sigma_1(x) = 0.25$, $\sigma_2(x) = 0.50$, $\sigma_3(x) = 0.75$ เมื่อการแจกแจงของความคลาดเคลื่อน ε_1 ~ Exponential (1) - 1, ε_2 ~ Exponential (1) - 1, ε_3 ~ Exponential (1) - 1 ที่ระดับนัยสำคัญเท่ากับ 0.05 และ 0.10

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

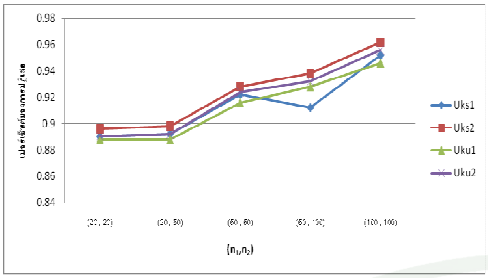
จากตารางที่ 13 และภาพที่ 17 พบว่าการควบคุมความคลาดเคลื่อนประเภทที่ 1 สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ จะทำได้ดีเมื่อตัวอย่างมีขนาดใหญ่ขึ้น เนื่องจากเมื่อกำหนดระดับนัยสำคัญเท่ากับ 0.05 ความคลาดเคลื่อนประเภทที่ 1 จะอยู่ในช่วง [0.036 , 0.063] ในทุกตัวแบบการถดถอย และเมื่อกำหนดระดับนัยสำคัญเท่ากับ 0.10 ความคลาดเคลื่อนประเภทที่ 1 จะอยู่ในช่วง [0.084 , 0.115] ในทุกตัวแบบการถดถอยซึ่งแสดงให้เห็นว่าสถิติทดสอบสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดีเมื่อตัวอย่างมีขนาดใหญ่

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

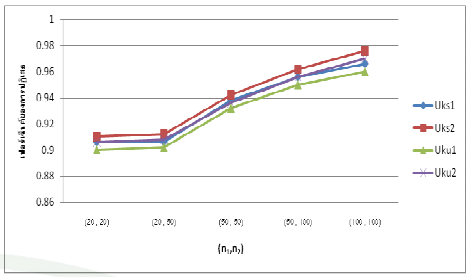
ตารางที่ 14 เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Constant Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \quad \sigma_2(x) = 0.50, \sigma_3(x) = 0.75$ เมื่อการแจกแจงของความคลาดเคลื่อน ε_1 ~ Exponential (1) - 1, ε_2 ~ Exponential (1) - 1, ε_3 ~ Exponential (1) - 1

(n_1, n_2, n_3)	ตัวแบบที่	U^1_{KS} ระดับนัยสำคัญ		U^2_{KS} ระดับนัยสำคัญ		U^1_{KU} ระดับนัยสำคัญ		U^2_{KU} ระดับนัยสำคัญ	
		0.05	0.10	0.05	0.10	0.05	0.10	0.05	0.10
(20 , 20, 20)	1)	0.890	0.910	0.896	0.916	0.888	0.906	0.890	0.910
	2)	0.906	0.916	0.910	0.922	0.900	0.908	0.906	0.918
	3)	0.872	0.922	0.880	0.932	0.868	0.916	0.872	0.926
	4)	0.396	0.412	0.406	0.420	0.396	0.412	0.396	0.412
(20 , 50 , 20)	1)	0.892	0.912	0.898	0.918	0.888	0.904	0.892	0.912
	2)	0.906	0.924	0.912	0.930	0.902	0.920	0.908	0.924
	3)	0.886	0.926	0.894	0.932	0.880	0.920	0.890	0.932
	4)	0.398	0.436	0.404	0.440	0.392	0.430	0.400	0.436
(50 , 50 ,50)	1)	0.922	0.936	0.928	0.442	0.916	0.926	0.924	0.938
	2)	0.938	0.946	0.942	0.950	0.932	0.942	0.936	0.946
	3)	0.902	0.922	0.908	0.930	0.898	0.920	0.904	0.928
	4)	0.406	0.418	0.416	0.454	0.402	0.440	0.412	0.450
(50 ,100 ,50)	1)	0.912	0.946	0.938	0.954	0.928	0.940	0.932	0.948
	2)	0.956	0.964	0.962	0.970	0.950	0.956	0.956	0.966
	3)	0.912	0.938	0.918	0.946	0.908	0.930	0.914	0.940
	4)	0.428	0.452	0.438	0.462	0.422	0.446	0.430	0.456
(100,100,100)	1)	0.952	0.964	0.962	0.978	0.946	0.960	0.956	0.972
	2)	0.966	0.974	0.976	0.982	0.960	0.966	0.970	0.976
	3)	0.926	0.932	0.934	0.936	0.918	0.926	0.930	0.934
	4)	0.450	0.460	0.454	0.466	0.444	0.456	0.450	0.456

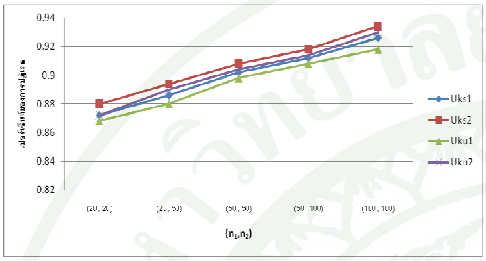
ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



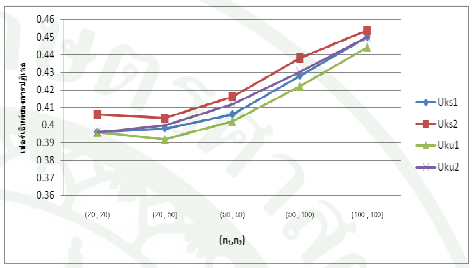
ตัวแบบที่ 1



ตัวแบบที่ 2

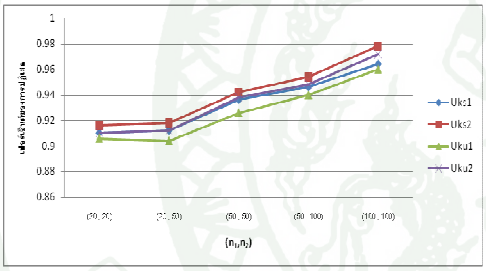


ตัวแบบที่ 3

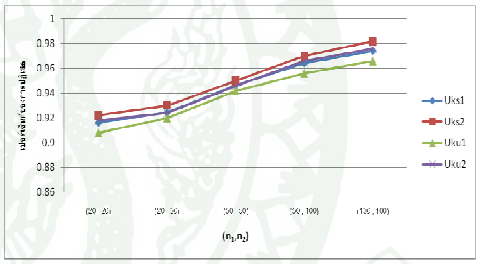


ตัวแบบที่ 4

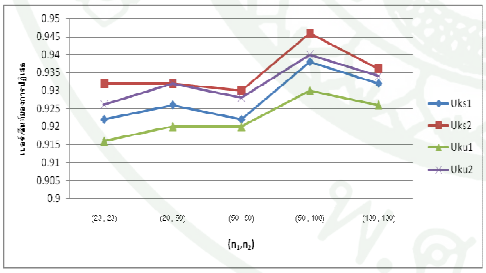
ระดับนัยสำคัญ 0.05



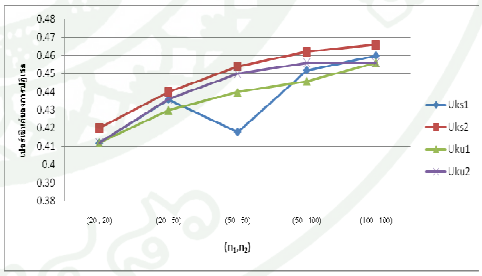
ตัวแบบที่ 1



ตัวแบบที่ 2



ตัวแบบที่ 3



ตัวแบบที่ 4

ระดับนัยสำคัญ 0.10

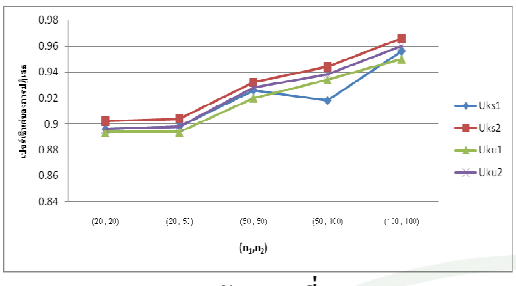
ภาพที่ 18 เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Constant Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25,\ \sigma_2(x) = 0.50, \sigma_3(x) = 0.75$ เมื่อการแจกแจงของความคลาดเคลื่อน ε_1 ~ Exponential (1) - 1, ε_2 ~ Exponential (1) - 1, ε_3 ~ Exponential (1) - 1 ที่ระดับนัยสำคัญเท่ากับ 0.05 และ 0.10

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

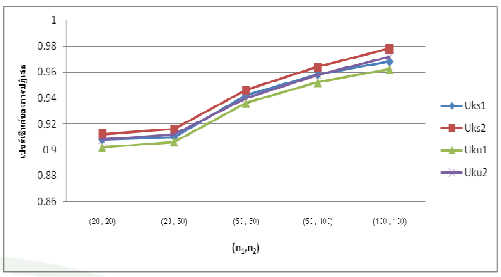
ตารางที่ 15 เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Affine Shift สำหรับสถิติทดสอบ U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU} ในกรณีที่ $\sigma_1(x) = 0.25$, $\sigma_2(x) = 0.50$, $\sigma_3(x) = 0.75$ เมื่อการแจกแจงของความคลาดเคลื่อน ε_1 ~ Exponential (1) - 1, ε_2 ~ Exponential (1) - 1, ε_3 ~ Exponential (1) - 1

(n_1, n_2, n_3)	ตัวแบบที่	U^1_{KS}		U^2_{KS}		U^1_{KU}		U^2_{KU}	
		ระดับนัยสำคัญ		ระดับนัยสำคัญ		ระดับนัยสำคัญ		ระดับนัยสำคัญ	
		0.05	0.10	0.05	0.10	0.05	0.10	0.05	0.10
(20 , 20, 20)	1)	0.896	0.912	0.902	0.918	0.894	0.908	0.896	0.912
	2)	0.908	0.918	0.912	0.924	0.902	0.910	0.908	0.920
	3)	0.876	0.926	0.884	0.936	0.872	0.920	0.876	0.930
	4)	0.398	0.416	0.408	0.424	0.398	0.416	0.398	0.416
(20 , 50 , 20)	1)	0.898	0.916	0.904	0.922	0.894	0.908	0.898	0.916
	2)	0.910	0.924	0.916	0.930	0.906	0.920	0.912	0.924
	3)	0.888	0.928	0.896	0.934	0.882	0.922	0.892	0.934
	4)	0.404	0.438	0.410	0.442	0.398	0.432	0.406	0.438
(50 , 50 ,50)	1)	0.926	0.938	0.932	0.944	0.920	0.928	0.928	0.944
	2)	0.942	0.948	0.946	0.952	0.936	0.944	0.940	0.948
	3)	0.908	0.926	0.914	0.934	0.904	0.924	0.910	0.932
	4)	0.410	0.418	0.420	0.454	0.406	0.440	0.416	0.450
(50 ,100 ,50)	1)	0.918	0.948	0.944	0.956	0.934	0.942	0.938	0.950
	2)	0.958	0.966	0.964	0.972	0.952	0.958	0.958	0.968
	3)	0.918	0.938	0.924	0.946	0.914	0.930	0.920	0.940
	4)	0.434	0.456	0.444	0.466	0.428	0.450	0.436	0.460
(100,100,100)	1)	0.956	0.964	0.966	0.978	0.950	0.960	0.960	0.972
	2)	0.968	0.976	0.978	0.984	0.962	0.968	0.972	0.978
	3)	0.928	0.934	0.936	0.938	0.920	0.928	0.932	0.936
	4)	0.450	0.462	0.454	0.468	0.444	0.458	0.450	0.458

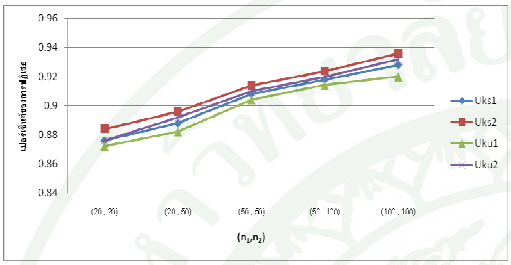
ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



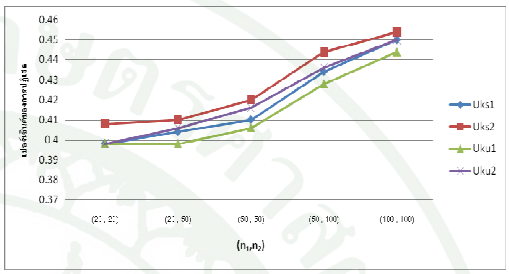
ตัวแบบที่ 1



ตัวแบบ 2

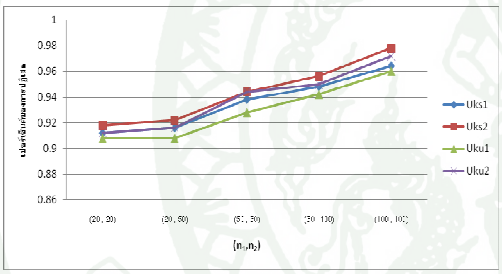


ตัวแบบที่ 3

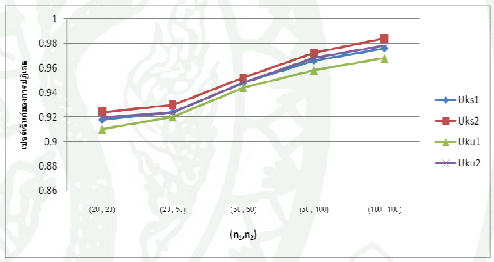


ตัวแบบ 4

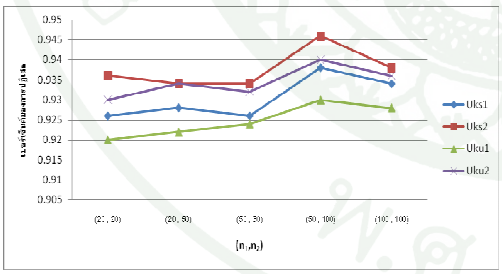
ระดับนัยสำคัญ 0.05



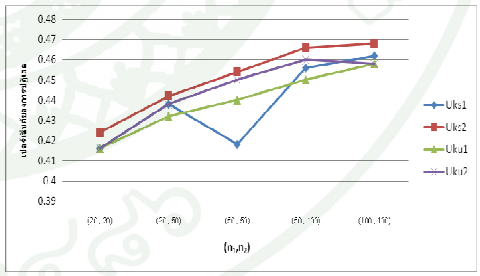
ตัวแบบที่ 1



ตัวแบบ 2



ตัวแบบที่ 3



ตัวแบบ 4

ระดับนัยสำคัญ 0.10

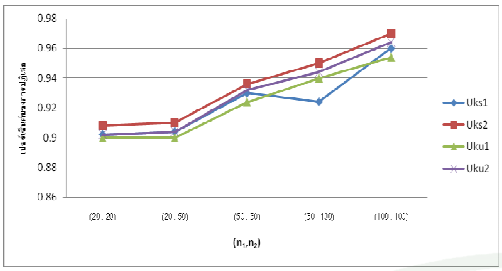
ภาพที่ 19 เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Affine Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25,\ \sigma_2(x) = 0.50, \sigma_3(x) = 0.75$ เมื่อการแจกแจงของความคลาดเคลื่อน ε_1 ~ Exponential (1) - 1, ε_2 ~ Exponential (1) - 1, ε_3 ~ Exponential (1) - 1 ที่ระดับนัยสำคัญเท่ากับ 0.05 และ 0.10

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

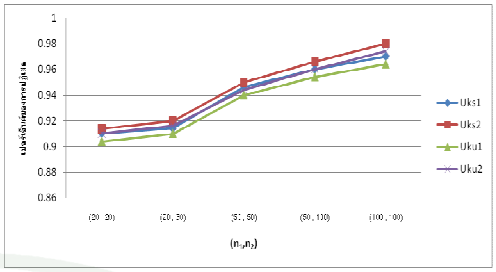
ตารางที่ 16 เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Quadratic Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25,\ \sigma_2(x) = 0.50, \sigma_3(x) = 0.75$ เมื่อการแจกแจงของความคลาดเคลื่อน ε_1 ~ Exponential (1) - 1, ε_2 ~ Exponential (1) - 1, ε_3 ~ Exponential (1) - 1

(n_1, n_2, n_3)	ตัวแบบที่	U^1_{KS} ระดับนัยสำคัญ		U^2_{KS} ระดับนัยสำคัญ		U^1_{KU} ระดับนัยสำคัญ		U^2_{KU} ระดับนัยสำคัญ	
		0.05	0.10	0.05	0.10	0.05	0.10	0.05	0.10
(20 , 20, 20)	1)	0.902	0.914	0.908	0.920	0.900	0.910	0.902	0.914
	2)	0.910	0.920	0.914	0.926	0.904	0.912	0.910	0.922
	3)	0.880	0.930	0.888	0.940	0.876	0.924	0.880	0.934
	4)	0.400	0.420	0.410	0.428	0.400	0.420	0.400	0.420
(20 , 50 , 20)	1)	0.904	0.920	0.910	0.926	0.900	0.912	0.904	0.920
	2)	0.914	0.924	0.920	0.930	0.910	0.920	0.916	0.924
	3)	0.890	0.930	0.898	0.936	0.884	0.924	0.894	0.936
	4)	0.410	0.440	0.416	0.444	0.404	0.434	0.412	0.440
(50 , 50 , 50)	1)	0.930	0.940	0.936	0.946	0.924	0.930	0.932	0.942
	2)	0.946	0.950	0.950	0.954	0.940	0.946	0.944	0.950
	3)	0.914	0.930	0.920	0.938	0.910	0.928	0.916	0.936
	4)	0.414	0.418	0.424	0.454	0.410	0.410	0.420	0.450
(50 ,100 , 50)	1)	0.924	0.950	0.950	0.958	0.940	0.944	0.944	0.952
	2)	0.960	0.968	0.966	0.974	0.954	0.960	0.960	0.970
	3)	0.924	0.938	0.930	0.946	0.920	0.930	0.926	0.940
	4)	0.440	0.460	0.450	0.470	0.434	0.454	0.442	0.464
(100,100,100)	1)	0.960	0.964	0.970	0.978	0.954	0.960	0.964	0.972
	2)	0.970	0.978	0.980	0.986	0.964	0.970	0.974	0.980
	3)	0.930	0.936	0.938	0.940	0.922	0.930	0.934	0.938
	4)	0.450	0.464	0.454	0.470	0.444	0.460	0.450	0.460

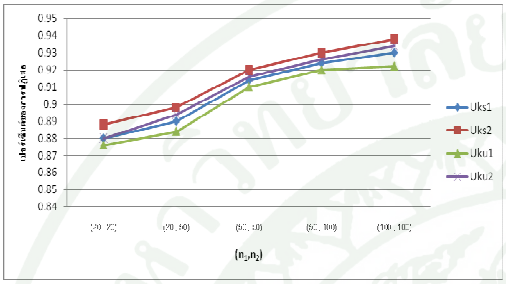
ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



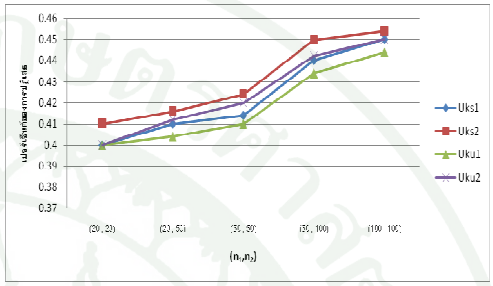
ตัวแบบที่ 1



ตัวแบบที่ 2

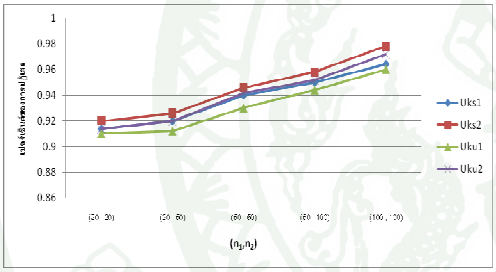


ตัวแบบที่ 3

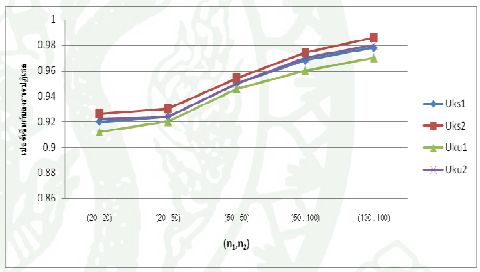


ตัวแบบที่ 4

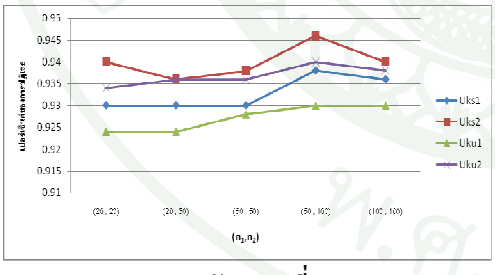
ระดับนัยสำคัญ 0.05



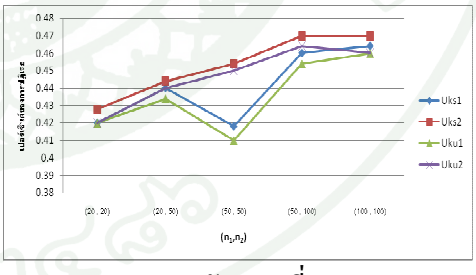
ตัวแบบที่ 1



ตัวแบบที่ 2



ตัวแบบที่ 1



ตัวแบบที่ 2

ระดับนัยสำคัญ 0.10

ภาพที่ 20 เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Quadratic Shift สำหรับสถิติทดสอบ U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU} ในกรณีที่ $\sigma_1(x) = 0.25$, $\sigma_2(x) = 0.50$, $\sigma_3(x) = 0.75$ เมื่อการแจกแจงของความคลาดเคลื่อน ε_1 ~ Exponential (1) - 1, ε_2 ~ Exponential (1) - 1, ε_3 ~ Exponential (1) - 1 ที่ระดับนัยสำคัญเท่ากับ 0.05 และ 0.10

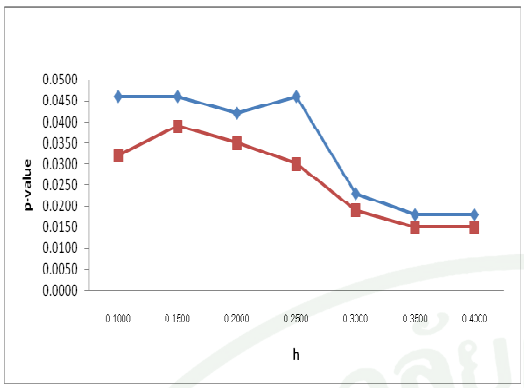
ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

จากตารางที่ 14 - 16 และภาพที่ 18 - 20 พบว่าที่ระดับนัยสำคัญเท่ากับ 0.05 และ 0.10 เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่ 1) - 4) กรณี Quadratic Shift สำหรับสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ในกรณีที่ $\sigma_1(x) = 0.25, \ \sigma_2(x) = 0.50, \ \sigma_3(x) = 0.75$ เมื่อการแจกแจงของความคลาดเคลื่อน $\varepsilon_1 \sim$ Exponential (1) - 1, $\varepsilon_2 \sim$ Exponential (1) - 1, $\varepsilon_3 \sim$ Exponential (1) - 1 จะมีค่าอำนาจการทดสอบสูงขึ้นเมื่อขนาดของตัวอย่างเพิ่มขึ้น โดยเฉพาะในตัวแบบที่ 1 ตัวแบบที่ 2 และตัวแบบที่ 3 ค่าอำนาจการทดสอบสูงขึ้นจนเกือบถึง 1 ในขณะที่ตัวแบบที่ 4 ซึ่งได้แก่ตัวแบบตรีโกณมิติจะให้ค่าอำนาจการทดสอบต่ำกว่าตัวแบบอื่น ๆ เมื่อพิจารณาที่การ Shift พบว่ากรณี Quadratic Shift ให้ค่าอำนาจการทดสอบสูงที่สุด รองลงมาได้แก่ กรณี Affine Shift และ Constant Shift ตามลำดับ

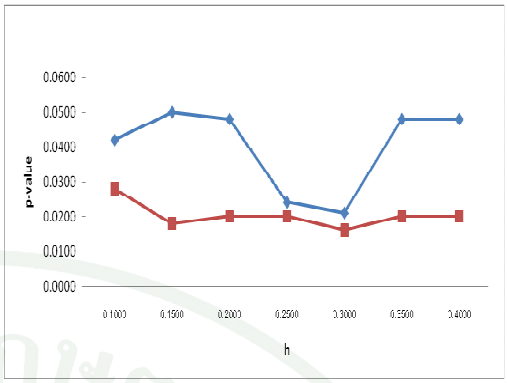
ส่วนที่ 3 การประยุกต์สถิติทดสอบการเท่ากันของฟังก์ชันการถดถอย k ฟังก์ชัน (k = 3) กับข้อมูลจริง

ผลการทดสอบการเท่ากันของ 3 ฟังก์ชันการถดถอย สำหรับสถิติทดสอบ U^1_{KS} และ U^1_{KU} เมื่อกำหนดให้พารามิเตอร์ปรับให้เรียบ มีค่าเท่ากับ 0.10, 0.15, 0.20, 0.25, 0.30, 0.35 และ 0.40 ดำเนินการทำซ้ำจำนวน 1,000 Bootstrap Replications ได้ค่า P - value แสดงดังภาพที่ 21

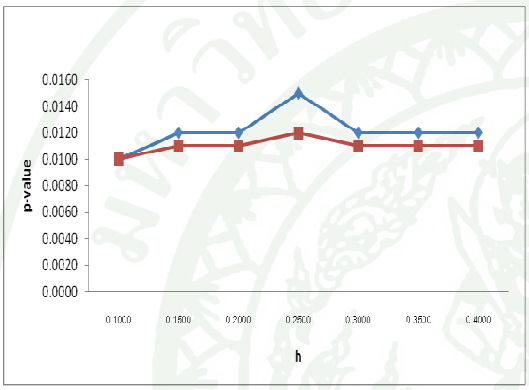
ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



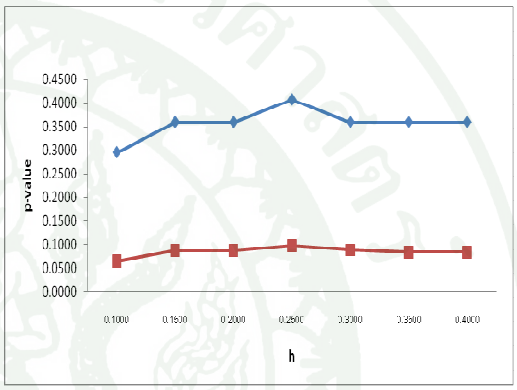
ปี พ.ศ. 2547, 2549 และ 2550



ปี พ.ศ. 2547 และ 2549



ปี พ.ศ. 2547 และ 2550



ปี พ.ศ. 2549 และ 2550

ภาพที่ 21 ค่า P - values ที่พารามิเตอร์ (h) เท่ากับ 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40 สำหรับสถิติทดสอบ U^1_{KS} (◆) และ U^1_{KU} (■)

ผลการทดสอบการเท่ากันของฟังก์ชันการถดถอย 3 ฟังก์ชัน สำหรับสถิติทดสอบ U^1_{KS} และ U^1_{KU} จากภาพที่ 21 แสดงให้เห็นว่า เมื่อทำการทดสอบการเท่ากันของฟังก์ชันการถดถอย 3 ฟังก์ชันการถดถอย (Over All-Test) โดย Y = log (ค่าใช้จ่ายเฉลี่ยต่อครัวเรือน) X = log (ค่าใช้จ่ายเฉลี่ยต่อครัวเรือน) ระหว่างปี พ.ศ. 2547 2549 และ 2550 สำหรับสถิติทดสอบ U^1_{KS} พบว่ามีค่า P-value อยู่ระหว่าง 0.0180 – 0.0460 ในขณะที่สถิติ U^1_{KU} มีค่า P-value อยู่ระหว่าง 0.0150 – 0.039

เมื่อทำการทดสอบการเท่ากันของฟังก์ชันการถดถอย 2 ฟังก์ชันการถดถอย โดย Y = log (ค่าใช้จ่ายเฉลี่ยต่อครัวเรือน) X= log (รายได้เฉลี่ยต่อครัวเรือน) ระหว่างปี พ.ศ. 2547 และ 2549 สำหรับสถิติทดสอบ U^1_{KS} พบว่ามีค่า P-value อยู่ระหว่าง 0.0180 – 0.0500 ในขณะที่สถิติ U^1_{KU} มีค่า P - value อยู่ระหว่าง 0.0160 – 0.0280

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

เมื่อทำการทดสอบการเท่ากันของฟังก์ชันการถดถอย 2 ฟังก์ชันการถดถอย โดย Y = log (ค่าใช้จ่ายเฉลี่ยต่อครัวเรือน) X = log (รายได้เฉลี่ยต่อครัวเรือน) ระหว่างปี พ.ศ. 2547 และ 2550 สำหรับสถิติทดสอบ U^1_{KS} พบว่ามีค่า P-value อยู่ระหว่าง 0.0100 – 0.0150 ในขณะที่สถิติ U^1_{KU} มีค่า P - value อยู่ระหว่าง 0.0100 – 0.0120

เมื่อทำการทดสอบการเท่ากันของฟังก์ชันการถดถอย 2 ฟังก์ชันการถดถอย โดย Y = log (ค่าใช้จ่ายเฉลี่ยต่อครัวเรือน) X = log (รายได้เฉลี่ยต่อครัวเรือน) ระหว่างปี พ.ศ. 2549 และ 2550 สำหรับสถิติทดสอบ U^1_{KS} พบว่ามีค่า P-value อยู่ระหว่าง 0.2950 – 0.4080 ในขณะที่สถิติ U^1_{KU} มีค่า P-value อยู่ระหว่าง 0.0630 – 0.0980

ซึ่งแสดงว่าในภาพรวม (Over all - test) เส้นโค้งการถดถอยทั้ง 3 เส้นมีความแตกต่างกัน แต่เมื่อพิจารณาเส้นโค้งการถดถอยแต่ละคู่ พบว่าเส้นโค้งการถดถอยคู่ที่มีความแตกต่างกันคือ คู่ที่ 1) ปี พ.ศ. 2547 และ ปี พ.ศ. 2550 คู่ที่ 2) ปี พ.ศ. 2547 และ ปี พ.ศ. 2549 ในขณะที่เส้นโค้งการถดถอยระหว่างปี พ.ศ. 2549 และ ปี พ.ศ. 2550 ไม่แตกต่างกันที่ระดับนัยสำคัญ 0.05

วิจารณ์

ในการศึกษาในครั้งนี้แบ่งผลการศึกษา ออกเป็น 3 ส่วนใหญ่ ๆ คือ

ส่วนที่ 1 ศึกษาสถิติทดสอบการเท่ากันของฟังก์ชันการถดถอย k ฟังก์ชัน เพื่อหาการแจกแจงของสถิติทดสอบภายใต้สมมติฐานหลัก $H_0: m_1 = m_2 = m_3 = \ldots = m_k$ เมื่อ $k \geq 2$ และสมมติฐานทางเลือก $H_1: m_i \neq m_j$ สำหรับ i, j บางค่า $i \neq j,\ i, j \in \{1, \ldots, k\}$

ส่วนที่ 2 การจำลองแบบในสถานการณ์ต่าง ๆ เพื่อศึกษาเปอร์เซ็นต์ของการปฏิเสธสมมติฐานภายใต้สมมติฐานหลักและเปอร์เซ็นต์ของการปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกหรืออำนาจของการทดสอบของสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ เมื่อตัวแบบของฟังก์ชันการถดถอยอยู่ในรูปของ 1) ฟังก์ชันคงที่ 2) ฟังก์ชันเชิงเส้นตรง 3) ฟังก์ชันเอ็กโพเนนเชียล และ 4) ฟังก์ชันตรีโกณมิติ โดยขนาดของตัวอย่างแบ่งเป็น ตัวอย่างขนาดเล็ก n = 20 ตัวอย่างขนาดกลาง n = 50 ตัวอย่างขนาดใหญ่ n = 100 และได้กำหนดให้ความคลาดเคลื่อนมีการแจกแจงใน 2 ลักษณะคือ N(0,1) และ Exponential (1) - 1

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

โดยได้แบ่งการจำลองแบบเป็น 2 ส่วนใหญ่ ๆ คือ การทดสอบการเท่ากันของฟังก์ชันการถดถอย 2 ฟังก์ชัน และ การทดสอบการเท่ากันของฟังก์ชันการถดถอย 3 ฟังก์ชัน

ส่วนที่ 3 การประยุกต์สถิติทดสอบการเท่ากันของ k ฟังก์ชันการถดถอย $k = 3$ กับข้อมูลจริง

ผลการศึกษาพบว่า

ส่วนที่ 1 ศึกษาสถิติทดสอบการเท่ากันของฟังก์ชันการถดถอย k ฟังก์ชัน เพื่อหาการแจกแจงของสถิติทดสอบภายใต้สมมติฐานหลัก $H_0: m_1 = m_2 = m_3 = \ldots = m_k$ เมื่อ $k \geq 2$ และสมมติฐานทางเลือก $H_1: m_i \neq m_j$ สำหรับ i, j บางค่า $i \neq j$, $i, j \in \{1, \ldots, k\}$

1. ภายใต้สมมติฐานหลัก สำหรับ $j = 1, \ldots, k$

1.1 สถิติทดสอบ $U^1_{KS} \xrightarrow{d} \sum_{j=1}^{k} \sup_y \left|U_j(y)\right|$, $U^1_{KS} = \sum_{j=1}^{k} \sup_y \left|\hat{U}_j(y)\right|$

1.2 สถิติทดสอบ $U^1_{KU} \xrightarrow{d} \sum_{i=1}^{k} \left[\sup_y \left|U_j(y)\right| - \inf_y \left|U_j(y)\right| \right]$,

$$U^1_{ku} = \sum_{i=1}^{k} \left[\sup_y \left|\hat{U}_j(y)\right| - \inf_y \left|\hat{U}_j(y)\right| \right]$$

1.3 สถิติทดสอบ $U^2_{KS} \xrightarrow{d} \sup_y \left|U(y)\right|$, $U^2_{KS} = \sup_y \left|\hat{U}(y)\right|$,

$$U(y) = \sum_{j=1}^{k} p_j^{1/2} U_j(y)$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

1.4 สถิติทดสอบ $U_{ku}^2 \xrightarrow{d} \left[\sup_y |U(y)| - \inf_y |U(y)| \right]$,

$$U_{ku}^2 = \left[\sup_y |\hat{U}(y)| - \inf_y |\hat{U}(y)| \right], \quad U(y) = \sum_{j=1}^{k} p_j^{1/2} U_j(y)$$

เมื่อ $U_j(y) = f_{\varepsilon j}(y) p_j \sum_{r=1}^{k} \sum_{i=1}^{n_r} \frac{Y_{ir} - m(X_{ir})}{\sigma_j(X_{ir})} \left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j)}{p_j} \right)$ คือตัวแปรสุ่มปกติ

ที่มีค่าเฉลี่ยเท่ากับ 0 และมีโครงสร้างความแปรปรวนร่วมคือ

$$Cov(U_j(y), U_{j'}(y)) = f_{\varepsilon j}(y) f_{\varepsilon j'}(y) p_j p_{j'} \sum_{r=1}^{k} E\left[\frac{\sigma_r^2(X_{ir})}{\sigma_j(X_{ir})\sigma_{j'}(X_{ir})} \left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j)}{p_j} \right) \right.$$

$$\left. \left(\frac{f_{j'}(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j')}{p_{j'}} \right) \right]$$

ดังนั้น $|U_j(y)|$ คือ ตัวแปรสุ่ม Half – Normal ที่มีค่าเฉลี่ย

$$\sqrt{f_{\varepsilon j}(y) f_{\varepsilon j'}(y) p_j p_{j'} \sum_{r=1}^{k} E\left[\frac{\sigma_r^2(X_{ir})}{\sigma_j(X_{ir})\sigma_{j'}(X_{ir})} \left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j)}{p_j} \right) \right.}$$

$$\left. \sqrt{\left(\frac{f_{j'}(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j')}{p_{j'}} \right)} \right] \sqrt{\frac{2}{\pi}}$$

และมีโครงสร้างความแปรปรวนร่วมคือ

$$Cov(U_j(y), U_{j'}(y)) = f_{\varepsilon j}(y) f_{\varepsilon j'}(y) p_j p_{j'} \sum_{r=1}^{k} E\left[\frac{\sigma_r^2(X_{ir})}{\sigma_j(X_{ir})\sigma_{j'}(X_{ir})} \left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j)}{p_j} \right) \right.$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

$$\left(\frac{f_{j'}(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j')}{p_{j'}}\right)\left[1 - \frac{2}{\pi}\right]$$

2. ภายใต้สมมติฐานทางเลือก สำหรับ $j = 1, \ldots, k$

2.1 สถิติทดสอบ $U^1_{KS} \xrightarrow{d} \sum_{j=1}^{k} \sup_y \left|U_j(y) + p_j^{1/2} e_j f_{\varepsilon j}(y)\right|$,

$$U^1_{KS} = \sum_{j=1}^{k} \sup_y \left|\hat{U}_j(y)\right|$$

2.2 สถิติทดสอบ

$$U^1_{KU} \xrightarrow{d} \sum_{j=1}^{k}\left[\sup_y \left|U_j(y) + p_j^{1/2} e_j f_{\varepsilon j}(y)\right| - \inf_y \left|U_j(y) + p_j^{1/2} e_j f_{\varepsilon j}(y)\right|\right],$$

$$U^1_{ku} = \sum_{i=1}^{k}\left[\sup_y \left|\hat{U}_j y\right| - \inf_y \left|\hat{U}_j y\right|\right]$$

2.3 สถิติทดสอบ $U^2_{KS} \xrightarrow{d} \sup_y \left|U(y) + e(y)\right|$, $U^2_{KS} = \sup_y \left|\hat{U}(y)\right|$,

$$U(y) = \sum_{j=1}^{k} p_j^{1/2} U_j(y), \quad e(y) = \sum_{j=1}^{k} p_j^{1/2} f_{\varepsilon j}(y) e_j$$

2.4 สถิติทดสอบ $U^2_{KU} \xrightarrow{d} \left[\sup_y \left|U(y) + e(y)\right| - \inf_y \left|U(y) + e(y)\right|\right]$,

$$U^2_{ku} = \left[\sup_y \left|\hat{U}(y)\right| - \inf_y \left|\hat{U}(y)\right|\right], \quad U(y) = \sum_{j=1}^{k} p_j^{1/2} U_j(y),$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

$$e(y) = \sum_{j=1}^{k} p_j^{1/2} f_{\varepsilon j}(y) e_j$$

เมื่อ $U_j(y) = f_{\varepsilon j}(y) p_j \sum_{r=1}^{k} \sum_{i=1}^{n_r} \frac{Y_{ir} - m(X_{ir})}{\sigma_j(X_{ir})} \left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j)}{p_j} \right)$ คือตัวแปรสุ่มปกติ

ที่มีค่าเฉลี่ยเท่ากับ 0 และมีโครงสร้างความแปรปรวนร่วมคือ

$$Cov(U_j(y), U_{j'}(y)) = f_{\varepsilon j}(y) f_{\varepsilon j'}(y) p_j p_{j'} \sum_{r=1}^{k} E\left[\frac{\sigma_r^2(X_{ir})}{\sigma_j(X_{ir}) \sigma_{j'}(X_{ir})} \left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j)}{p_j} \right) \right.$$

$$\left. \left(\frac{f_{j'}(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j')}{p_{j'}} \right) \right]$$

$$E(y) = (p_1^{1/2} f_{\varepsilon 1}(y) e_1 , \ldots , p_k^{1/2} f_{\varepsilon k}(y) e_k) , \; e_j = E\left[\frac{S(X_{ij}) - s_j(X_{ij})}{\sigma_j(X_{ij}) / \sqrt{n_j}} \right],$$

$$S(u) = \sum_{j=1}^{k} p_j \frac{f_j(u)}{f_{mix}(u)} s_j(u)$$

ส่วนที่ 2 การจำลองแบบ (Simulation)

การทดสอบการเท่ากันของฟังก์ชันการถดถอย 2 ฟังก์ชันและการทดสอบการเท่ากันของฟังก์ชันการถดถอย 3 ฟังก์ชันให้ผลในลักษณะเช่นเดียวกัน กล่าวคือ

1. ภายใต้สมมติฐานหลัก สำหรับ $j = 1, \ldots, k$ พบว่าที่ระดับนัยสำคัญเท่ากับ 0.05 และ 0.10 สถิติทดสอบ $U_{KS}^1, U_{KU}^1, U_{KS}^2, U_{KU}^2$ สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดีเมื่อตัวอย่างมีขนาดใหญ่ในทุกตัวแบบการถดถอย

2. ภายใต้สมมติฐานทางเลือกสำหรับ $j = 1, \ldots, k$ พบว่าที่ระดับนัยสำคัญเท่ากับ 0.05 และ 0.10 เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่อยู่ในรูปของ 1) ฟังก์ชันคงที่ 2) ฟังก์ชันเชิงเส้นตรง 3) ฟังก์ชันเอ็กโพเนนเชียล และ 4) ฟังก์ชันตรีโกณมิติ สำหรับสถิติทดสอบ $U_{KS}^1, U_{KU}^1, U_{KS}^2, U_{KU}^2$ มีค่าอำนาจการทดสอบสูงขึ้นเมื่อ

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

ขนาดของตัวอย่างเพิ่มขึ้น โดยเฉพาะในตัวแบบที่อยู่ในรูปของ ฟังก์ชันเชิงเส้นตรง ฟังก์ชันคงที่ และฟังก์ชันเอ็กโพเนนเชียล ในขณะที่ตัวแบบที่อยู่ในรูปของตัวแบบตรีโกณมิติให้ค่าอำนาจการทดสอบต่ำกว่าตัวแบบอื่น ๆ แต่เมื่อเทียบระหว่างกรณี Constant Shift , Affine Shift และ Quadratic Shift พบว่ากรณี Quadratic Shift ให้ค่าอำนาจการทดสอบสูงที่สุด รองลงมาได้แก่กรณี Affine Shift และกรณี Constant Shift ตามลำดับ

ส่วนที่ 3 การประยุกต์สถิติทดสอบการเท่ากันของ k ฟังก์ชันการถดถอย $k = 3$ กับข้อมูลจริง

ผลการทดสอบการเท่ากันของฟังก์ชันการถดถอย 3 ฟังก์ชัน สำหรับสถิติทดสอบ U^1_{KS} และ U^1_{KU} โดยการประยุกต์ใช้สถิติทดสอบกับข้อมูลจริงให้ผลการทดสอบสอดคล้องกับภาพที่ 4 กล่าวคือ ในภาพรวมเส้นโค้งการถดถอยทั้ง 3 เส้นมีความแตกต่างกัน แต่เมื่อพิจารณาเส้นโค้งการถดถอยแต่ละคู่ พบว่าเส้นโค้งการถดถอยคู่ที่มีความแตกต่างกันคือ คู่ที่ 1) ปี พ.ศ. 2547 และ ปี พ.ศ. 2550 คู่ที่ 2) ปี พ.ศ. 2547 และ ปี พ.ศ. 2549 ในขณะที่เส้นโค้งการถดถอยระหว่างปี พ.ศ. 2549 และ ปี พ.ศ. 2550 ไม่แตกต่างกันที่ระดับนัยสำคัญ 0.05

จากผลการศึกษาทั้งหมดแสดงให้เห็นว่า สถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ ที่พัฒนาขึ้นมีความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดีเมื่อตัวอย่างมีขนาดใหญ่และนอกจากนั้นสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ มีอำนาจการทดสอบสูงเมื่อใช้กับข้อมูลที่มีตัวอย่างขนาดใหญ่ และจะให้ค่าอำนาจการทดสอบสูงในกรณีที่ฟังก์ชันการถดถอยอยู่ในรูปของฟังก์ชันเส้นตรง ฟังก์ชันคงที่ ฟังก์ชันเอ็กโพเนนเชียลตามลำดับ แต่เมื่อใช้ทดสอบกับฟังก์ชันการถดถอยที่อยู่ในรูปแบบฟังก์ชันตรีโกณมิติ จะให้ค่าอำนาจการทดสอบค่อนข้างต่ำ นอกจากนั้นการที่ขนาดตัวอย่างของข้อมูลแต่ละกลุ่มแตกต่างกันไม่ส่งผลต่อความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 และค่าอำนาจการทดสอบของสถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

สรุปและข้อเสนอแนะ

สรุป

จากผลการศึกษาดังที่กล่าวมาแล้วข้างต้นสรุปได้ว่า ภายใต้สมมติฐานหลักเป็นจริง สถิติทดสอบ $U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU}$ เป็นฟังก์ชันของ $\left|U_j(y)\right|$ แต่เนื่องจาก

$$U_j(y) = f_{\varepsilon j}(y)p_j\sum_{r=1}^{k}\sum_{i=1}^{n_r}\frac{Y_{ir} - m(X_{ir})}{\sigma_j(X_{ir})}\left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j)}{p_j}\right)$$ มีการแจกแจงแบบปกติที่

มีค่าเฉลี่ยเท่ากับ 0 และโครงสร้างของความแปรปรวนร่วมคือ

$$Cov(U_j(y),U_{j'}(y)) = f_{\varepsilon j}(y)f_{\varepsilon j'}(y)p_j p_{j'}\sum_{r=1}^{k}E\left[\frac{\sigma_r^2(X_{ir})}{\sigma_j(X_{ir})\sigma_{j'}(X_{ir})}\left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j)}{p_j}\right)\right.$$

$$\left.\left(\frac{f_{j'}(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j')}{p_{j'}}\right)\right]$$

ดังนั้น $\left|U_j(y)\right|$ คือ ตัวแปรสุ่ม Half – Normal ที่มีค่าเฉลี่ย

$$\sqrt{f_{\varepsilon j}(y)f_{\varepsilon j'}(y)p_j p_{j'}\sum_{r=1}^{k}E\left[\frac{\sigma_r^2(X_{ir})}{\sigma_j(X_{ir})\sigma_{j'}(X_{ir})}\left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j)}{p_j}\right)\right.}$$

$$\left.\sqrt{\left(\frac{f_{j'}(X_{ir})}{f_{mix}(X_{ir})} - \frac{I(r=j')}{p_{j'}}\right)}\right]\sqrt{\frac{2}{\pi}}$$

และ มีโครงสร้างความแปรปรวนร่วมคือ

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

$$Cov(U_j(y),U_{j'}(y)) = f_{\varepsilon j}(y)f_{\varepsilon j'}(y)p_j p_{j'}\sum_{r=1}^{k}E\left[\frac{\sigma_r^2(X_{ir})}{\sigma_j(X_{ir})\sigma_{j'}(X_{ir})}\left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})}-\frac{I(r=j)}{p_j}\right)\right.$$

$$\left.\left(\frac{f_{j'}(X_{ir})}{f_{mix}(X_{ir})}-\frac{I(r=j')}{p_{j'}}\right)\right]\left[1-\frac{2}{\pi}\right]$$

หรืออาจกล่าวได้ว่าสถิติทดสอบ $U_{KS}^1, U_{KU}^1, U_{KS}^2, U_{KU}^2$ มีการแจกแจงที่เป็นฟังก์ชันของการแจกแจงแบบ Half - Normal

ภายใต้สมมติฐานทางเลือกเป็นจริง สถิติทดสอบ $U_{KS}^1, U_{KU}^1, U_{KS}^2, U_{KU}^2$ เป็นฟังก์ชันของ $\left|U_j(y) + p_j^{1/2}e_j f_{\varepsilon j}(y)\right|$ โดยที่ $U_j(y)$ มีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเท่ากับ 0 และโครงสร้างของความแปรปรวนร่วมคือ

$$Cov(U_j(y),U_{j'}(y)) = f_{\varepsilon j}(y)f_{\varepsilon j'}(y)p_j p_{j'}\sum_{r=1}^{k}E\left[\frac{\sigma_r^2(X_{ir})}{\sigma_j(X_{ir})\sigma_{j'}(X_{ir})}\left(\frac{f_j(X_{ir})}{f_{mix}(X_{ir})}-\frac{I(r=j)}{p_j}\right)\right.$$

$$\left.\left(\frac{f_{j'}(X_{ir})}{f_{mix}(X_{ir})}-\frac{I(r=j')}{p_{j'}}\right)\right]$$

$p_j = \frac{n_j}{n}$, $e_j = E\left[\frac{S(X_{ij}) - s_j(X_{ij})}{\sigma_j(X_{ij})/\sqrt{n_j}}\right]$, $f_{\varepsilon j}(y)$ คือฟังก์ชันความหนาแน่นน่าจะเป็นของความคลาดเคลื่อนกลุ่มที่ j, $f_{\varepsilon j}(y) = F'_{\varepsilon j}(y)$

เมื่อพิจารณาในแง่ของเปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานหลักและเปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกพบว่า ภายใต้สมมติฐานหลัก สำหรับ

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

$j = 1, \ldots, k$ เมื่อกำหนดระดับนัยสำคัญเท่ากับ 0.05 และ 0.10 สถิติทดสอบ U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU} สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดีเมื่อตัวอย่างมีขนาดใหญ่ในทุกตัวแบบการถดถอย ในขณะเดียวกันภายใต้สมมติฐานทางเลือก สำหรับ $j = 1, \ldots, k$ เปอร์เซ็นต์การปฏิเสธสมมติฐานภายใต้สมมติฐานทางเลือกของตัวแบบที่อยู่ในรูปของ 1) ฟังก์ชันคงที่ 2) ฟังก์ชันเชิงเส้นตรง 3) ฟังก์ชันเอ็กโพเนนเชียล และ 4) ฟังก์ชันตรีโกณมิติ สำหรับสถิติทดสอบ U^1_{KS}, U^1_{KU}, U^2_{KS}, U^2_{KU} จะมีค่าอำนาจการทดสอบสูงขึ้นเมื่อขนาดของตัวอย่างเพิ่มขึ้น โดยเฉพาะในตัวแบบที่อยู่ในรูปของ ฟังก์ชันเชิงเส้นตรง ฟังก์ชันคงที่ และฟังก์ชันเอ็กโพเนนเชียล ในขณะที่ตัวแบบที่อยู่ในรูปของตัวแบบตรีโกณมิติให้ค่าอำนาจการทดสอบต่ำกว่าตัวแบบอื่น ๆ แต่เมื่อเปรียบเทียบระหว่างกรณี Constant Shift , Affine Shift และ Quadratic Shift พบว่ากรณี Quadratic Shift ให้ค่าอำนาจการทดสอบสูงที่สุด รองลงมาได้แก่กรณี Affine Shift และกรณี Constant Shift ตามลำดับ

ข้อเสนอแนะ

1. ในการศึกษาในครั้งนี้เป็นการพัฒนาสถิติเพื่อใช้ในการทดสอบสมมติฐาน $H_0: m_1 = m_2 = m_3 = \ldots = m_k$ เมื่อ $k \geq 2$ และ $H_1: m_i \neq m_j$ สำหรับ i, j บางค่า $i \neq j$, $i, j \in \{1, \ldots, k\}$, $k \geq 2$ โดยกำหนดตัวแบบการถดถอยอยู่ในรูปของ $Y_{ij} = m_j(X_{ij}) + \sigma_j(X_{ij})\varepsilon_{ij}$, $i = 1, \ldots, n_j$ และ $j = 1, \ldots, k$ โดยที่ $m_j(X_{ij})$ คือฟังก์ชันการถดถอยแบบไม่อิงพารามิเตอร์ ซึ่งเป็นฟังก์ชันการถดถอยที่ไม่ทราบรูปแบบของพารามิเตอร์ในฟังก์ชันการถดถอย โดยประมาณค่าฟังก์ชันการถดถอยด้วยวิธีการของ Nadaraya – Watson Estimator ดังนั้นจึงควรประมาณฟังก์ชันการถดถอยด้วยวิธีการอื่น เพื่อเปรียบเทียบว่ามีความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 ตลอดจนให้ค่าอำนาจการทดสอบเหมือนหรือแตกต่างกันอย่างไร

2. ในการศึกษาในครั้งนี้ ได้มีการกำหนดวิธีการเลือกพารามิเตอร์ปรับให้เรียบ h_j โดยกำหนดให้ $h_j = cn_j^{-3/10}$ เมื่อ $c = 1$ เพียงกรณีเดียว ดังนั้นในการศึกษาต่อไปจึงควรมีการกำหนดรูปแบบพารามิเตอร์ปรับให้เรียบในรูปแบบอื่น เพื่อเปรียบเทียบว่ารูปแบบพารามิเตอร์ปรับให้เรียบแต่ละรูปแบบ ส่งผลต่อความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 ตลอดจนให้ค่าอำนาจการทดสอบเหมือนหรือแตกต่างกันอย่างไร

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

เอกสารและสิ่งอ้างอิง

สำนักงานสถิติแห่งชาติ. 2552. **การสำรวจภาวะเศรษฐกิจและสังคมของครัวเรือน พ.ศ. 2545 2547 2549 และ 2550**. แหล่งที่มา: http://www. nso.go.th/nso/nsopublish/download/files /ictDev53.pdf, 17 มิถุนายน 2553.

Akritas, M.G. and I. Van Keilegom. 2001. Non-Parametric Estimation of the Residual Distribution. **Scand.J.Statist**. 28 : 549 – 567.

Brame, R., R. Paternoster, P. Mazerolle and A. Piquero. 1998. Testing for the Equality of Maximum-Likelihood Regression Coefficients Between Two Independent Equations. **Journal of Quantitative Criminology**. 14: 245 – 261.

Brame, R., R. Paternoster, P. Mazerolle and A. Piquero. 1998. Testing for the Equality of Maximum-Likelihood Regression Coefficients Between Two Independent Equations. **Journal of Quantitative Criminology**. 14: 245 – 261. *Cited* N.L. Johnson and S. Kotz. 1970. **Continuous Univariate Distribution – I (Distributions in Statistics)**. John Wiley & sons, Inc., New York.

Brame, R., R. Paternoster, P. Mazerolle and A. Piquero. 1998. Testing for the Equality of Maximum-Likelihood Regression Coefficients Between Two Independent Equations. **Journal of Quantitative Criminology**. 14: 245 – 261. *Cited* H. Koo And D.S. Hong. 1980. Class and income inequality in Korea. **Am. Sociol. Rev**. 45: 610 – 626.

Delgado, M.A. 1993. Testing the Equality of Nonparametric Regression Curves. **Statistics & Probability Letters**. 17: 199 – 204.

Hardle, W. 1993. **Applied Nonparametric Regression.** Cambridge University Press., United Kingdom.

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

Hegazy, A.S. and J.R. Green. 1975. **Some New Goodness-of-fit tests Using Order Statistics. Applied Statistics**. 24 : 299 – 308.

Keselman, H.J., G. Robert and H. Burt. 2002. Controlling the Rate of Type I Error Over a Large Set of Statistical tests. **British Journal of Mathematical and Statistical Psychology**. 55 : 27 – 39.

Koul, L.H. and A. Schick. 1997. Testing for the Equality of Two Nonparametric Regression Curves. **Journal of Statistical planning and Inference**. 65: 293 – 314.

Munk, A. and H. Dette. 1998. Nonparametric Comparison of Several Regression Function : Exact nd Asymptotic Theory. **Annals of Statistics**. 26: 2339 – 2368.

Neumeyer, N. and H. Dette. 2003. Nonparametric Comparison of Regression Curves : An Empirical Process Approach. **Annals of Statistics**. 31: 880 – 920.

Rardo-Fernandez, J.C. 2007. Comparison of Error Distributions in Nonparametric Regression. **Statistics &Probability**. 77 : 350 – 356.

Serfling, R.J. 1980. **Approximation of Mathematical Statistics**. John Wiley & sons, Inc., New York.

Stone, C.J. 1977. Consistent Nonparametric Regression. **The annals of statistics.** 5: 595 – 620.

Takezawa, K.1999. **Introduction to Nonparametric Regression.** John Wiley & sons, Inc., New York.

Van der Vaart, A.W. and J.A. Wellner. 1996. **Weak Convergence and Empirical Process**. Springer, Inc., New York.

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

Young, S. and A.W. Bowman. 1995. Nonparametric Analysis of Covariance. **Biometrics**. 51: 920 – 931.

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

ภาคผนวก

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

ภาคผนวก ก

ความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

ความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1

การศึกษาความสามารถในการควบคุมระดับนัยสำคัญของสถิติหรือการควบคุมความคลาดเคลื่อนประเภทที่ 1 ใช้ตัวสถิติทดสอบ Z ในการทดสอบสมมติฐานแบบ 2 ด้าน (Keselman *et.al.*, 2002)

กำหนด α แทน ความคลาดเคลื่อนประเภทที่ 1 ที่เกิดขึ้นจริง

$\hat{\alpha}$ แทน ค่าประมาณของความคลาดเคลื่อนประเภทที่ 1 ที่เกิดขึ้นจริง

α_0 แทน ความคลาดเคลื่อนประเภทที่ 1 ที่กำหนด ในการศึกษานี้กำหนดความคลาดเคลื่อนประเภทที่ 1 2 ระดับคือ 0.05 และ 0.10

n แทน จำนวนตัวอย่างที่ใช้ เท่ากับ 1,000

สมมติฐานของการทดสอบคือ

$H_0: \alpha = \alpha_0$

$H_1: \alpha \neq \alpha_0$

สถิติที่ใช้ทดสอบคือ

$$Z = \frac{\hat{\alpha} - \alpha_0}{\sqrt{\frac{\alpha_0(1 - \alpha_0)}{n}}}$$

1. เมื่อกำหนดความคลาดเคลื่อนประเภทที่ 1 ของการทดสอบเป็น 0.05 จะยอมรับสมมติฐานหลักถ้า $\hat{\alpha}$ อยู่ในช่วง

$$\hat{\alpha} \leq 0.05 + Z_{\frac{\alpha}{2}}\sqrt{\frac{(0.05)(0.95)}{1000}} \quad \text{และ} \quad \hat{\alpha} \geq 0.05 - Z_{\frac{\alpha}{2}}\sqrt{\frac{(0.05)(0.95)}{1000}}$$

$$\hat{\alpha} \leq 0.05 + 1.96\sqrt{\frac{(0.05)(0.95)}{1000}} \quad \text{และ} \quad \hat{\alpha} \geq 0.05 - 1.96\sqrt{\frac{(0.05)(0.95)}{1000}}$$

$$\hat{\alpha} \leq 0.063 \qquad \text{และ} \qquad \hat{\alpha} \geq 0.036$$

$$\text{ดังนั้น} \quad \hat{\alpha} \in [0.036,\ 0.063]$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

2. เมื่อกำหนดความคลาดเคลื่อนประเภทที่ 1 ของการทดสอบเป็น 0.10 จะยอมรับสมมติฐานหลักถ้า $\hat{\alpha}$ อยู่ในช่วง

$$\hat{\alpha} \leq 0.10 + Z_{\frac{\alpha}{2}}\sqrt{\frac{(0.10)(0.90)}{1000}} \quad \text{และ} \quad \hat{\alpha} \geq 0.10 - Z_{\frac{\alpha}{2}}\sqrt{\frac{(0.10)(0.90)}{1000}}$$

$$\hat{\alpha} \leq 0.10 + 1.645\sqrt{\frac{(0.10)(0.90)}{1000}} \quad \text{และ} \quad \hat{\alpha} \geq 0.10 - 1.645\sqrt{\frac{(0.10)(0.90)}{1000}}$$

$$\hat{\alpha} \leq 0.115 \qquad \text{และ} \qquad \hat{\alpha} \geq 0.084$$

$$\text{ดังนั้น} \quad \hat{\alpha} \in [0.084,\ 0.115]$$

สถิติทดสอบใดที่มีค่าความคลาดเคลื่อนประเภทที่ 1 อยู่ในช่วงดังกล่าวจะถือว่าสถิติทดสอบนั้นสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

ภาคผนวก ข

การเขียนโปรแกรม

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

การเขียนโปรแกรม

ทดสอบการเท่ากันของฟังก์ชันการถดถอย 2 ฟังก์ชันโดยใช้สถิติทดสอบ U^{1}_{KS}

```
title 'Generate distribution of  y1 = m(x)=2X+0.25e1 ,
y2=m(x)=2X+0.5e2,  n=20';
%macro boott(trial,numberobs);
%do i = 1 %to &trial;
%let t=1;
%let a1=20**-0.3;
%let numObs=20;
%let seed2=2345;
%let seed3=3456;
%let seed4=4567;
%let seed5=5678;
%let seed6=6789;
 data one&i;
Do i = 1 to 20;
v1=rannor(-1);
z1=rannor(&seed2);
x1=ranuni(&seed3);
e1=((1-(&a1**2))**0.5)*v1+(&a1*z1);
y1=(2*x1)+(0.25*e1);
v2=rannor(&seed4);
z2=rannor(&seed5);
x2=ranuni(&seed6);
e2=((1-(&a1**2))**0.5)*v2+(&a1*z2);
y2=(2*x2) +(0.50*e2);
output;
end;
keep i x1 x2 y1 y2 e1 e2;
proc print data=one&i;
run;
proc means data=one&i;
var y1 y2;
run;
%end;
run;
%mend;
%boott(1000,20);
proc insight tools data=one1;
run;
option ps=65 ls=80 nodate;
DM 'log;clear;output;clear';
ods trace on;
data empi;
input fe1 fe2 feo1 feo2;
datalines;

⋮       ⋮          ⋮          ⋮
;
proc print data=empi;
```

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

```
run;
proc rank
      data=empi
      out=order   /* highly recommended */
      fraction
      ties=high;
   var fe1 fe2 feo1 feo2;
   ranks fe1Rank fe2Rank feo1Rank feo2Rank ;
run;
data order;
set order;
a = feo1Rank-fe1Rank;
b = feo2Rank-fe2Rank;
run;
data order;
set order;
c=ABS(a);
d=ABS(b);
run;
proc univariate data=order;
var c;
Output out=em1ma max=MAX;
run;
proc univariate data=order;
var d;
Output out=em2ma max=MAX;
run;
data em1ma;
set em1ma;
e=(20**(1/2))*(max);
run;
data em2ma;
set em2ma;
f=(20**(1/2))*(max);
run;
data em12ma;
merge em1ma em2ma;
run;
data em12ma;
set em12ma;
g=e+f;
run;
/*Macro program of bootstrap for empirical distribution function of
residual*/
%macro bootT(num=);
/*DM 'log;clear;output;clear';*/
/*-----data file for bootstrap-----*/
PROC SURVEYSELECT DATA=empi METHOD=urs
n=20 REP=&num seed=12345 OUT=empi1;
 run;
 /**proc print data=empi1;**/
 /**run;**/
data empi2(DROP=i);
set empi1;
DO i =1 TO numberhits;
output;
end;
```

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

```
run;
/**proc print data=empi2;**/
 /**run;**/
%do i=1 %to &num;
data set&i;
set empi2;
where replicate=&i;
%end;
%do i=1 %to &num;
proc rank
      data=set&i
      out=order&i /* highly recommended */
      fraction
      ties=high;
   var fe1 fe2 feo1 feo2;
   ranks fe1Rank fe2Rank feo1Rank feo2Rank ;
run;
data order&i;
set order&i;
a = feo1Rank-fe1Rank;
b = feo2Rank-fe2Rank;
run;
data order&i;
set order&i;
c=ABS(a);
d=ABS(b);
run;
proc univariate data=order&i;
var c;
Output out=em1ma&i max=MAX;
run;
proc univariate data=order&i;
var d;
Output out=em2ma&i max=MAX;
run;
data em1ma&i;
set em1ma&i;
e=(20**(1/2))*(max);
run;
data em2ma&i;
set em2ma&i;
f=(20**(1/2))*(max);
run;
data em12ma&i;
merge em1ma&i em2ma&i;
run;
data em12ma&i;
set em12ma&i;
g=e+f;
run;
%end;
data outall;
set
%do i=1 %to &num;
em12ma&i
%end;
;
```

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

```
run;
data em12ma;
set em12ma;
h=g;
run;
data mergeoutall;
merge outall em12ma;
run;
data mergeoutall;
set mergeoutall;
h=5.8137767415;
run;
data mergeoutallnew;
set mergeoutall;
if g>h then g=1;
else  g=0;
run;
proc print data=mergeoutallnew;
proc freq;  tables g;
where g=1;
run;
%mend;
%boott(num=200);
```

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

ทดสอบการเท่ากันของฟังก์ชันการถดถอย 2 ฟังก์ชันโดยใช้สถิติทดสอบ U^{1}_{KU}

```
option ps=65 ls=80 nodate;
DM 'log;clear;output;clear';
ods trace on;
data empi;
input fe1 fe2 feo1 feo2;
datalines;
⋮       ⋮          ⋮          ⋮
;
proc print data=empi;
run;
proc rank
      data=empi
      out=order   /* highly recommended */
      fraction
      ties=high;
   var fe1 fe2 feo1 feo2;
   ranks fe1Rank fe2Rank feo1Rank feo2Rank ;
run;
data order;
set order;
a = feo1Rank-fe1Rank;
b = feo2Rank-fe2Rank;
run;
data order;
set order;
c=ABS(a);
d=ABS(b);
run;
proc univariate data=order;
var c;
Output out=em1ma max=MAX min=MIN;
run;
proc univariate data=order;
var d;
Output out=em2ma max=MAX min=MIN ;
run;
data em1ma;
set em1ma;
e=(20**(1/2))*(max-min);
run;
data em2ma;
set em2ma;
f=(20**(1/2))*(max-min);
run;
data em12ma;
merge em1ma em2ma;
run;
data em12ma;
set em12ma;
g=e+f;
run;
data em12mb;
merge em1ma em2ma;
run;
```

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

```
data em12mb;
set em12mb;
g=e+f;
run;
data em12mb;
set em12mb;
m=e+f;
run;
data em12mbnew;
merge  em12mb em12ma;
run;
/*Macro program of bootstrap for empirical distribution function of
residual*/
%macro bootт(num=);
/*DM 'log;clear;output;clear';*/
/*-----data file for bootstrap-----*/
PROC SURVEYSELECT DATA=empi METHOD=urs
n=20 REP=&num seed=12345 OUT=empi1;
 run;
 /**proc print data=empi1;**/
 /**run;**/
data empi2(DROP=i);
set empi1;
DO i =1 TO numberhits;
output;
end;
run;
/**proc print data=empi2;**/
 /**run;**/
%do i=1 %to &num;
data set&i;
set empi2;
where replicate=&i;
%end;
%do i=1 %to &num;
proc rank
      data=set&i
      out=order&i /* highly recommended */
      fraction
      ties=high;
   var fe1 fe2 feo1 feo2;
   ranks fe1Rank fe2Rank feo1Rank feo2Rank ;
run;
data order&i;
set order&i;
a = feo1Rank-fe1Rank;
b = feo2Rank-fe2Rank;
run;
data order&i;
set order&i;
c=ABS(a);
d=ABS(b);
run;
proc univariate data=order&i;
var c;
Output out=em1ma&i max=MAX min = MIN;
run;
```

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

```
proc univariate data=order&i;
var d;
Output out=em2ma&i max=MAX min = MIN;
run;
data em1ma&i;
set em1ma&i;
e=(20**(1/2))*(max-min);
run;
data em2ma&i;
set em2ma&i;
f=(20**(1/2))*(max-min);
run;
data em12ma&i;
merge em1ma&i em2ma&i;
run;
data em12ma&i;
set em12ma&i;
g=e+f;
run;
data em12mbnew
merge em12ma em12mb;
run;
%end;
data outall;
set
%do i=1 %to &num;
em12ma&i
%end;
;
run;
data em12ma;
set em12ma;
h=g;
run;
data mergeoutall;
merge outall em12ma;
run;
data mergeoutall;
set mergeoutall;
h=13.191;
run;
data mergeoutallnew;
set mergeoutall;
if g>h then g=1;
else  g=0;
run;
proc print data=mergeoutallnew;
proc freq;  tables g;
where g=1;
run;
%mend;
%boott(num=1000);
```

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

การประยุกต์สถิติทดสอบ U^1_{KS} และ U^1_{KU} กับข้อมูลจริง

สถิติทดสอบ U^1_{KS} (overall test p-value 0.10)

```
option ps=65 ls=80 nodate;
DM 'log;clear;output;clear';
ods trace on;
data empi;
input fe1 feo1 fe2 feo2 fe3 feo3;
datalines;

-4246.43133 -108702.962 -17.62317275 -476.4807154   -13.84455833   -359.5348675
⋮           ⋮           ⋮            ⋮             ⋮              ⋮
-29.62867424 -1838.898856 -1.413770857 -106.8281027 -11.6681134    -699.6150028;

proc print data=empi;
run;
proc rank
      data=empi
      out=order   /* highly recommended */
      fraction
      ties=high;
   var fe1 feo1 fe2 feo2 fe3 feo3;
   ranks fe1Rank fe2Rank fe3Rank feo1Rank feo2Rank feo3Rank  ;
run;
data order;
set order;
a = feo1Rank-fe1Rank;
b = feo2Rank-fe2Rank;
c = feo3Rank-fe3Rank;
run;
data order;
set order;
d=ABS(a);
e=ABS(b);
f=ABS(c);
run;
proc univariate data=order;
var d;
Output out=em1ma max=MAX;
run;
proc univariate data=order;
var e;
Output out=em2ma max=MAX;
run;
proc univariate data=order;
var f;
Output out=em3ma max=MAX;
run;
data em1ma;
set em1ma;
g=(76**(1/2))*(max);
run;
data em2ma;
```

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

```
set em2ma;
h=(76**(1/2))*(max);
run;
data em3ma;
set em3ma;
i=(76**(1/2))*(max);
run;
data em123ma;
merge em1ma em2ma em3ma;
run;
data em123ma;
set em123ma;
j=g+h+i;
run;
/*Macro program of bootstrap for empirical distribution function of
residual*/
%macro bootT(num=);
/*DM 'log;clear;output;clear';*/
/*-----data file for bootstrap-----*/
PROC SURVEYSELECT DATA=empi METHOD=urs
n=76 REP=&num seed=12345 OUT=empi1;
 run;
 /**proc print data=empi1;**/
 /**run;**/
data empi2(DROP=i);
set empi1;
DO i =1 TO numberhits;
output;
end;
run;
/**proc print data=empi2;**/
 /**run;**/
%do i=1 %to &num;
data set&i;
set empi2;
where replicate=&i;
%end;
%do i=1 %to &num;
proc rank
      data=set&i
      out=order&i /* highly recommended */
      fraction
      ties=high;
   var fe1 fe2 fe3 feo1 feo2 feo3;
   ranks fe1Rank fe2Rank fe3Rank feo1Rank feo2Rank feo3Rank  ;
run;
data order&i;
set order&i;
a = feo1Rank-fe1Rank;
b = feo2Rank-fe2Rank;
c = feo3Rank-fe3Rank;
run;
data order&i;
set order&i;
d=ABS(a);
e=ABS(b);
f=ABS(c);
```

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

```
run;
proc univariate data=order&i;
var d;
Output out=em1ma&i max=MAX;
run;
proc univariate data=order&i;
var e;
Output out=em2ma&i max=MAX;
run;
proc univariate data=order&i;
var f;
Output out=em3ma&i max=MAX;
run;
data em1ma&i;
set em1ma&i;
g=(76**(1/2))*(max);
run;
data em2ma&i;
set em2ma&i;
h=(76**(1/2))*(max);
run;
data em3ma&i;
set em3ma&i;
i=(76**(1/2))*(max);
run;
data em123ma&i;
merge em1ma&i em2ma&i em3ma&i ;
run;
data em123ma&i;
set em123ma&i;
j=g+h+i;;
run;
%end;
data outall;
set
%do i=1 %to &num;
em123ma&i
%end;
;
run;
data em123ma;
set em123ma;
k=j;
run;
data mergeoutall;
merge outall em123ma;
run;
data mergeoutall;
set mergeoutall;
k=20.9915;
run;
data mergeoutallnew;
set mergeoutall;
if j>k then j=1;
else  j=0;
run;
proc print data=mergeoutallnew;
proc freq;  tables j;
```

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

```
where j=1;
run;
%mend;
%boott(num=1000);
```


ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

สถิติทดสอบ U^{1}_{KU} (overall test p-value 0.10)

```
option ps=65 ls=80 nodate;
DM 'log;clear;output;clear';
ods trace on;
data empi;
input fe1 feo1 fe2 feo2 fe3 feo3;
datalines;

-4246.43133 -108702.962 -17.62317275  -476.4807154   -13.84455833   -359.5348675
⋮           ⋮           ⋮             ⋮             ⋮              ⋮
-29.62867424 -1838.898856 -1.413770857 -106.8281027  -11.6681134    -699.6150028;
;
proc print data=empi;
run;
proc rank
      data=empi
      out=order   /* highly recommended */
      fraction
      ties=high;
   var fe1 fe2 fe3 feo1 feo2  feo3;
   ranks fe1Rank fe2Rank fe3Rank feo1Rank feo2Rank feo3Rank  ;
run;
data order;
set order;
a = feo1Rank-fe1Rank;
b = feo2Rank-fe2Rank;
c = feo3Rank-fe3Rank;
run;
data order;
set order;
d=ABS(a);
e=ABS(b);
f=ABS(c);
run;
proc univariate data=order;
var d;
Output out=em1ma max=MAX min=MIN;
run;
proc univariate data=order;
var e;
Output out=em2ma max=MAX min=MIN;
run;
proc univariate data=order;
var f;
Output out=em3ma max=MAX min=MIN;
run;
data em1ma;
set em1ma;
g=(76**(1/2))*(max-min);
run;
data em2ma;
set em2ma;
h=(76**(1/2))*(max-min);
run;
data em3ma;
```

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

```
set em3ma;
i=(76**(1/2))*(max-min);
run;
data em123ma;
merge em1ma em2ma em3ma;
run;
data em123ma;
set em123ma;
j=g+h+i;
run;
/*Macro program of bootstrap for empirical distribution function of
residual*/
%macro bootT(num=);
/*DM 'log;clear;output;clear';*/
/*-----data file for bootstrap-----*/
PROC SURVEYSELECT DATA=empi METHOD=urs
n=76 REP=&num seed=12345 OUT=empi1;
 run;
 /**proc print data=empi1;**/
 /**run;**/
data empi2(DROP=i);
set empi1;
DO i =1 TO numberhits;
output;
end;
run;
/**proc print data=empi2;**/
 /**run;**/
%do i=1 %to &num;
data set&i;
set empi2;
where replicate=&i;
%end;
%do i=1 %to &num;
proc rank
      data=set&i
      out=order&i /* highly recommended */
      fraction
      ties=high;
   var fe1 fe2 fe3 feo1 feo2 feo3;
   ranks fe1Rank fe2Rank fe3Rank feo1Rank feo2Rank feo3Rank  ;
run;
data order&i;
set order&i;
a = feo1Rank-fe1Rank;
b = feo2Rank-fe2Rank;
c = feo3Rank-fe3Rank;
run;
data order&i;
set order&i;
d=ABS(a);
e=ABS(b);
f=ABS(c);
run;
proc univariate data=order&i;
var d;
Output out=em1ma&i max=MAX min=MIN;
```

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

```
run;
proc univariate data=order&i;
var e;
Output out=em2ma&i max=MAX min=MIN;
run;
proc univariate data=order&i;
var f;
Output out=em3ma&i max=MAX min=MIN;
run;
data em1ma&i;
set em1ma&i;
g=(76**(1/2))*(max-min);
run;
data em2ma&i;
set em2ma&i;
h=(76**(1/2))*(max-min);
run;
data em3ma&i;
set em3ma&i;
i=(76**(1/2))*(max-min);
run;
data em123ma&i;
merge em1ma&i em2ma&i em3ma&i ;
run;
data em123ma&i;
set em123ma&i;
j=g+h+i;;
run;
%end;
data outall;
set
%do i=1 %to &num;
em123ma&i
%end;
;
run;
data em123ma;
set em123ma;
k=j;
run;
data mergeoutall;
merge outall em123ma;
run;
data mergeoutall;
set mergeoutall;
k=20.9915;
run;
data mergeoutallnew;
set mergeoutall;
if j>k then j=1;
else  j=0;
run;
proc print data=mergeoutallnew;
proc freq;  tables j;
where j=1;
run;
%mend;
%boott(num=1000);
```

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

ประวัติการศึกษา และการทำงาน

ชื่อ –นามสกุล	นางสาวอัญชลี ทองกำเหนิด
วัน เดือน ปี ที่เกิด	26 มกราคม 2520
สถานที่เกิด	78 หมู่ 7 ต. พิกุลออก อ.บ้านนา จ. นครนายก 26110
ประวัติการศึกษา	- พ.ศ. 2542 ปริญญาตรี กศ.บ. (วิทยาศาสตร์-คณิตศาสตร์) เกียรตินิยมอันดับ 1 มหาวิทยาลัยบูรพา - พ.ศ. 2545 ปริญญาโท วท.ม.(สถิติ) มหาวิทยาลัยเกษตรศาสตร์ - พ.ศ. 2554 ปร.ด. (สถิติ) มหาวิทยาลัยเกษตรศาสตร์
ตำแหน่งหน้าที่การงานปัจจุบัน	อาจารย์ระดับ 6 ประจำสาขาวิชาสถิติประยุกต์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
สถานที่ทำงานปัจจุบัน	สาขาวิชาสถิติประยุกต์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี ถ. รังสิต – นครนายก อ. ธัญบุรี จ. ปทุมธานี 12110
ผลงานดีเด่นและรางวัลทางวิชาการ	-
ทุนการศึกษาที่ได้รับ	-

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์