การทดสอบพารามิเตอร์การกระจายของตัวแบบการถดถอยปัวซงนัยทั่วไป

โดย
นางสาวเมษิยา แย้มเจริญกิจ

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาสถิติประยุกต์
ภาควิชาสถิติ
บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร
ปีการศึกษา 2554

การทดสอบพารามิเตอร์การกระจายของตัวแบบการถดถอยปัวซงนัยทั่วไป

โดย
นางสาวเมษิยา แย้มเจริญกิจ

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาสถิติประยุกต์
ภาควิชาสถิติ
บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร
ปีการศึกษา 2554

TESTS OF DISPERSION PARAMETER IN GENERALIZED POISSON REGRESSION MODELS

By

Maysiya Yamjaroenkit

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

Department of Statistics

Graduate School

SILPAKORN UNIVERSITY

2011

บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร อนุมัติให้วิทยานิพนธ์เรื่อง " การทดสอบพารามิเตอร์การกระจายของตัวแบบการถดถอยปัวซงนัยทั่วไป " เสนอโดย นางสาวเมษิยา แย้มเจริญกิจ เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาสถิติประยุกต์

……..
(ผู้ช่วยศาสตราจารย์ ดร.ปานใจ ธารทัศนวงศ์)
คณบดีบัณฑิตวิทยาลัย
วันที่.........เดือน................... พ.ศ..........

อาจารย์ที่ปรึกษาวิทยานิพนธ์

รองศาสตราจารย์วีรานันท์ พงศาภักดี

คณะกรรมการตรวจสอบวิทยานิพนธ์

.... ประธานกรรมการ
(อาจารย์ ดร.ไพโรจน์ ขาวสิทธิวงษ์)
............/......................../.............

.... กรรมการ
(รองศาสตราจารย์ ดร.กมล บุษบา)
............/......................../.............

.... กรรมการ
(รองศาสตราจารย์วีรานันท์ พงศาภักดี)
............/......................../.............

52304203 : สาขาวิชาสถิติประยุกต์

คำสำคัญ : ตัวแบบการถดถอยปัวซง/ ตัวแบบการถดถอยปัวซงนัยทั่วไป/ การทดสอบภาวะสารูปดี/ OVERDISPERSION/ UNDERDISPERSION

เมษิยา แย้มเจริญกิจ : การทดสอบพารามิเตอร์การกระจายของตัวแบบการถดถอยปัวซงนัยทั่วไป. อาจารย์ที่ปรึกษาวิทยานิพนธ์ : รศ.วีรานันท์ พงศาภักดี. 95 หน้า.

งานวิจัยนี้เสนอการทดสอบภาวะสารูปดี 2 วิธีเรียกว่า การทดสอบ $Z_{\bar{\mu}}$ และการทดสอบ $Z_{\bar{Y}}$ ที่มีการแจกแจงค่อนข้างสมมาตร เพื่อใช้ในการทดสอบ Overdispersion และ Underdispersion ในตัวแบบการถดถอยปัวซงเทียบกับตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 และเปรียบเทียบวิธีการทดสอบต่าง ๆ ได้แก่ การทดสอบ Z_0 การทดสอบวาลด์ที การทดสอบอัตราส่วนภาวะน่าจะเป็น การทดสอบสกอร์ และการทดสอบ Q^* โดยจำลองแบบข้อมูล 5,000 ชุดในแต่ละเงื่อนไขของการทดลอง ภายใต้ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 และเปรียบเทียบกำลังการทดสอบที่ได้จากการทดลองของการทดสอบทั้งหมด 7 วิธี โดยเปรียบเทียบในสถานการณ์ที่แตกต่างกันของขนาดตัวอย่างและพารามิเตอร์ของการกระจาย

ผลการวิจัย ในกรณี Overdispersion การทดสอบ $Z_{\bar{\mu}}$ และการทดสอบ $Z_{\bar{Y}}$ ที่ผู้วิจัยเสนอขึ้นให้ผลการทดสอบที่เด่นกว่าการทดสอบอื่น ๆ เมื่อพิจารณาในเทอมของกำลังการทดสอบ แต่การทดสอบ $Z_{\bar{Y}}$ มีรูปแบบที่ง่ายกว่าการทดสอบ $Z_{\bar{\mu}}$ ทั้งสองวิธีให้ผลใกล้เคียงกัน ดังนั้น การทดสอบ $Z_{\bar{Y}}$ จึงเหมาะสมมากกว่าเมื่อนำไปใช้ในการทดสอบ Overdispersion ในตัวแบบการถดถอยปัวซงโดยเทียบกับ ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 ในกรณี Underdispersion เมื่อตัวอย่างมีขนาดเล็ก การทดสอบสกอร์เหมาะสมกว่าการทดสอบอื่น ๆ แต่เมื่อตัวอย่างมีขนาดใหญ่ การทดสอบ $Z_{\bar{\mu}}$ และการทดสอบ $Z_{\bar{Y}}$ ยังให้กำลังการทดสอบสูงและใกล้เคียงกับการทดสอบสกอร์ แต่การทดสอบสกอร์ให้กำลังการทดสอบที่ผิดปกติในบางกรณี และเนื่องจากการทดสอบ $Z_{\bar{Y}}$ มีรูปแบบที่ง่ายกว่าการทดสอบ $Z_{\bar{\mu}}$ ดังนั้น เมื่อตัวอย่างมีขนาดใหญ่ การทดสอบ $Z_{\bar{Y}}$ จึงเหมาะสมมากกว่าเมื่อนำไปใช้ในการทดสอบ Underdispersion ในตัวแบบการถดถอยปัวซงโดยเทียบกับตัวแบบปัวซงนัยทั่วไปแบบที่ 2

นอกจากนั้นยังพบว่า ขนาดตัวอย่างมีอิทธิพลต่อกำลังการทดสอบอย่างชัดเจนเมื่อตัวอย่างมีขนาดใหญ่ขึ้นกำลังการทดสอบจะเพิ่มขึ้นด้วย โดยพบว่า เมื่อตัวอย่างมีขนาดใหญ่ กำลังการทดสอบจะเพิ่มขึ้น อย่างรวดเร็วและเข้าใกล้ 1.0 แต่เมื่อตัวอย่างมีขนาดเล็ก กำลังการทดสอบจะเพิ่มขึ้นอย่างช้าๆในช่วงแรก และต่อมากำลังการทดสอบจะเพิ่มขึ้นอย่างรวดเร็วและเข้าใกล้ 1.0

ภาควิชาสถิติ บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร ปีการศึกษา 2554

ลายมือชื่อนักศึกษา..

ลายมือชื่ออาจารย์ที่ปรึกษาวิทยานิพนธ์ ....

52304203 : MAJOR : APPLIED STATISTICS
KEY WORD : POISSON MODEL/GENERALIZED POISSON REGRESSION MODEL/ GOODNESS-OF-FIT TESTS/ OVERDISPERSION/ UNDERDISPERSION

MAYSIYA YAMJAROENKIT : TESTS OF DISPERSION PARAMETER IN GENERALIZED POISSON REGRESSION MODELS. THESIS ADVISOR : ASSOC.PROF. VEERANUN PONGSAPUKDEE. 95 pp.

This research proposes two quite symmetric distributed goodness-of-fit tests called the $Z_{\bar{\hat{\mu}}}$ test and the $Z_{\bar{Y}}$ test for overdispersion and underdispersion in Poisson regression models versus generalized Poisson regression model type II, and compares these tests with the Z_0 test, Wald-t test, Likelihood ratio test, Score test and Q^* test. In this simulation study, 5,000 sets of data in each condition of study under generalized Poisson regression model type II are simulated and these seven goodness-of-fit tests based on the power of the tests under different situations of sample sizes and dispersion parameters are compared.

The research results show that for overdispersion case the proposed $Z_{\bar{\hat{\mu}}}$ test and $Z_{\bar{Y}}$ test dominate uniformly over other tests in term of power but the $Z_{\bar{Y}}$ test is more simple than $Z_{\bar{\hat{\mu}}}$ test. Therefore, In all overdispersion test, the $Z_{\bar{Y}}$ test is most appropriate for general application to detect overdispersion in Poisson regression model versus generalized Poisson regression model type II. For underdispersion case, when the sample size is small, the score test has advantage over other tests. When the sample size is large, the power of the test from the $Z_{\bar{\hat{\mu}}}$ test and $Z_{\bar{Y}}$ test are approaching to those from the score test. However, the score test has outliers in some cases and that the $Z_{\bar{Y}}$ test is more simple than $Z_{\bar{\hat{\mu}}}$ test. Thus, it's confirmed that the $Z_{\bar{Y}}$ test is most appropriate for general applications to detect underdispersion in Poisson regression model versus generalized Poisson regression model type II.

Furthermore, The result indicate that the power is in increasing pattern clearly for all the large sample size (n) : for large sample, the power increases and gets to 1.0 very fast, for small sample the power increases slowly only in the beginning and then quickly gets to 1.0 .

Department of Statistics Graduate School, Silpakorn University Academic Year 2011
Student's signature ....
Thesis Advisor's signature ....

กิตติกรรมประกาศ

การศึกษา ค้นคว้าและเรียบเรียงวิทยานิพนธ์ฉบับนี้ สามารถสำเร็จลุล่วงไปด้วยดี ผู้วิจัยได้รับความอนุเคราะห์และความเมตตาจากรองศาสตราจารย์วีรานันท์ พงศาภักดี อาจารย์ที่ปรึกษาวิทยานิพนธ์ ที่ได้ให้คำแนะนำและความรู้ต่างๆในการจัดทำและการเรียบเรียงวิทยานิพนธ์และช่วยแก้ไขข้อบกพร่อง ด้วยความเมตตาและเอาใจใส่เป็นอย่างยิ่ง ผู้วิจัยจึงขอกราบขอบพระคุณรองศาสตราจารย์วีรานันท์ พงศาภักดี อาจารย์ที่ปรึกษาวิทยานิพนธ์ ไว้ ณ ที่นี้

ขอกราบขอบพระคุณ อาจารย์ ดร. ไพโรจน์ ขาวสิทธิวงษ์ และรองศาสตราจารย์ ดร. กมล บุษบา กรรมการตรวจสอบวิทยานิพนธ์ ที่ให้คำแนะนำต่าง ๆ เพื่อแก้ไขข้อบกพร่อง เพื่อให้วิทยานิพนธ์ฉบับนี้สมบูรณ์ยิ่งขึ้น

ขอขอบคุณบัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากรที่กรุณาให้ความอนุเคราะห์ สนับสนุนและอนุมัติให้ดำเนินงานวิจัยในครั้งนี้

สุดท้ายนี้ขอกราบขอบพระคุณบิดา มารดา คณาจารย์ ที่ให้กำลังใจและให้การสนับสนุนทางด้านการศึกษา ตลอดจนเพื่อน ๆ ที่ให้คำแนะนำต่าง ๆ ในช่วงเวลาของการเรียนและการเรียบเรียงวิทยานิพนธ์ตลอดมา

สารบัญ

หน้า

บทที่		หน้า

สารบัญตาราง

สารบัญภาพ

ภาพที่ หน้า

ภาพที่ หน้า

บทที่ 1

บทนำ

1. ความเป็นมาและความสำคัญของปัญหา

ในการสำรวจหรือการวิเคราะห์ข้อมูลจากตัวอย่างสุ่มที่เก็บรวบรวมมาได้นั้น หากตัวแปรสุ่มจำแนกประเภท (Y) ที่เก็บรวบรวมมาเป็นจำนวนนับและยังสามารถจำแนกตามตัวแปรอธิบายแบบจำแนกประเภทหรือแบบเชิงกลุ่ม (X) นักวิจัยมักใช้ตัวแบบปัวซงเป็นตัวแบบพื้นฐานในการวิเคราะห์ แต่ตัวแบบปัวซงมีคุณสมบัติคือ การกระจายต้องเท่ากัน (equi-dispersion) นั่นคือค่าเฉลี่ยมีค่าเท่ากับความแปรปรวน แต่ในทางปฏิบัติมักพบว่า ข้อมูลมีความแปรปรวนมากกว่าค่าเฉลี่ย ซึ่งกรณีที่ความแปรปรวนมากกว่าค่าเฉลี่ยนี้เรียกว่า Overdispersion และกรณีที่ความแปรปรวนน้อยกว่าค่าเฉลี่ยจะเรียกกรณีนี้ว่า Underdispersion ดังนั้น หากข้อมูลเกิด Overdispersion หรือ Underdispersion แล้วผู้วิจัยเลือกใช้ตัวแบบปัวซงจะทำให้ผลที่ได้ไม่ถูกต้องเท่าที่ควร เพราะฉะนั้นเมื่อเกิด Overdispersion หรือ Underdispersion ในตัวแบบปัวซง ผู้ทำการวิจัยควรพิจารณาตัวแบบอื่น ๆ ที่เหมาะสมมากกว่า เช่น ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 1 หรือ ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 เป็นต้น

ในปี 1973 Consul และ Jain เสนอการแจกแจงแบบปัวซงนัยทั่วไปเพื่อใช้แทนการแจกแจงแบบปัวซงเมื่อเกิด Overdispersion หรือ Underdispersion ซึ่งงานวิจัยของ Consul และ Jain ได้เสนอการสร้างตัวแบบสำหรับข้อมูลจำนวนนับที่มี Overdispersion โดยใช้การแจกแจงแบบปัวซงนัยทั่วไปเป็นพื้นฐานของตัวแบบการถดถอยซึ่งใช้สำหรับกรณี Overdispersion เมื่อความแปรปรวนมากกว่าค่าเฉลี่ย และใช้สำหรับกรณี Underdispersion เมื่อความแปรปรวนน้อยกว่าค่าเฉลี่ย ซึ่งตัวแบบการถดถอยปัวซงนัยทั่วไปที่ใช้กันทั่วไปมี 2 รูปแบบ โดยแต่ละรูปแบบจะขึ้นอยู่กับลักษณะความสัมพันธ์ระหว่างค่าเฉลี่ยและความแปรปรวน ซึ่งมีรายละเอียดดังนี้

ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 1 (generalized Poisson regression model Type 1: GP1) เป็นตัวแบบที่ค่าความแปรปรวนเป็นฟังก์ชันเชิงเส้นของค่าเฉลี่ย โดยตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 1 มีรูปแบบฟังก์ชันมวลความน่าจะเป็นของตัวแปรตอบสนอง Y ดังนี้

$$P\left(Y = y_i \mid \boldsymbol{x}_i, \boldsymbol{\beta}, \varphi\right) = \left(\left(1-\varphi\right)\mu_i + \varphi y_i\right)^{y_i - 1} \frac{\left(1-\varphi\right)\mu_i}{y_i\,!}\, exp\left(-\left(1-\varphi\right)\mu_i - \varphi y_i\right),\ y_i = 0,1,2,\ldots$$

โดยค่าเฉลี่ยเท่ากับ $\mu_i = exp\left(\boldsymbol{X}'\boldsymbol{\beta}\right)$ และความแปรปรวนเท่ากับ $\phi\mu_i$

ซึ่ง ϕ แทน ปัจจัยการกระจาย (dispersion factor) โดย $\phi = \dfrac{1}{(1-\varphi)^2}$ และ

φ แทน พารามิเตอร์การกระจาย (dispersion parameter)

โดย φ เป็นค่าคงที่ ซึ่ง $\varphi < 0$ เมื่อความแปรปรวนน้อยกว่าค่าเฉลี่ย (กรณี Underdispersion) และ $\varphi > 0$ เมื่อความแปรปรวนมากกว่าค่าเฉลี่ย (กรณี Overdispersion) และเมื่อ $\varphi = 0$ ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 1 จะลดรูปเป็นตัวแบบการถดถอยปัวซง

ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 (generalized Poisson regression model Type 2: GP2) เป็นตัวแบบที่ค่าความแปรปรวนไม่เป็นฟังก์ชันเชิงเส้นของค่าเฉลี่ย (อยู่ในรูปแบบกำลังสาม) โดยตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 มีรูปแบบฟังก์ชันมวลความน่าจะเป็นของตัวแปรตอบสนอง Y ดังนี้

$$P\left(Y_i = y_i \mid \boldsymbol{x_i}, \boldsymbol{\beta}, \varphi\right) = \left(\frac{\mu_i}{1+\varphi\mu_i}\right)^{y_i} \frac{\left(1+\varphi y_i\right)^{y_i-1}}{y_i\,!}\; exp\left(-\frac{\mu_i\left(1+\varphi y_i\right)}{1+\varphi\mu_i}\right) \quad , \quad y_i = 0,1,...$$

โดยค่าเฉลี่ยเท่ากับ $\mu_i = \exp\left(\boldsymbol{X'\beta}\right)$ และความแปรปรวนเท่ากับ $\mu_i\left(1+\varphi\mu_i\right)^2$

ซึ่งค่า φ แทน พารามิเตอร์การกระจาย (dispersion parameter) โดย φ เป็นค่าคงที่ ซึ่ง $\varphi < 0$ เมื่อความแปรปรวนน้อยกว่าค่าเฉลี่ย (กรณี Underdispersion) จะทำให้ $1+\varphi\mu_i > 0$ และ $1+\varphi y_i > 0$ นั่นคือ ค่า φ ที่สมการต้องการคือ $\varphi > min\left(-1/max(\mu_i), -1/max(y_i)\right)$ และเมื่อ $\varphi > 0$ จะทำให้ความแปรปรวนมากกว่าค่าเฉลี่ย ดังนั้นจะเรียกกรณีนี้ว่า Overdispersion และเมื่อ $\varphi = 0$ ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 จะลดรูปเป็นตัวแบบการถดถอยปัวซง

จากการศึกษาค้นคว้า พบว่า เมื่อตัวแปรสุ่มจำแนกประเภท (Y) ที่เก็บรวบรวมมาเป็นจำนวนนับและยังสามารถจำแนกตามตัวแปรอธิบาย (X) โดยที่ตัวแปรอธิบายอย่างน้อย 1 ตัวที่เป็นตัวแปรต่อเนื่อง หากข้อมูลที่เก็บรวบรวมมาได้นั้นเกิด Overdispersion หรือเกิด Underdispersion นักวิจัยจะเลือกใช้ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 1 หรือตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 แทนตัวแบบการถดถอยปัวซง แต่ตัวแบบที่นักวิจัยนิยมเลือกใช้คือ ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 เนื่องจากข้อมูลที่นำมาวิเคราะห์มักมีค่าความแปรปรวนไม่เป็นฟังก์ชันเชิงเส้นของค่าเฉลี่ย ซึ่งในที่นี้จะยกตัวอย่างงานวิจัยที่นักวิจัยเลือกใช้ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 ในการวิเคราะห์ข้อมูลในกรณีที่ข้อมูลเกิด Overdispersion หรือ Underdispersion ดังนี้ ในปี 1997 Wang และ Famoye ใช้ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 ในการวิเคราะห์ข้อมูลจำนวนบุตรในแต่ละครอบครัว (household fertility) ในปี 1999 Singh, Wulu, Bartolucci และ Valappil ใช้ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบทั่วไปแบบที่ 2 ในการวิเคราะห์ข้อมูลความถี่ของเหตุการณ์เกี่ยวกับ Gay Men's Sexual ในปี 2004 Famoye, Wulu และ Singh ใช้ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 ในการวิเคราะห์ข้อมูลการเกิดอุบัติเหตุ โดยพิจารณาความเกี่ยวพันระหว่างจำนวนการเกิดอุบัติเหตุและตัวแปรอธิบายแบบต่าง ๆ และใน

ปี 2007 Özmen และ Famoye ใช้ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 ในการวิเคราะห์ข้อมูลลักษณะต่าง ๆ ของสัตว์ จากที่กล่าวมาจะเห็นว่านักวิจัยมักเลือกใช้ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 ในการวิเคราะห์ข้อมูล ดังนั้นในงานวิจัยนี้ผู้วิจัยจึงสนใจศึกษาเฉพาะตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2

การเลือกใช้ตัวแบบทางสถิติเช่น ตัวแบบการถดถอยปัวซง ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 1 ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 หรือตัวแบบทางสถิติอื่น ๆ มีความสำคัญเป็นอย่างยิ่งต่อผลลัพธ์ที่จะได้จากการวิเคราะห์ ซึ่ง Dean and Lawless (1989) และ Dean (1992) กล่าวไว้ว่า เมื่อวิเคราะห์ข้อมูลโดยมองข้ามกรณีที่มี Overdispersion จะนำไปสู่การประมาณค่าความคลาดเคลื่อนมาตรฐานที่ต่ำกว่าความเป็นจริงและทำให้การอนุมานค่าพารามิเตอร์ของการถดถอยผิดพลาดไปด้วย ดังนั้น การทดสอบที่ใช้เป็นเกณฑ์ในการเลือกใช้ตัวแบบจึงมีความสำคัญเป็นอย่างยิ่งต่อการเลือกใช้ตัวแบบที่เหมาะสม หากผู้วิจัยเลือกใช้การทดสอบที่ไม่เหมาะสมอาจจะทำให้ผู้วิจัยเลือกใช้ตัวแบบที่ไม่เหมาะสมตามไปด้วย เพราะฉะนั้นการเลือกใช้การทดสอบเพื่อตรวจสอบความเหมาะสมของตัวแบบจึงมีความสำคัญเป็นอย่างยิ่ง ซึ่งการทดสอบที่ใช้ในการเปรียบเทียบความเหมาะสมของตัวแบบการถดถอยปัวซงเทียบกับตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 มีด้วยกันหลายการทดสอบเช่น การทดสอบวาลด์ที การทดสอบอัตราส่วนภาวะน่าจะเป็น การทดสอบสกอร์และการทดสอบเพียร์สันไควสแควร์

นอกจากนั้นยังมีนักวิจัยหลายท่านได้เสนอการทดสอบทางเลือกเพื่อใช้เป็นเกณฑ์ในการเปรียบเทียบความเหมาะสมของตัวแบบการถดถอยปัวซงเทียบกับตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 ดังนี้ ในปี 1965 Tiago de Oliveira เสนอการทดสอบ O_T เพื่อใช้ในการทดสอบ Overdispersion ในตัวแบบปัวซง ในปี 1994 Böhning ได้อธิบายข้อผิดพลาดของการทดสอบ O_T และเสนอการทดสอบ $\mathbf{Z_o}$ เพื่อใช้แทนการทดสอบ O_T โดยการทดสอบ $\mathbf{Z_o}$ นี้ใช้สำหรับการทดสอบ Overdispersion ในตัวแบบปัวซงและแสดงว่า การทดสอบ $\mathbf{Z_o}$ นี้มีการแจกแจงแบบปกติมาตรฐานเมื่อใกล้อนันต์ ในปี 1999 Jani, Shanubhogue และ Muralidharan เสนอการทดสอบ Q^* เพื่อใช้ในการเปรียบเทียบการแจกแจงแบบปัวซงเทียบกับการแจกแจงแบบปัวซงนัยทั่วไป แต่หาการแจกแจงที่แท้จริงของการทดสอบ Q^* ยาก นักวิจัยจึงใช้การจำลองแบบมอนติคาร์โลในการหาค่าวิกฤตที่จะใช้ในการปฏิเสธสมมติฐานว่างคือ $H_0 : \varphi = 0$ และในปี 2009 Yang, Hardin และ Addy เสนอการทดสอบสกอร์สำหรับ Overdispersion บนพื้นฐานของตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 ซึ่งการทดสอบสกอร์นี้ มีการแจกแจงเมื่อใกล้อนันต์คือ การแจกแจงแบบปกติมาตรฐาน

จากการศึกษาของงานวิจัยนี้ด้วยวิธีการจำลองแบบข้อมูลผู้วิจัยพบว่า การทดสอบ $\mathbf{Z_o}$ ไม่มีการแจกแจงแบบปกติมาตรฐานเมื่อใกล้อนันต์ ดังภาพที่ 20-23 ซึ่งแสดงการแจกแจงของการทดสอบ $\mathbf{Z_o}$ เมื่อกำหนดค่าพารามิเตอร์การกระจาย (φ) เท่ากับ 0.0 และขนาดตัวอย่างเท่ากับ 30, 50, 100 และ 200 ตามลำดับ (แสดงรายละเอียดในภาคผนวก ก) จากปัญหาที่พบนี้ ผู้วิจัยจึงขยายการศึกษาเกี่ยวกับการทดสอบ $\mathbf{Z_o}$ ซึ่งมีรูปแบบคือ $Z_o = \sqrt{(n-1)/2}\ \left[\left(S^2/\bar{Y}\right)-1\right]$ โดยที่ $\bar{Y}$ แทน ค่าเฉลี่ยของตัวอย่างสุ่ม

และ S^2 แทน ความแปรปรวนของตัวอย่าง ซึ่งภายใต้ $H_0 : \varphi = 0$ เป็นจริง ความแปรปรวนจะมีค่าเท่ากับค่าเฉลี่ย ซึ่งคล้ายกับการทดสอบสมมติฐานว่างคือ ตัวแบบการถดถอยปัวซง เทียบกับสมมติฐานแย้งคือ ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2

ดังนั้น ผู้วิจัยจึงเสนอใช้ค่าเฉลี่ยจากตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 นั่นคือ $\bar{\hat{\mu}}$ แทน $\bar{Y}$ และใช้ความแปรปรวนของตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 นั่นคือ $\bar{\hat{\mu}}\left(1+\hat{\varphi}\,\bar{\hat{\mu}}\right)^2$ แทน S^2 ซึ่งเมื่อแทนค่า $\bar{\hat{\mu}}$ และ $\bar{\hat{\mu}}\left(1+\hat{\varphi}\,\bar{\hat{\mu}}\right)^2$ จะเรียกการทดสอบใหม่นี้ว่า การทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) ซึ่งมีรูปแบบดังนี้

$$Z_{\bar{\hat{\mu}}} = \sqrt{\frac{n-1}{2}}\left(\frac{\bar{\hat{\mu}}\left(1+\hat{\varphi}\,\bar{\hat{\mu}}\right)^2}{\bar{\hat{\mu}}} - 1\right)$$

$$= \sqrt{\frac{n-1}{2}}\left(\left(1+\hat{\varphi}\,\bar{\hat{\mu}}\right)^2 - 1\right)$$

โดยที่ $\bar{\hat{\mu}}$ แทน ค่าเฉลี่ยของค่าที่พยากรณ์ที่ได้จากตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2

$\hat{\varphi}$ แทน ค่าประมาณพารามิเตอร์การกระจายภายใต้ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2

ผลลัพธ์จากการจำลองแบบข้อมูลพบว่า การทดสอบ $Z_{\bar{\hat{\mu}}}$ มีการแจกแจงแบบสมมาตรและลู่เข้าการแจกแจงแบบปกติมาตรฐานเมื่อใกล้อนันต์ ดังภาพที่ 24-27 ซึ่งภาพที่ 24-27 แสดงการแจกแจงของการทดสอบ $Z_{\bar{\hat{\mu}}}$ เมื่อกำหนดค่าพารามิเตอร์การกระจาย (φ) เท่ากับ 0.0 และขนาดตัวอย่างเท่ากับ 30, 50, 100 และ 200 ตามลำดับ (แสดงรายละเอียดในภาคผนวก ก)

เมื่อพิจารณาการทดสอบ $Z_{\bar{\hat{\mu}}}$ จะพบว่า ค่าเฉลี่ยของค่าที่พยากรณ์ที่ได้จากตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 $\left(\bar{\hat{\mu}}\right)$ คำนวณค่อนข้างยุ่งยาก ดังนั้นผู้วิจัยจึงเสนอใช้ค่าเฉลี่ยของตัวอย่างสุ่ม นั่นคือ $\bar{Y}$ แทน $\bar{\hat{\mu}}$ ซึ่งเมื่อแทนค่า $\bar{Y}$ จะเรียกการทดสอบใหม่นี้ว่า การทดสอบ $Z_{\bar{Y}}$ (Propose2) ซึ่งมีรูปแบบดังนี้

$$Z_{\bar{Y}} = \sqrt{\frac{n-1}{2}}\left(\left(1+\hat{\varphi}\bar{Y}\right)^2 - 1\right)$$

โดยที่ $\bar{Y}$ แทน ค่าเฉลี่ยของตัวอย่างสุ่ม ซึ่งคำนวณได้จาก $\sum_{i=1}^{n} y_i \Big/ n$

$\hat{\varphi}$ แทน ค่าประมาณพารามิเตอร์การกระจายภายใต้ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2

ผลลัพธ์จากการจำลองแบบข้อมูลพบว่า การทดสอบ $Z_{\bar{Y}}$ มีการแจกแจงแบบสมมาตรและลู่เข้าการแจกแจงแบบปกติมาตรฐานเมื่อใกล้อนันต์ ดังภาพที่ 28-31 ซึ่งภาพที่ 28-31 แสดงการแจกแจงของการทดสอบ $Z_{\bar{Y}}$ เมื่อกำหนดค่าพารามิเตอร์การกระจาย (φ) เท่ากับ 0.0 และขนาดตัวอย่างเท่ากับ 30, 50, 100 และ 200 ตามลำดับ (แสดงรายละเอียดในภาคผนวก ก)

ดังนั้น ในงานวิจัยนี้ผู้วิจัยจะเสนอการทดสอบ 2 การทดสอบคือ การทดสอบ $Z_{\bar{\hat{\mu}}}$ และการทดสอบ $Z_{\bar{Y}}$ ซึ่งมีการแจกแจงแบบสมมาตรและลู่เข้าการแจกแจงแบบปกติมาตรฐานเมื่อใกล้อนันต์ เพื่อใช้ในการเปรียบเทียบตัวแบบการถดถอยปัวซงเทียบกับตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 ทั้งในกรณี Overdispersion และ Underdispersion โดยทำการศึกษาการทดสอบและเปรียบเทียบผลลัพธ์กับการทดสอบอื่น ๆ ดังนี้ การทดสอบ Z_0 ซึ่งเสนอโดย Böhning (1994) การทดสอบวาลด์ที (Wald-t test) อ้างถึงใน Yang, Hardin and Addy (2009) และ Wang and Famoye (1997) การทดสอบอัตราส่วนภาวะน่าจะเป็น (Likelihood ratio test) อ้างถึงใน Yang, Hardin and Addy (2009) การทดสอบสกอร์ (Score test) ซึ่งเสนอโดย Yang, Hardin and Addy (2009) และการทดสอบ Q^* ซึ่งเสนอโดย Jani, Shanubhogue and Muralidharan (1999) โดยในงานวิจัยครั้งนี้จะทำการเปรียบเทียบกำลังการทดสอบที่ได้จากการทดลองของการทดสอบ $Z_{\bar{\hat{\mu}}}$ การทดสอบ $Z_{\bar{Y}}$ การทดสอบ Z_0 การทดสอบวาลด์ที การทดสอบอัตราส่วนภาวะน่าจะเป็น การทดสอบสกอร์ และการทดสอบ Q^* ภายใต้ตัวแบบ 2 ตัวแบบที่จะใช้ในงานวิจัยนี้คือ ตัวแบบการถดถอยปัวซงและตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 จากการจำลองแบบข้อมูลทำภายใต้ตัวแบบจริงคือ ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 ในแต่ละสถานการณ์จำนวน 5,000 รอบ ภายใต้เงื่อนไขของขนาดตัวอย่างและพารามิเตอร์การกระจาย นอกจากนี้ ผู้วิจัยสนใจศึกษาอิทธิพลของขนาดตัวอย่างต่าง ๆ ที่มีผลกระทบต่อตัวแบบการถดถอยปัวซงและตัวแบบปัวซงนัยทั่วไปแบบที่ 2 ด้วย เนื่องจากตัวแบบทั้งสองอยู่ภายใต้การแจกแจงปัวซงที่โดยทั่วไปควรใช้ขนาดตัวอย่างที่มากพอ ภายใต้กรณีการเกิด Overdispersion และ Underdispersion

2. วัตถุประสงค์ของการวิจัย

1. เพื่อเปรียบเทียบการทดสอบ $Z_{\bar{\hat{\mu}}}$ การทดสอบ $Z_{\bar{Y}}$ การทดสอบ Z_0 การทดสอบวาลด์ที การทดสอบอัตราส่วนภาวะน่าจะเป็น การทดสอบสกอร์ และการทดสอบ Q^* ซึ่งใช้ทดสอบภาวะสารูปดีของตัวแบบการถดถอยปัวซงและตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2

2. เพื่อศึกษาอิทธิพลของขนาดตัวอย่างที่มีผลต่อกำลังการทดสอบภายใต้กรณี Overdispersion และกรณี Underdispersion

3. สมมติฐานของการวิจัย

1. การทดสอบ $Z_{\bar{Y}}$ ให้กำลังการทดสอบที่ได้จากการทดลองมากกว่าการทดสอบอื่น ๆ

2. การทดสอบ $Z_{\bar{\hat{\mu}}}$ และการทดสอบ $Z_{\bar{Y}}$ ให้กำลังการทดสอบที่ได้จากการทดลองใกล้เคียงกัน

3. เมื่อเกิด Overdispersion หรือ Underdispersion กำลังการทดสอบที่ได้จากการทดลองในแต่ละขนาดตัวอย่างจะให้ผลใกล้เคียงกันในทุก ๆ การทดสอบ

4. ขอบเขตของการวิจัย

งานวิจัยนี้ใช้การจำลองแบบข้อมูลเพื่อเปรียบเทียบกำลังการทดสอบที่ได้จากการทดลองของการทดสอบต่าง ๆ โดยขอบเขตการวิจัยประกอบด้วยขอบเขตการจำลองแบบข้อมูล ขอบเขตของการทดสอบ และขอบเขตตัวแบบที่ใช้ในการวิเคราะห์ข้อมูล ซึ่งมีรายละเอียดดังนี้

4.1 ขอบเขตการจำลองแบบข้อมูล

1. ตัวแปรตอบสนอง (Y)

ตัวแปรตอบสนอง (Y) จำลองมาจากตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 ซึ่งมีค่าเฉลี่ยเท่ากับ μ_i และความแปรปรวนเท่ากับ $\mu_i\left(1+\varphi\mu_i\right)^2$ โดยคำนวณ μ_i จากเซตของพารามิเตอร์ $\{\beta_0, \beta_1\}$ ดังนี้ $\{2, -0.5\}$ ซึ่งสมการที่ใช้ในการคำนวณคือ $log\ \mu_i = 2-0.5x_i$

2. ตัวแปรอธิบาย (X)

ตัวแปรอธิบาย (X) ที่ใช้ในงานวิจัยนี้มี 1 ตัวแปร โดยตัวแปรอธิบาย (X) มีการแจกแจงแบบ Uniform [0,1] แบบต่อเนื่อง

4.2 ขอบเขตการทดสอบ

ขอบเขตการทดสอบแบ่งออกเป็น 2 ส่วน ดังนี้

<u>ส่วนที่ 1</u> การทดสอบที่ผู้วิจัยเสนอ ได้แก่

1. การทดสอบ $Z_{\bar{\hat{\mu}}}$
2. การทดสอบ $Z_{\bar{Y}}$

<u>ส่วนที่ 2</u> การทดสอบที่นำมาเปรียบเทียบกับการทดสอบที่ผู้วิจัยเสนอ ได้แก่

1. การทดสอบ $Z_{\mathbf{o}}$
2. การทดสอบวาลด์ที
3. การทดสอบอัตราส่วนภาวะน่าจะเป็น
4. การทดสอบสกอร์
5. การทดสอบ Q^*

4.3 ขอบเขตตัวแบบที่ใช้ในการวิเคราะห์ข้อมูล

1. ตัวแบบการถดถอยปัวซง
2. ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2

5. นิยามศัพท์เฉพาะ

1. กำลังการทดสอบที่ได้จากการทดลอง

1.1 กำลังการทดสอบที่ได้จากการทดลอง กรณี Overdispersion หมายถึง สัดส่วนของจำนวนครั้งที่ S มากกว่าค่าวิกฤต C หารด้วยจำนวนของการทำซ้ำ คือ

$$\frac{\#(S>C)}{R}$$

โดยที่ S แทน ค่าสถิติทดสอบที่คำนวณได้จากแต่ละการทดสอบ

C แทน ค่าวิกฤตที่ใช้ในการปฏิเสธสมมติฐานว่าง H_0

R แทน จำนวนรอบของการทำซ้ำ ในที่นี้ $R = 5,000$

1.2 กำลังการทดสอบที่ได้จากการทดลอง กรณี Underdispersion หมายถึง สัดส่วนของจำนวนครั้งที่ S น้อยกว่าค่าวิกฤต C หารด้วยจำนวนของการทำซ้ำ คือ

$$\frac{\#(S<C)}{R}$$

โดยที่ S แทน ค่าสถิติทดสอบที่คำนวณได้จากแต่ละการทดสอบ

C แทน ค่าวิกฤตที่ใช้ในการปฏิเสธสมมติฐานว่าง H_0

R แทน จำนวนรอบของการทำซ้ำ ในที่นี้ $R = 5,000$

2. ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2

ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 เป็นตัวแบบที่ค่าความแปรปรวนไม่เป็นฟังก์ชันเชิงเส้นของค่าเฉลี่ย (อยู่ในรูปแบบกำลังสาม) โดยตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 มีรูปแบบฟังก์ชันมวลความน่าจะเป็นของตัวแปรตอบสนอง Y ดังนี้

$$P\left(Y_i = y_i \mid \boldsymbol{x_i}, \boldsymbol{\beta}, \varphi\right) = \left(\frac{\mu_i}{1+\varphi\mu_i}\right)^{y_i} \frac{(1+\varphi y_i)^{y_i-1}}{y_i\,!}\, exp\left(-\frac{\mu_i(1+\varphi y_i)}{1+\varphi\mu_i}\right) \quad , \quad y_i = 0,1,...$$

เมื่อค่าเฉลี่ยเท่ากับ $\mu_i = exp(\boldsymbol{X'\beta})$ และความแปรปรวนเท่ากับ $\mu_i\,(1+\varphi\mu_i)^2$

ซึ่งค่า φ แทน พารามิเตอร์การกระจาย (dispersion parameter)

6. ประโยชน์ที่คาดว่าจะได้รับ

1. เป็นแนวทางในการเลือกใช้การทดสอบที่เหมาะสมในการเปรียบเทียบตัวแบบการถดถอยปัวซงเทียบกับตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2

2. เป็นแนวทางในการเลือกใช้ขนาดตัวอย่างเมื่อเกิด Overdispersion หรือ Underdispersion สำหรับการวิเคราะห์ตัวแบบการถดถอยปัวซงเทียบกับตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2

บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในบทนี้แบ่งออกเป็น 3 ส่วนดังนี้ ส่วนที่ 1 คือ ทฤษฎีที่เกี่ยวข้อง ส่วนที่ 2 คือ การทดสอบที่ใช้ในงานวิจัย และส่วนที่ 3 คือ งานวิจัยที่เกี่ยวข้อง

1. ทฤษฎีที่เกี่ยวข้อง

1.1 การแจกแจงแบบปัวซง (Poisson Distribution)

ให้เวกเตอร์ตอบสนองคือ $\mathbf{Y}=\left(Y_1,...,Y_n\right)'$ โดยที่ n แทน ขนาดของตัวอย่างและ Y_i , Y_j มีการแจกแจงเหมือนกันและเป็นอิสระต่อกันสำหรับทุก ๆ $i \neq j$ โดยฟังก์ชันมวลความน่าจะเป็น (probability mass function (p.m.f.)) ของการแจกแจงแบบปัวซงสำหรับตัวแปรสุ่ม Y_i (Agresti 2002 : 7, อ้างถึงใน Poisson 1837 : 206) คือ

$$P(y_i\,;\mu)=\frac{\mu^{y_i}\,e^{-\mu}}{y_i!} \qquad , \quad y_i=0,1,2,... \quad , \quad \mu>0$$

โดยค่าเฉลี่ยและความแปรปรวนมีค่าเท่ากับ μ

การประมาณค่าพารามิเตอร์ $\hat{\beta}$ สามารถประมาณด้วยวิธีภาวะน่าจะเป็นสูงสุด (Method of Maximum Likelihood Estimate: MLE) โดย

$$L(\beta)=\prod_{i=1}^{n}\frac{exp(-\mu_i)\mu_i^{y_i}}{y_i\,!}$$

ซึ่ง Log-likelihood function คือ

$$ln\,L(\beta)=\sum_{i=1}^{n}\left\{-\mu_i+y_i\,ln\,\mu_i-ln\,y_i!\right\}$$

การประมาณพารามิเตอร์ของตัวแบบปัวซงอาศัยการวนซ้ำและ $\mu_i = exp\left(\boldsymbol{X}_{\mathbf{i}}'\boldsymbol{\beta}\right)$ โดย $\mathbf{X}$ แทน เวกเตอร์ของตัวแปรร่วมและ $\boldsymbol{\beta}$ แทน เวกเตอร์ของพารามิเตอร์ที่ไม่ทราบค่าและต้องการประมาณ ซึ่งค่า $\hat{\beta}$ หาได้จากการหาอนุพันธ์ย่อยของฟังก์ชันล็อกภาวะน่าจะเป็น(Log-likelihood function) เทียบกับ β และให้สมการอนุพันธ์มีค่าเท่ากับ 0 ดังนี้

$$\frac{\partial \ln L(\beta)}{\partial \beta_j} = \sum_{i=1}^{n}\left(y_i - \exp\left(\mathbf{X}_i'\boldsymbol{\beta}\right) X_i\right) = 0 \qquad j=1, 2,\ldots, k.$$

เนื่องจากสมการที่กล่าวข้างต้นไม่สามารถแก้ได้โดยตรง จึงต้องใช้การแก้สมการ k สมการ โดยวิธีวนซ้ำแบบนิวตันราฟสัน (Newton-Raphson) ในการประมาณค่าพารามิเตอร์ β

1.2 การแจกแจงแบบปัวซงนัยทั่วไป (generalized Poisson distribution)

ในปี 1973 Consul และ Jain ได้เสนอการแจกแจงแบบปัวซงนัยทั่วไปสำหรับการเกิด Overdispersion ในข้อมูลจำนวนนับ ซึ่งมีพารามิเตอร์ 2 ตัวคือ (μ,φ) โดยฟังก์ชันมวลความน่าจะเป็นของการแจกแจงแบบปัวซงนัยทั่วไปสำหรับตัวแปรสุ่ม Y_i คือ

$$P(y_i ;\mu ,\varphi)=\frac{\mu(\mu+\varphi y_i)^{y_i-1} e^{-\mu-\varphi y_i}}{y_i!} \qquad , \quad y_i = 0,1,2,\ldots$$

โดย μ แทน ค่าเฉลี่ย ซึ่ง $\mu > 0$

φ แทน พารามิเตอร์การกระจาย ซึ่งมีค่าอยู่ระหว่าง $max(-1,-\mu/4) < \varphi < 1$

ซึ่งการแจกแจงแบบปัวซงนัยทั่วไปนี้ มีค่าเฉลี่ยเท่ากับ $\frac{\mu}{1-\varphi}$

และความแปรปรวนเท่ากับ $\frac{\mu}{(1-\varphi)^3} = \frac{1}{(1-\varphi)^2} E(Y_i) = \phi E(Y_i)$

(Yang, Hardin and Addy 2009 : 1515, อ้างถึงใน Joe and Zhu 2005) โดยเทอม $\phi = \frac{1}{(1-\varphi)^2}$ แสดงถึงบทบาทของปัจจัยการกระจาย (dispersion factor) โดยที่ $\varphi = 0$ การแจกแจงแบบปัวซงนัยทั่วไปจะลดรูปเป็นการแจกแจงแบบปัวซงที่มีพารามิเตอร์ μ ถ้า $\varphi < 0$ จะเกิด Underdispersion ในตัวแบบ แต่ถ้า $\varphi > 0$ จะเกิด Overdispersion ในตัวแบบ

การประมาณค่าพารามิเตอร์ μ และ φ สำหรับการแจกแจงแบบปัวซงนัยทั่วไป (He, Xie, Goh and Tsui 2006 : 384) ทำได้โดยใช้วิธีการประมาณแบบโมเมนต์และวิธีการประมาณแบบภาวะน่าจะเป็นสูงสุด ซึ่ง log-likelihood function คือ

$$L(\mu,\varphi)=n \ln \mu+\sum_{i=1}^{n}(y_i-1)\ln(\mu+y_i\varphi)-n\mu-n\bar{y}\varphi-\sum_{i=1}^{n}\ln y_i!$$

โดยที่ $\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n}$

1.3 ตัวแบบการถดถอยปัวซง (Poisson regression model)

ตัวแบบการถดถอยที่ใช้กันโดยทั่วไปสำหรับข้อมูลจำนวนนับคือ ตัวแบบการถดถอยปัวซง ซึ่งเป็นการศึกษาความสัมพันธ์ระหว่างตัวแปรตอบสนอง Y_i ที่มีการแจกแจงแบบปัวซงและมีตัวแปรอธิบาย $\mathbf{x_i}$ อย่างน้อย 1 ตัวที่เป็นตัวแปรต่อเนื่อง โดยเชื่อมความสัมพันธ์ของตัวแปรตอบสนองและตัวแปรอธิบายด้วย log link function ผ่านค่าเฉลี่ย μ_i ถ้าตัวแปรอธิบายทุกตัวเป็นตัวแปรจำแนกประเภท (categorical variable) เรียกว่า log-linear model ซึ่งรูปแบบคือ

$$\ln(\mu_i) = \boldsymbol{X'\beta}$$

โดย $\mu_i = \exp(\boldsymbol{X'\beta})$ ซึ่ง $\mathbf{X}$ แทน เวกเตอร์ของตัวแปรร่วมและ $\boldsymbol{\beta}$ แทน เวกเตอร์ของพารามิเตอร์ที่ไม่ทราบค่าและต้องการประมาณ (Cameron and Trivedi 1999 : 2-3)

1.4 ตัวแบบการถดถอยปัวซงนัยทั่วไป (generalized Poisson (GP) regression model)

ตัวแบบการถดถอยปัวซงนัยทั่วไปมี 2 รูปแบบ โดยแต่ละรูปแบบจะขึ้นอยู่กับลักษณะความสัมพันธ์ระหว่างค่าเฉลี่ยและความแปรปรวน ซึ่งมีรายละเอียดดังนี้

ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 1 (generalized Poisson regression model Type 1: GP1) เป็นตัวแบบที่ค่าความแปรปรวนเป็นฟังก์ชันเชิงเส้นของค่าเฉลี่ย โดยตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 1 มีรูปแบบดังนี้

$$P(Y = y_i \mid \boldsymbol{x_i}, \boldsymbol{\beta}, \varphi) = ((1-\varphi)\mu_i + \varphi y_i)^{y_i - 1} \frac{(1-\varphi)\mu_i}{y_i\,!} \exp\left(-(1-\varphi)\mu_i - \varphi y_i\right) \quad , \quad y_i = 0,1,2,\ldots$$

โดยค่าเฉลี่ยเท่ากับ $\mu_i = exp(\boldsymbol{X'\beta})$ และความแปรปรวนเท่ากับ $\phi\mu_i$

ซึ่ง ϕ แทน บทบาทของปัจจัยการกระจาย (dispersion factor) โดย $\phi = \dfrac{1}{(1-\varphi)^2}$ และ

φ แทน พารามิเตอร์การกระจาย (dispersion parameter)

โดย φ เป็นค่าคงที่ ซึ่ง $\varphi < 0$ เมื่อความแปรปรวนน้อยกว่าค่าเฉลี่ย (กรณี Underdispersion) และ $\varphi > 0$ เมื่อความแปรปรวนมากกว่าค่าเฉลี่ย (กรณี Overdispersion) และเมื่อ $\varphi = 0$ ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 1 จะลดรูปเป็นตัวแบบการถดถอยปัวซง

ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 (generalized Poisson regression model Type 2: GP2) เป็นตัวแบบที่ค่าความแปรปรวนไม่เป็นฟังก์ชันเชิงเส้นของค่าเฉลี่ย (อยู่ในรูปแบบกำลังสาม) โดยตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 มีรูปแบบดังนี้

$$P(Y_i = y_i \mid \boldsymbol{x_i}, \boldsymbol{\beta}, \varphi) = \left(\frac{\mu_i}{1+\varphi\mu_i}\right)^{y_i} \frac{(1+\varphi y_i)^{y_i - 1}}{y_i\,!}\, exp\left(-\frac{\mu_i(1+\varphi y_i)}{1+\varphi\mu_i}\right) \quad , \quad y_i = 0,1,\ldots$$

โดยค่าเฉลี่ยเท่ากับ $\mu_i = \exp(\boldsymbol{X'\beta})$ และความแปรปรวนเท่ากับ $\mu_i\,(1+\varphi\mu_i)^2$

ซึ่งค่า φ แทน พารามิเตอร์การกระจาย (dispersion parameter)

โดย φ เป็นค่าคงที่ ซึ่ง $\varphi < 0$ (กรณี Underdispersion) จะทำให้ $1+\varphi\mu_i > 0$ และ $1+\varphi y_i > 0$ นั่นคือค่า φ ที่สมการต้องการคือ $\varphi > \min\left(-1/\max(\mu_i)\ ,\ -1/\max(y_i)\right)$ และเมื่อ $\varphi > 0$ จะทำให้ความแปรปรวนมากกว่าค่าเฉลี่ย ดังนั้นจะเรียกกรณีนี้ว่า Overdispersion และเมื่อ $\varphi = 0$ ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 จะลดรูปเป็นตัวแบบการถดถอยปัวซง

การประมาณค่าพารามิเตอร์ $\boldsymbol{\beta}$ และ φ สำหรับตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 (Wang and Famoye 1997 : 277-278) ซึ่งมีรูปแบบคือ

$$P\left(Y_i = y_i \mid \boldsymbol{x_i}, \boldsymbol{\beta}, \varphi\right) = \left(\frac{\mu_i}{1+\varphi\mu_i}\right)^{y_i} \frac{\left(1+\varphi y_i\right)^{y_i-1}}{y_i\,!}\ exp\left(-\frac{\mu_i\left(1+\varphi y_i\right)}{1+\varphi\mu_i}\right) \quad , \quad y_i = 0,1,...$$

เราสามารถประมาณค่าพารามิเตอร์การกระจาย φ พร้อมกับสัมประสิทธิ์ β ได้ โดยใช้วิธีภาวะน่าจะเป็นสูงสุด (Method of Maximum Likelihood Estimate: MLE) ซึ่งฟังก์ชันล็อกภาวะน่าจะเป็น (Log- Likelihood Function) คือ

$$\ln L\left(\varphi, \beta; y_i\right) = \sum_{i=1}^{n}\left\{y_i \ln\left(\frac{\mu_i}{1+\varphi\mu_i}\right) + \left(y_i - 1\right)\ln\left(1+\varphi y_i\right) - \frac{\mu_i\left(1+\varphi y_i\right)}{1+\varphi\mu_i} - \ln\left(y_i!\right)\right\}$$

การประมาณค่าพารามิเตอร์การกระจาย (φ) ทำได้โดยการหาอนุพันธ์ย่อยของฟังก์ชันล็อกภาวะน่าจะเป็นเทียบกับ φ และให้สมการอนุพันธ์มีค่าเท่ากับ 0 ดังนี้

$$\frac{\partial\, ln\, L}{\partial\varphi} = \sum_{i=1}^{n}\left\{\frac{-y_i\mu_i}{1+\varphi\mu_i} + \frac{y_i\left(y_i - 1\right)}{1+\varphi y_i} - \frac{\mu_i\left(y_i - \mu_i\right)}{\left(1+\varphi\mu_i\right)^2}\right\} \quad = \quad 0$$

และการประมาณค่าพารามิเตอร์ β ทำได้โดยการหาอนุพันธ์ย่อยของฟังก์ชันล็อกภาวะน่าจะเป็นเทียบกับ β_r และให้สมการอนุพันธ์มีค่าเท่ากับ 0 ดังนี้

$$\frac{\partial\, ln\, L}{\partial\beta_r} = \sum_{i=1}^{n}\left\{\frac{y_i - \mu_i}{\mu_i\left(1+\varphi\mu_i\right)^2}\ \frac{\partial\mu_i}{\partial\beta_r}\right\} \quad = \quad 0 \qquad , r = 1,2,\ldots k \qquad (1)$$

โดย $\mu_i = exp(\boldsymbol{X'\beta})$ ซึ่งสมการที่ (1) สามารถเขียนใหม่ได้ดังนี้

$$\frac{\partial\, ln\, L}{\partial\beta_1} = \sum_{i=1}^{n}\frac{y_i - \mu_i}{\left(1+\varphi\mu_i\right)^2} \quad = \quad 0$$

$$\frac{\partial\, ln\, L}{\partial\beta_r} = \sum_{i=1}^{n}\left\{\frac{\left(y_i - \mu_i\right)x_i}{\left(1+\varphi\mu_i\right)^2}\right\} \quad = \quad 0 \qquad , r = 2,3,\ldots k$$

เนื่องจากสมการที่กล่าวข้างต้นไม่สามารถแก้ได้โดยตรง จึงต้องใช้การแก้สมการไม่เชิงเส้นแบบมีกระบวนการย้อนซ้ำ โดยวิธีนิวตันราฟสัน (Newton-Raphson)

นอกจากนั้น เรายังสามารถประมาณพารามิเตอร์ φ โดยวิธีประมาณแบบโมเมนต์ได้อีกด้วย ซึ่งวิธีประมาณแบบโมเมนต์นี้จะให้ค่าเท่ากับเพียร์สันไคสแควร์ที่มีองศาอิสระเท่ากับ n-k ซึ่ง

$$\sum_{i=1}^{n}\frac{(y_i-\mu_i)^2}{\mu_i(1+\varphi\mu_i)^2}=n-k$$

โดย n แทน ขนาดตัวอย่าง

k แทน จำนวนพารามิเตอร์ของการถดถอย (Rashwan and Kamel 2011 : 216)

ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 1 และตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 เป็นตัวแบบที่ขยายจากตัวแบบการถดถอยปัวซง ความแตกต่างของตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 1 และตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 คือ ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 1 มีความแปรปรวนเป็นฟังก์ชันเชิงเส้นของค่าเฉลี่ย ในขณะที่ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 ความสัมพันธ์ของความแปรปรวนซึ่งอยู่ในฟังก์ชันของค่าเฉลี่ยไม่เป็นเชิงเส้น (อยู่ในรูปแบบกำลังสาม) (Yang, Hardin and Addy 2009 : 1515-1516)

2. การทดสอบที่ใช้ในงานวิจัย

การทดสอบที่ใช้เปรียบเทียบตัวแบบการถดถอยปัวซงเทียบกับตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 ที่ใช้ในงานวิจัยแบ่งออกเป็น 2 ส่วน ดังนี้

<u>**ส่วนที่ 1**</u> การทดสอบที่ผู้วิจัยเสนอ ได้แก่

1. การทดสอบ $\mathbf{Z}_{\bar{\hat{\mu}}}$ (Propose1)

จากการศึกษาของงานวิจัยนี้ด้วยวิธีการจำลองแบบข้อมูลผู้วิจัยพบว่า การทดสอบ $Z_{\mathbf{o}}$ ไม่มีการแจกแจงแบบปกติมาตรฐานเมื่อใกล้อนันต์ ดังภาพที่ 20-23 ซึ่งเมื่อพิจารณาภาพที่ 20-23 ซึ่งแสดงการแจกแจงของการทดสอบ $Z_{\mathbf{o}}$ เมื่อกำหนดค่าพารามิเตอร์การกระจาย (φ) เท่ากับ 0.0 และขนาดตัวอย่างเท่ากับ 30, 50, 100 และ 200 ตามลำดับ (แสดงรายละเอียดในภาคผนวก ก) จากปัญหาที่พบนี้ ผู้วิจัยจึงขยายการศึกษาเกี่ยวกับการทดสอบ $Z_{\mathbf{o}}$ ซึ่งมีรูปแบบคือ $Z_o = \sqrt{(n-1)/2}\,\left[\left(S^2/\bar{Y}\right)-1\right]$ โดยที่ $\bar{Y}$ แทน ค่าเฉลี่ยของตัวอย่างสุ่ม และ S^2 แทน ความแปรปรวนของตัวอย่าง ซึ่งภายใต้ $H_0 : \varphi = 0$ เป็นจริง ความแปรปรวนจะมีค่าเท่ากับค่าเฉลี่ย ซึ่งคล้ายกับการทดสอบสมมติฐานว่างคือ ตัวแบบการถดถอยปัวซง เทียบกับสมมติฐานแย้งคือ ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2

ดังนั้น ผู้วิจัยจึงเสนอใช้ค่าเฉลี่ยจากตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 นั่นคือ $\bar{\hat{\mu}}$ แทน $\bar{Y}$ และใช้ความแปรปรวนของตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 นั่นคือ $\bar{\hat{\mu}}\left(1+\hat{\varphi}\,\bar{\hat{\mu}}\right)^2$ แทน S^2 ซึ่งเมื่อแทนค่า $\bar{\hat{\mu}}$ และ $\bar{\hat{\mu}}\left(1+\hat{\varphi}\,\bar{\hat{\mu}}\right)^2$ จะเรียกการทดสอบใหม่นี้ว่า การทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) ซึ่งมีรูปแบบดังนี้

$$
\begin{aligned}
Z_{\bar{\hat{\mu}}} &= \frac{\bar{\hat{\mu}}\left(1+\hat{\varphi}\,\bar{\hat{\mu}}\right)^2 - \bar{\hat{\mu}}}{\sqrt{2\bar{\hat{\mu}}^2/(n-1)}} \\
&= \sqrt{\frac{n-1}{2}}\left(\frac{\bar{\hat{\mu}}\left(1+\hat{\varphi}\,\bar{\hat{\mu}}\right)^2}{\bar{\hat{\mu}}} - 1\right) \\
&= \sqrt{\frac{n-1}{2}}\left(\left(1+\hat{\varphi}\,\bar{\hat{\mu}}\right)^2 - 1\right)
\end{aligned}
$$

โดยที่ $\bar{\hat{\mu}}$ แทน ค่าเฉลี่ยของค่าที่พยากรณ์ที่ได้จากตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 และ

$\hat{\varphi}$ แทน ค่าประมาณพารามิเตอร์การกระจายภายใต้ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2

ผลลัพธ์จากการจำลองแบบข้อมูลพบว่า การทดสอบ $Z_{\bar{\hat{\mu}}}$ มีการแจกแจงแบบสมมาตรและลู่เข้าการแจกแจงแบบปกติมาตรฐานเมื่อใกล้อนันต์ ดังภาพที่ 24-27 ซึ่งภาพที่ 24-27 แสดงการแจกแจงของการทดสอบ $Z_{\bar{\hat{\mu}}}$ เมื่อกำหนดค่าพารามิเตอร์การกระจาย (φ) เท่ากับ 0.0 และขนาดตัวอย่างเท่ากับ 30, 50, 100 และ 200 ตามลำดับ (แสดงรายละเอียดในภาคผนวก ก)

ในงานวิจัยนี้จะใช้การทดสอบ $Z_{\bar{\hat{\mu}}}$ ในการเปรียบเทียบตัวแบบการถดถอยปัวซงเทียบกับตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 โดยใช้อักษรย่อคือ Propose1

2. การทดสอบ $Z_{\bar{Y}}$ (Propose2)

เมื่อพิจารณาการทดสอบ $Z_{\bar{\hat{\mu}}}$ จะพบว่า ค่าเฉลี่ยของค่าที่พยากรณ์ที่ได้จากตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 $\left(\bar{\hat{\mu}}\right)$ คำนวณค่อนข้างยุ่งยาก ดังนั้นผู้วิจัยจึงเสนอใช้ค่าเฉลี่ยของตัวอย่างสุ่ม นั่นคือ $\bar{Y}$ แทน $\bar{\hat{\mu}}$ ซึ่งเมื่อแทนค่า $\bar{Y}$ จะเรียกการทดสอบใหม่นี้ว่า การทดสอบ $Z_{\bar{Y}}$ (Propose2) ซึ่งมีรูปแบบดังนี้

$$
Z_{\bar{Y}} = \sqrt{\frac{n-1}{2}}\left(\left(1+\hat{\varphi}\bar{Y}\right)^2 - 1\right)
$$

โดยที่ $\bar{Y}$ แทน ค่าเฉลี่ยของตัวอย่างสุ่ม ซึ่งคำนวณได้จาก $\sum_{i=1}^{n} y_i \Big/ n$ และ

$\hat{\varphi}$ แทน ค่าประมาณพารามิเตอร์การกระจายภายใต้ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2

ผลลัพธ์จากการจำลองแบบข้อมูลพบว่า การทดสอบ $Z_{\bar{Y}}$ มีการแจกแจงแบบสมมาตรและลู่เข้าการแจกแจงแบบปกติมาตรฐานเมื่อใกล้อนันต์ ดังภาพที่ 28-31 ซึ่งภาพที่ 28-31 แสดงการแจกแจงของการทดสอบ $Z_{\bar{Y}}$ เมื่อกำหนดค่าพารามิเตอร์การกระจาย (φ) เท่ากับ 0.0 และขนาดตัวอย่างเท่ากับ 30, 50, 100 และ 200 ตามลำดับ (แสดงรายละเอียดในภาคผนวก ก)

ในการเปรียบเทียบตัวแบบการถดถอยปัวซงเทียบกับตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 โดยใช้อักษรย่อคือ Propose2

ส่วนที่ 2 การทดสอบที่นำมาเปรียบเทียบกับการทดสอบที่ผู้วิจัยเสนออีก 5 การทดสอบได้แก่

1. การทดสอบ $\mathbf{Z_o}$ เสนอโดย Böhning (1994)

ในปี 1994 Böhning เสนอการทดสอบ Z_o เพื่อใช้ในการทดสอบ Overdispersion ในตัวแบบการถดถอยปัวซง ซึ่งการทดสอบ Z_o มีรูปแบบดังนี้

$$Z_o = \frac{S^2 - \bar{X}}{\sqrt{var\left(S^2 - \bar{X}\right)}}$$

โดย $var\left(S^2 - \bar{X}\right)$ มีค่าเท่ากับ $2\mu^2/(n-1)$ และแทนค่า μ ด้วย $\bar{Y}$ จะได้การทดสอบ Z_o ที่มีรูปแบบดังนี้

$$Z_o = \frac{S^2 - \bar{X}}{\sqrt{2\bar{Y}^2/(n-1)}} = \sqrt{\frac{n-1}{2}}\left(\frac{S^2}{\bar{Y}} - 1\right)$$

โดยที่ $\bar{Y}$ แทน ค่าเฉลี่ยของตัวอย่างสุ่ม ซึ่งคำนวณได้จาก $\sum_{i=1}^{n} y_i \Big/ n$

S^2 แทน ความแปรปรวนของตัวอย่างสุ่ม ซึ่งคำนวณได้จาก $\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2 \Big/ n-1$

ซึ่ง Böhning กล่าวว่า การทดสอบ Z_o มีการแจกแจงแบบปกติมาตรฐานเมื่อใกล้อนันต์ แต่จากการจำลองแบบข้อมูลผู้วิจัยพบว่า การทดสอบ Z_o ไม่มีการแจกแจงแบบสมมาตรหรือปกติมาตรฐานเมื่อใกล้อนันต์

ในงานวิจัยนี้จะใช้การทดสอบ Z_o ภายใต้การแจกแจงแบบปกติมาตรฐานเมื่อใกล้อนันต์ เพื่อเปรียบเทียบกับการทดสอบ $Z_{\bar{\mu}}$ การทดสอบ $Z_{\bar{Y}}$ และการทดสอบอื่น ๆ ในการเปรียบเทียบตัวแบบการถดถอยปัวซงเทียบกับตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 ซึ่งการทดสอบ Z_o จะแทนด้วยอักษรย่อคือ Z_o

2. การทดสอบวาลด์ที อ้างถึงใน Yang, Hardin and Addy (2009) และ Wang and Famoye (1997)

ในงานวิจัยนี้จะใช้ตัวสถิติ t ในรูปแบบวาลด์ที ซึ่งมีการแจกแจงแบบปกติเมื่อใกล้อนันต์ โดยอยู่ในรูปของอัตราส่วนของตัวประมาณของ φ กับความคลาดเคลื่อนมาตรฐานของตัวมันเอง ดังนั้นการทดสอบวาลด์ที มีรูปแบบดังนี้

$$t = \frac{\hat{\varphi} - \varphi}{SE\left(\hat{\varphi}\right)}$$

โดยที่ $\hat{\varphi}$ แทน ค่าประมาณพารามิเตอร์การกระจายภายใต้ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2

$SE\left(\hat{\varphi}\right)$ แทน ความคลาดเคลื่อนมาตรฐานของค่าประมาณพารามิเตอร์การกระจาย

ในงานวิจัยนี้จะใช้การทดสอบวาลด์ทีนี้ในการเปรียบเทียบตัวแบบการถดถอยปัวซงเทียบกับตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 โดยใช้อักษรย่อคือ Wald-t

3. การทดสอบอัตราส่วนภาวะน่าจะเป็น อ้างถึงใน Yang, Hardin and Addy (2009)

การทดสอบอัตราส่วนภาวะน่าจะเป็นสำหรับ φ คือ

$$LRT_{\varphi} = -2\left[\ell(\hat{\boldsymbol{\mu}}) - \ell(\hat{\boldsymbol{\mu}}, \hat{\varphi})\right]$$

โดยที่ $\ell(\hat{\boldsymbol{\mu}})$ และ $\ell(\hat{\boldsymbol{\mu}}, \hat{\varphi})$ แทน ฟังก์ชัน Log ของ likelihood ภายใต้ตัวแบบการถดถอยปัวซงและตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 ตามลำดับ ทฤษฎีเชิงเส้นกำกับเมื่อใกล้อนันต์มาตรฐาน (standard asymptotic theory) แสดงให้เห็นว่า ภายใต้ H_0

$$sgn(\hat{\varphi})\sqrt{LRT_{\varphi}} = sgn(\hat{\varphi})\sqrt{-2\left[\ell(\hat{\boldsymbol{\mu}}) - \ell(\hat{\boldsymbol{\mu}}, \hat{\varphi})\right]}$$

ซึ่งเมื่อใกล้อนันต์จะมีการแจกแจงแบบปกติมาตรฐาน โดยที่ $\mathrm{sgn}(\cdot)$ แทน ฟังก์ชัน sign ซึ่งจะเท่ากับ 1 เมื่อ $\hat{\varphi} \geq 0$ และเท่ากับ -1 เมื่อ $\hat{\varphi} < 0$ ในงานวิจัยนี้จะใช้การทดสอบ Signed square-root ของ LRT นี้ในการเปรียบเทียบตัวแบบการถดถอยปัวซงเทียบกับตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 โดยใช้อักษรย่อคือ SSR-LRT

4. การทดสอบสกอร์ (Score test statistic) เสนอโดย Yang, Hardin and Addy (2009)

การทดสอบสกอร์มี 2 รูปแบบดังนี้

รูปแบบที่ 1 มีรูปแบบคือ

$$S_1(\hat{\boldsymbol{\beta}}) = \left(\sum_{i=1}^{n} 2\hat{\mu}_i^2\right)^{-1} \left[\sum_{i=1}^{n} \left((y_i - \hat{\mu}_i)^2 - y_i\right)\right]^2$$

โดยที่ y_i แทน ค่าของตัวอย่างสุ่ม

$\hat{\mu}_i$ แทน ค่าที่พยากรณ์ได้จากตัวแบบการถดถอยปัวซง

ภายใต้สมมติฐานว่างคือ ลักษณะของข้อมูลเหมาะสำหรับตัวแบบการถดถอยปัวซง ซึ่งการทดสอบสกอร์รูปแบบนี้มีการแจกแจงแบบไคสแควร์ (χ^2) ที่มีองศาอิสระ df = 1

รูปแบบที่ 2 มีรูปแบบคือ

$$S_2(\hat{\boldsymbol{\beta}}) = \left(\sqrt{2\sum_{i=1}^{n} \hat{\mu}_i^2}\right)^{-1} \sum_{i=1}^{n} \left((y_i - \hat{\mu}_i)^2 - y_i\right)$$

โดยที่ y_i แทน ค่าของตัวอย่างสุ่ม

$\hat{\mu}_i$ แทน ค่าที่พยากรณ์ได้จากตัวแบบการถดถอยปัวซง

ภายใต้สมมติฐานว่างคือ ลักษณะของข้อมูลเหมาะสำหรับตัวแบบการถดถอยปัวซง ซึ่งการทดสอบสกอร์ในรูปแบบนี้มีการแจกแจงเมื่อใกล้อนันต์คือ การแจกแจงแบบปกติมาตรฐาน

ในงานวิจัยนี้จะใช้การทดสอบสกอร์รูปแบบที่ 2 ในการเปรียบเทียบตัวแบบการถดถอยปัวซงเทียบกับตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 โดยใช้อักษรย่อคือ Score

5. การทดสอบ $\mathbf{Q^*}$ เสนอโดย Jani, Shanubhogue and Muralidharan (1999)

ในปี 1999 Jani, Shanubhogue and Muralidharan เสนอการทดสอบ Q^* ซึ่งใช้เป็นเกณฑ์ในการเลือกตัวแบบต่าง ๆ ซึ่งการทดสอบ Q^* สามารถคำนวณได้จาก

$$Q_g^* = \frac{1}{n\nu(T)}\left[\sum_{i=1}^{n} d^2(Y_i) - \frac{\left(\sum_{i=1}^{n} d(Y_i)\right)^2}{n}\right] \tag{2}$$

โดย T แทนตัวสถิติพอเพียงของ $\mu e^{-\varphi\mu}$ ซึ่ง $T = \sum_{i=1}^{n} d(Y_i) = \sum_{i=1}^{n} Y_i$

การแจกแจงแบบมีเงื่อนไขของ Y_i เมื่อกำหนด $T = t$ มีรูปแบบดังนี้

$$P(y_i|t) = \frac{n-1}{n}\binom{t}{y}\frac{(1+\varphi y)^{y-1}}{(n+\varphi t-\varphi y-1)^{y}}\left[1-\frac{1+\varphi y}{n+\varphi t}\right]^{t-1} \quad ; \quad y = 0,1,\ldots,t$$

ดังนั้น ภายใต้ H_0 การแจกแจงแบบมีเงื่อนไขของ Y_i เมื่อกำหนด $T = t$ คือ

$$P_{Y_i|T}(y_i|t) = \binom{t}{y}\left(\frac{1}{n}\right)^{y}\left[1-\frac{1}{n}\right]^{t-1} \quad ; \quad y = 0,1,\ldots,t$$

โดย $\sigma_{ii} = V_{H_0}(Y|\,t) = (n-1)\nu(t) = \frac{(n-1)t}{n^2}$

ดังนั้น $v(t) = \frac{t}{n^2}$

เมื่อแทนค่าต่าง ๆ ลงในสมการที่ (2) จะได้การทดสอบ Q^* ที่ใช้เปรียบเทียบตัวแบบการถดถอยปัวซงเทียบกับตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 ซึ่งมีรูปแบบดังสมการที่ (3)

$$Q^* = \frac{n\sum_{i=1}^{n} Y_i^2}{\sum_{i=1}^{n} Y_i} - \sum_{i=1}^{n} Y_i \tag{3}$$

เนื่องจากหาการแจกแจงที่แท้จริงของการทดสอบ Q^* ยาก ผู้วิจัยจึงใช้การจำลองแบบมอนติคาร์โลในการหาค่าวิกฤตที่จะใช้ปฏิเสธสมมติฐานว่าง ซึ่งค่าวิกฤตที่ใช้ปฏิเสธสมมติฐานว่าง $H_0 : \varphi = 0$ แสดงดังตารางที่ 1 ซึ่งจะปฏิเสธสมมติฐานว่าง $H_0 : \varphi = 0$ และยอมรับ $H_1 : \varphi > 0$ เมื่อค่า Q^* ที่คำนวณได้มากกว่า Percentage points ที่ระดับนัยสำคัญ α หรือ $q_{GPD}(\alpha)$ และจะปฏิเสธสมมติฐานว่าง $H_0 : \varphi = 0$ และยอมรับ $H_1 : \varphi < 0$ เมื่อค่า Q^* ที่คำนวณได้น้อยกว่า Percentage points ที่ระดับนัยสำคัญ $1-\alpha$ หรือ $q_{GPD}(1-\alpha)$ ซึ่งค่า Percentage points แสดงดังตารางที่ 1

ตารางที่ 1 Percentage points of Q^*

$n \downarrow$ $\alpha \longrightarrow$	1%	5%	95%	99%
10	2.000	3.588	16.556	20.818
20	8.524	10.778	29.143	37.095
30	14.714	18.000	42.714	50.500
40	21.778	26.308	55.111	62.538
50	29.076	34.231	66.583	74.854

3. งานวิจัยที่เกี่ยวข้อง

Consul and Jain (1973) เสนอการแจกแจงแบบปัวซงนัยทั่วไปสำหรับกรณีที่ความแปรปรวนมีค่ามากกว่า เท่ากับ หรือน้อยกว่าค่าเฉลี่ย ซึ่งการแจกแจงแบบปัวซงนัยทั่วไปมีพารามิเตอร์ 2 ตัวคือ (μ,φ) โดยฟังก์ชันมวลความน่าจะเป็นของการแจกแจงแบบปัวซงนัยทั่วไปสำหรับตัวแปรสุ่ม Y_i คือ $P(y_i ; \mu,\varphi) = \left[\mu(\mu+\varphi y_i)^{y_i-1} e^{-\mu-\varphi y_i} / y_i!\right]$, $y_i = 0,1,2,...$ โดย μ แทน ค่าเฉลี่ย ซึ่ง $\mu > 0$ และ φ แทน พารามิเตอร์การกระจาย เมื่อ $\varphi = 0$ การแจกแจงแบบปัวซงนัยทั่วไปจะลดรูปเป็นการแจกแจงแบบปัวซงที่มีพารามิเตอร์ μ ถ้า $\varphi < 0$ จะเกิด Underdispersion ในตัวแบบการถดถอยปัวซง แต่ถ้า $\varphi > 0$ จะเกิด Overdispersion ในตัวแบบการถดถอยปัวซง

Böhning (1994) อธิบายข้อผิดพลาดที่เกิดขึ้นเมื่อใช้การทดสอบ O_T ของ Tiago de Oliveira ซึ่งเสนอไว้เมื่อค.ศ.1965 ซึ่งการทดสอบ O_T นี้ใช้ในการทดสอบ Overdispersion ในตัวแบบปัวซง ซึ่งความผิดพลาดของการทดสอบ O_T คือ การทดสอบ O_T นี้ไม่มีการแจกแจงแบบปกติและค่าเฉลี่ยของการแจกแจงแบบปัวซงไม่เป็นอิสระกัน นอกจากนั้น Böhning ยังเสนอการทดสอบ O_T^{new} เพื่อใช้ในการทดสอบ Overdispersion ในตัวแบบปัวซง ซึ่งการทดสอบ O_T^{new} นี้มีการแจกแจงแบบปกติเมื่อใกล้อนันต์หรือการแจกแจงแบบปกติเชิงเส้นกำกับ

Jani, Shanubhogue and Muralidharan (1999) เสนอการทดสอบ Q^* ซึ่งใช้เป็นเกณฑ์ในการเลือกการแจกแจงที่เหมาะสม สำหรับการทดสอบการแจกแจงปัวซงนัยทั่วไป (generalized Poisson distribution) การทดสอบการแจกแจงทวินามลบนัยทั่วไป (generalized negative binomial distribution) การทดสอบการแจกแจงไวบูลล์ (Weibull distribution) และการทดสอบการแจกแจงแบบปกติ (Normal distribution)

Singh, Wulu, Bartolucci and Valappil (1999) ทำการวิเคราะห์ข้อมูลความถี่ของเหตุการณ์เกี่ยวกับ Gay Men's Sexual โดยเปรียบเทียบตัวแบบการถดถอยปัวซง ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 และตัวแบบการถดถอยทวินามลบ โดยใช้มาตรวัดภาวะสารูปดี (Goodness of fit-test) คือ Pearson's Chi-Square, Generalized Chi-Square, Deviance และ Log-Likelihood ในการตรวจสอบความเหมาะสมของตัวแบบ ซึ่งจากการศึกษาพบว่า ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 มีความเหมาะสมกว่าตัวแบบการถดถอยปัวซงและตัวแบบการถดถอยทวินามลบในเชิงสถิติ

Singh, Wulu and Bartolucci (2001) ได้ทำการวิเคราะห์ผลกระทบที่เกิดจากการเลือกตัวแปรอธิบายเข้าตัวแบบในข้อมูลความถี่ของการเดินทางของสมาชิกในครอบครัว (Household trip frequencies) โดยใช้ตัวแบบการถดถอยปัวซงและตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 และทำการประมาณค่าพารามิเตอร์การกระจายสำหรับตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 โดยใช้ Wald t-statistic สำหรับการทดสอบสมมติฐานว่าง $H_0: \varphi = 0$ เทียบกับ $H_1: \varphi \neq 0$ ซึ่งจากการทดสอบพบว่า φ แตกต่างจากศูนย์อย่างมีนัยสำคัญ และใช้มาตรวัดภาวะสารูปดี (Goodness of fit-test) คือ Pearson's Chi-square, Deviance และ Log-Likelihood ในการตรวจสอบความเหมาะสมของตัวแบบ ซึ่งจากการศึกษาพบว่า ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 เหมาะสมกว่าตัวแบบการถดถอยปัวซง เนื่องจากตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 มีค่า Pearson's Chi-square มากกว่าตัวแบบการถดถอยปัวซง และตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 มีค่า Deviance น้อยกว่าตัวแบบการถดถอยปัวซง และการประมาณค่า Log-Likelihood พบว่า ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 มีค่า Log-Likelihood เท่ากับ -1374.8166 และค่า Log-Likelihood ของตัวแบบการถดถอยปัวซงเท่ากับ -1636.1103 ดังนั้นตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 มีความเหมาะสมกว่าตัวแบบการถดถอยปัวซงในเชิงสถิติ

Xie, He and Goh (2001) ทำการเปรียบเทียบตัวแบบปัวซงเทียบกับตัวแบบปัวซงที่มีศูนย์มากโดยใช้กำลังการทดสอบของตัวสถิติทดสอบที่เป็นไปได้คือ การทดสอบสกอร์ (Score test) การทดสอบอัตราส่วนภาวะน่าจะเป็น (Likelihood ratio test) การทดสอบไคสแควร์ (Chi-square test) การทดสอบบนพื้นฐานของช่วงความเชื่อมั่นของ p (A test base on a confidence interval of p) การทดสอบ C (Cochran test) การทดสอบ R (Rao-Chakravarti) ซึ่งทำการศึกษาจากการจำลองแบบข้อมูล (simulation study) สำหรับแต่ละ p และ μ โดยใช้ขนาดตัวอย่างที่แตกต่างกัน 3 ขนาดคือ 10, 20 และ 50 โดยแต่ละขนาดตัวอย่างทำซ้ำ 1,000 ครั้ง จากการวิเคราะห์พบว่า การทดสอบบนพื้นฐานของช่วงความเชื่อมั่นสำหรับ p (A test base on a confidence interval of p) มีกำลังการทดสอบน้อยกว่าการทดสอบอื่น ๆ

Singh, Wulu , Bae, Bartolucci and Trevino (2003) ได้ทำการวิเคราะห์ข้อมูลการเปลี่ยนโรงพยาบาลโดยใช้ตัวแบบการถดถอยปัวซงและตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 และตัวแบบการถดถอยทวินามลบ และทำการประมาณค่าพารามิเตอร์การกระจายสำหรับตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 โดยใช้ Wald t-statistic ซึ่งมีการแจกแจงแบบปกติเมื่อใกล้อนันต์ และการทดสอบทางเลือกคือ การทดสอบ Log-Likelihood ซึ่งการทดสอบทั้งสองใช้ในการทดสอบสมมติฐานว่าง $H_0 : \varphi = 0$ เทียบกับ $H_1 : \varphi \neq 0$ นอกจากนี้ยังใช้มาตรวัดภาวะสารูปดี (Goodness of fit-test) คือ Chi-square, Deviance และ Log-Likelihood ในการตรวจสอบความเหมาะสมของตัวแบบทั้งสาม ซึ่งจากการศึกษาพบว่า ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 เหมาะสมกว่าตัวแบบการถดถอยปัวซง และเหมาะสมกว่าตัวแบบการถดถอยทวินามลบ

Famoye, Wulu and Singh, Jr. , K.P. (2004) ทำการวิเคราะห์ข้อมูลการเกิดอุบัติเหตุโดยใช้ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 ในการพิจารณาความสัมพันธ์ระหว่างจำนวนการเกิดอุบัติเหตุและตัวแปรอธิบายต่างๆ ซึ่งจากการประมาณค่าพารามิเตอร์การกระจาย (φ) พบว่า $\varphi =$ 0.0794±0.0296 โดยใช้ตัวสถิติ Wald type “t” พบว่า t = 2.68 ดังนั้น φ แตกต่างจากศูนย์อย่างมีนัยสำคัญ และการทดสอบโดยใช้มาตรวัดภาวะสารูปดี (Goodness of fit-test) พบว่า ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 มีค่า Log-Likelihood เท่ากับ -667.0 และตัวแบบการถดถอยปัวซงมีค่า Log-Likelihood เท่ากับ -673.3 ดังนั้นตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 เหมาะสมกับข้อมูลการเกิดอุบัติเหตุมากกว่าตัวแบบการถดถอยปัวซง

Özmen and Famoye (2007) ได้ทำการวิเคราะห์ข้อมูลลักษณะต่าง ๆ ของสัตว์ ซึ่งข้อมูลที่ใช้ในการวิเคราะห์มีค่าเป็นศูนย์ประกอบอยู่ด้วย จึงทำการเปรียบเทียบตัวแบบการถดถอยปัวซง ตัวแบบทวินามลบ ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 ตัวแบบการถดถอยปัวซงที่มีศูนย์มาก และตัวแบบการถดถอยปัวซงนัยทั่วไปที่มีศูนย์มาก โดยใช้การทดสอบสกอร์ (Score test), Log-likelihood และ Deviance จากการศึกษาพบว่า ตัวแบบการถดถอยปัวซงนัยทั่วไปที่มีศูนย์มากเหมาะสมกับข้อมูล (แต่ไม่เหมาะสมกว่าตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2) ตัวแบบทวินามลบและตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 เหมาะสมกว่าตัวแบบการถดถอยปัวซงและตัวแบบการถดถอยปัวซงที่มีศูนย์มาก ดังนั้นสรุปได้ว่า ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 เหมาะสมกว่าตัวแบบอื่น ๆ

Yang, Hardin and Addy (2009) เสนอการทดสอบสกอร์สำหรับ Overdispersion บนพื้นฐานของตัวแบบการถดถอยปัวซงนัยทั่วไป (ศึกษาเฉพาะตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2) และนำการทดสอบสกอร์มาเปรียบเทียบกำลังการทดสอบกับการทดสอบอัตราส่วนภาวะน่าจะเป็นและการทดสอบวาลด์ที โดยศึกษาจากการจำลองแบบข้อมูล (simulation study) ซึ่งจากการศึกษาพบว่าการทดสอบสกอร์ที่เสนอบนพื้นฐานของการแจกแจงแบบปกติมาตรฐานเชิงเส้นกำกับหรือเมื่อใกล้อนันต์ (asymptotic standard normal distribution) มีความเหมาะสมมากกว่าการทดสอบอัตราส่วนภาวะน่าจะเป็นและการทดสอบวาลด์ที เมื่อนำไปใช้ในทางปฏิบัติ

Annafari (2010) ทำการวิเคราะห์ปัจจัยที่มีอิทธิพลต่อความต้องการในการใช้โทรศัพท์ของชาวสวีเดน ซึ่งตัวแปรตอบสนองที่ใช้คือ จำนวนโทรศัพท์ต่อคนหนึ่งคน โดยจำนวนโทรศัพท์ต่อคนหนึ่งคนของตัวอย่างที่ถูกสุ่มมามีค่าอยู่ระหว่าง 0 ถึง 9 และตัวแปรอธิบายที่ใช้แบ่งออกเป็นตัวแปรแบบต่อเนื่องและตัวแปรแบบไม่ต่อเนื่อง โดยใช้ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 ในการวิเคราะห์ และตรวจสอบความเหมาะสมของตัวแบบโดยใช้ Deviance, Scaled Deviance, Pearson Chi-Square, Scaled Pearson Chi-Square และ Log Likelihood ซึ่งค่าของ Deviance และ Pearson Chi-Square หารด้วยองศาอิสระซึ่งมีค่าเท่ากับ 0.4579 และ 0.4785 ตามลำดับ ซึ่งค่าทั้งสองมีค่าต่ำกว่า 1 ชี้ให้เห็นว่า ข้อมูลเกิด Underdispersion ดังนั้นจึงเลือกใช้ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 ในแก้ปัญหาข้อมูลชุดนี้ และทำการตรวจสอบความเหมาะสมของตัวแบบโดยใช้ค่า Scaled Deviance และ Scaled Pearson Chi-Square หารด้วยองศาอิสระซึ่งมีค่าเท่ากับ 0.9569 และ 1 ตามลำดับ แสดงให้เห็นว่า เมื่อใช้ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 ข้อมูลชุดนี้ไม่เกิด Overdispersion และ Underdispersion

บทที่ 3

วิธีดำเนินการวิจัย

งานวิจัยนี้เป็นการศึกษาวิธีการทดสอบภาวะสารูปดีที่เหมาะสมสำหรับการเปรียบเทียบตัวแบบการถดถอยปัวซงเทียบกับตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 โดยศึกษาจากการจำลองแบบข้อมูลเพื่อเปรียบเทียบกำลังการทดสอบที่ได้จากการทดลองของการทดสอบต่าง ๆ ซึ่งมีวิธีดำเนินการวิจัยดังนี้

1. ข้อมูลที่ใช้ในการวิจัย

ข้อมูลที่นำมาใช้ในการวิจัยนี้เป็นข้อมูลที่ได้จากการจำลองแบบ ในแต่ละเงื่อนไข 5,000 รอบ ซึ่งตัวแบบที่ใช้ในการวิจัยนี้คือ ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 โดยจะใช้ตัวแปรอธิบาย 1 ตัว ในการจำลองแบบ ดังสมการที่ (9)

$$\log \mu_i = 2 - 0.5x_i \tag{9}$$

โดยที่ x_i สุ่มมาจากการแจกแจงแบบ Uniform $[0,1]$ แบบต่อเนื่อง
ภายใต้ตัวแบบเชิงเส้นนัยทั่วไป 2 ตัวแบบคือ ตัวแบบการถดถอยปัวซงและตัวแบบการถดถอยปัวซงนัยทั่วไป ใน (10) - (11) ตามลำดับ ดังนี้

$$\ln(\mu_i) = \boldsymbol{X'\beta} \tag{10}$$

โดยค่าเฉลี่ยและความแปรปรวนเท่ากับ $\mu_i = \exp(\boldsymbol{X'\beta})$

$$P(Y_i = y_i \mid \boldsymbol{x_i}, \boldsymbol{\beta}, \varphi) = \left(\frac{\mu_i}{1+\varphi\mu_i}\right)^{y_i} \frac{(1+\varphi y_i)^{y_i-1}}{y_i\,!}\, exp\left(-\frac{\mu_i(1+\varphi y_i)}{1+\varphi\mu_i}\right) \quad , \quad y_i = 0,1,... \tag{11}$$

โดยค่าเฉลี่ยเท่ากับ $\mu_i = exp(\boldsymbol{X'\beta})$ และความแปรปรวนเท่ากับ $\mu_i\,(1+\varphi\mu_i)^2$
ซึ่งค่า φ แทน พารามิเตอร์การกระจาย (dispersion parameter)

ภายใต้เงื่อนไขของขนาดตัวอย่างเท่ากับ 30, 50, 100 และ 200 และกำหนดค่าพารามิเตอร์การกระจาย (φ) ดังนี้

กรณี Overdispersion กำหนดค่าพารามิเตอร์การกระจายเท่ากับ 0.0, 0.015, 0.02, 0.025, 0.03, 0.035, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.10 และ 0.2

กรณี Underdispersion กำหนดค่าพารามิเตอร์การกระจายเท่ากับ 0.0, -0.015, -0.02, -0.025, -0.03, -0.035, -0.04, -0.05, -0.06, -0.07, -0.08, -0.09 , -0.10 และ -0.2

2. เครื่องมือที่ใช้ในการวิจัย

โปรแกรมสำเร็จรูป SAS® version 9.1

3. การทดสอบที่ใช้ในการวิจัย

การทดสอบที่ใช้เปรียบเทียบตัวแบบการถดถอยปัวซงเทียบกับตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 ที่ใช้ในงานวิจัยแบ่งออกเป็น 2 ส่วน ดังนี้

<u>ส่วนที่ 1</u> การทดสอบที่ผู้วิจัยเสนอ ได้แก่

1. การทดสอบ $\mathrm{Z}_{\bar{\hat{\mu}}}$ (Propose1)

การทดสอบ $\mathrm{Z}_{\bar{\hat{\mu}}}$ มีรูปแบบคือ

$$Z_{\bar{\hat{\mu}}} = \sqrt{\frac{n-1}{2}}\left(\left(1+\hat{\varphi}\,\bar{\hat{\mu}}\right)^2 - 1\right)$$

โดยที่ $\bar{\hat{\mu}}$ แทน ค่าเฉลี่ยของค่าที่พยากรณ์ได้จากตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2

$\hat{\varphi}$ แทน ค่าประมาณพารามิเตอร์การกระจายภายใต้ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2

2. การทดสอบ $\mathrm{Z}_{\bar{\mathrm{Y}}}$ (Propose2)

การทดสอบ $\mathrm{Z}_{\bar{\mathrm{Y}}}$ ซึ่งมีรูปแบบคือ

$$Z_{\bar{Y}} = \sqrt{\frac{n-1}{2}}\left(\left(1+\hat{\varphi}\bar{Y}\right)^2 - 1\right)$$

โดยที่ $\bar{Y}$ แทน ค่าเฉลี่ยของตัวอย่างสุ่ม ซึ่งคำนวณได้จาก $\sum_{i=1}^{n} y_i \Big/ n$

$\hat{\varphi}$ แทน ค่าประมาณพารามิเตอร์การกระจายภายใต้ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2

<u>**ส่วนที่ 2**</u> การทดสอบที่นำมาเปรียบเทียบกับการทดสอบที่ผู้วิจัยเสนออีก 5 การทดสอบได้แก่

1. การทดสอบ Z_o เสนอโดย Böhning (1994)

การทดสอบ Z_o มีรูปแบบคือ

$$Z_o = \sqrt{\frac{n-1}{2}}\left(\frac{S^2}{\bar{Y}} - 1\right)$$

โดยที่ $\bar{Y}$ แทน ค่าเฉลี่ยของตัวอย่างสุ่มซึ่งคำนวณได้จาก $\sum_{i=1}^{n} y_i \big/ n$

S^2 แทน ความแปรปรวนของตัวอย่างสุ่มซึ่งคำนวณได้จาก $\sum_{i=1}^{n}(y_i - \bar{y}) \big/ n-1$

2. การทดสอบวาลด์ที อ้างถึงใน Yang, Hardin and Addy (2009) และ Wang and Famoye (1997)

การทดสอบวาลด์ที มีรูปแบบคือ

$$t = \frac{\hat{\varphi}-\varphi}{SE(\hat{\varphi})}$$

โดยที่ $\hat{\varphi}$ แทน ค่าประมาณพารามิเตอร์การกระจายภายใต้ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2

$SE(\hat{\varphi})$ แทน ความคลาดเคลื่อนมาตรฐานของค่าประมาณพารามิเตอร์การกระจาย

3. การทดสอบอัตราส่วนภาวะน่าจะเป็น อ้างถึงใน Yang, Hardin and Addy (2009)

การทดสอบอัตราส่วนภาวะน่าจะเป็น มีรูปแบบคือ

$$sgn(\hat{\varphi})\sqrt{LRT_{\varphi}} = sgn(\hat{\varphi})\sqrt{-2[\ell(\hat{\boldsymbol{\mu}}) - \ell(\hat{\boldsymbol{\mu}},\hat{\varphi})]}$$

โดยที่ $\ell(\hat{\boldsymbol{\mu}})$ แทน ฟังก์ชัน Log ของ likelihood ภายใต้ตัวแบบการถดถอยปัวซง

$\ell(\hat{\boldsymbol{\mu}},\hat{\varphi})$ แทน ฟังก์ชัน Log ของ likelihood ภายใต้ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2

$\mathrm{sgn}(\cdot)$ แทน ฟังก์ชัน sign ซึ่งจะเท่ากับ 1 เมื่อ $\hat{\varphi} \geq 0$ และเท่ากับ -1 เมื่อ $\hat{\varphi} < 0$

4. การทดสอบสกอร์ เสนอโดย Yang, Hardin and Addy (2009)

การทดสอบสกอร์ มีรูปแบบคือ

$$S_2(\hat{\beta}) = \left(\sqrt{2\sum_{i=1}^{n}\hat{\mu}_i^2}\right)^{-1} \sum_{i=1}^{n}\left((y_i - \hat{\mu}_i)^2 - y_i\right)$$

โดยที่ y_i แทน ค่าของตัวอย่างสุ่ม

$\hat{\mu}_i$ แทน ค่าที่พยากรณ์ได้จากตัวแบบปัวซง

5. การทดสอบ Q^* เสนอโดย Jani, Shanubhogue and Muralidharan (1999)

การทดสอบ Q^* มีรูปแบบคือ

$$Q^* = \frac{n\sum_{i=1}^{n} Y_i^2}{\sum_{i=1}^{n} Y_i} - \sum_{i=1}^{n} Y_i$$

โดยที่ Y_i แทน ค่าของตัวอย่างสุ่ม

4. วิธีการจำลองแบบข้อมูล

วิธีการจำลองแบบข้อมูลมีขั้นตอน ดังนี้

1. กำหนดขนาดตัวอย่าง (n) เท่ากับ 30, 50, 100 และ 200

2. กำหนดค่าพารามิเตอร์ของการกระจาย (φ) ดังนี้

กรณี Overdispersion กำหนดค่าพารามิเตอร์การกระจายเท่ากับ 0.0, 0.015, 0.02, 0.025, 0.03, 0.035, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.10 และ 0.2

กรณี Underdispersion กำหนดค่าพารามิเตอร์การกระจายเท่ากับ 0.0, -0.015, -0.02, -0.025, -0.03, -0.035, -0.04, -0.05, -0.06, -0.07, -0.08, -0.09 , -0.10 และ -0.2

3. สุ่มตัวแปรอธิบาย x_i จำนวน n ค่า จากการแจกแจงแบบ Uniform $[0,1]$ แบบต่อเนื่อง

4. คำนวณค่าเฉลี่ย $E(Y_i) = \mu_i = \exp(\boldsymbol{X'\beta})$ จากสมการ $log\ \mu_i = 2 - 0.5x_i$

5. สุ่มตัวแปรสุ่ม Y จากการแจกแจงแบบปัวซงที่มีค่าเฉลี่ยเท่ากับ $\mu_i / (1+\varphi\mu_i)^2$

6. นำตัวแปรสุ่ม Y ที่ได้จากขั้นตอนที่ 4 มาคูณด้วย $(1+\varphi\mu_i)^2$ และปรับเศษทศนิยมให้เป็นจำนวนเต็ม โดยกำหนดว่า ถ้าเลขทศนิยมมากกว่า 0.5 ปัดเป็น 1 และถ้าเลขทศนิยมน้อยกว่า 0.5 ตัดเศษทศนิยมทิ้ง ดังนั้นจะได้ตัวแปรสุ่ม Y ที่มีการแจกแจงแบบปัวซงนัยทั่วไปแบบที่ 2 ซึ่งมีค่าเฉลี่ยเท่ากับ μ_i และความแปรปรวนเท่ากับ $\mu_i(1+\varphi\mu_i)^2$ ซึ่งเกิด Overdispersion เมื่อ $\varphi > 0$ และเกิด Underdispersion เมื่อ $\varphi < 0$ (Heinzl and Mittlböck 2003 : 258)

7. กำหนดการจำลองแบบข้อมูลจำนวน 5,000 ครั้งในแต่ละเงื่อนไข

5. วิธีการวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลแบ่งออกเป็น 2 กรณี ดังนี้

1. กรณี Overdispersion มีขั้นตอนดังนี้

1.1 กำหนด $\alpha = 0.05$ ขนาดตัวอย่าง (n) เท่ากับ 30, 50, 100 และ 200 และค่าพารามิเตอร์การกระจาย (φ) สำหรับตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 เท่ากับ 0.0, 0.015, 0.02, 0.025, 0.03, 0.035, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.10 และ 0.2

1.2 สร้างตัวแบบการถดถอยปัวซงและตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 ที่มีตัวแปรอธิบายเหมือนกัน

1.3 คำนวณค่าของการทดสอบต่าง ๆ ดังนี้

1.3.1 การทดสอบ $Z_{\bar{\hat{\mu}}}$

1.3.2 การทดสอบ $Z_{\bar{Y}}$

1.3.3 การทดสอบ $Z_{\mathbf{o}}$

1.3.4 การทดสอบวาลด์ที

1.3.5 การทดสอบอัตราส่วนภาวะน่าจะเป็น

1.3.6 การทดสอบสกอร์

1.3.7 การทดสอบ Q^*

1.4 สรุปผลการยอมรับหรือการปฏิเสธสมมติฐานว่างของการทดสอบแต่ละตัว ซึ่งสมมติฐานที่ใช้ในการทดสอบคือ $H_0 : \varphi = 0$ vs. $H_1 : \varphi > 0$

1.5 ทำซ้ำขั้นตอนที่ 1.1-1.4 จำนวน 5,000 ครั้ง สำหรับข้อมูลที่ได้จากการจำลองแบบในแต่ละเงื่อนไข

1.6 หากำลังการทดสอบที่ได้จากการทดลองของการทดสอบทั้งเจ็ด โดยคำนวณจากสัดส่วนของจำนวนครั้งที่ S มากกว่าค่าวิกฤต C หารด้วยจำนวนของการทำซ้ำ คือ

$$\frac{\#(S > C)}{R}$$

โดยที่ S แทน ค่าสถิติทดสอบที่คำนวณได้จากแต่ละการทดสอบ

□ แทน จำนวนรอบของการทำซ้ำ ในที่นี้ □= 5,000

C แทน ค่าวิกฤตที่ใช้ในการปฏิเสธสมมติฐานว่าง H_0 โดยการทดสอบ $Z_{\bar{\hat{\mu}}}$ การทดสอบ $Z_{\bar{Y}}$ การทดสอบ $Z_{\mathbf{o}}$ การทดสอบวาลด์ที การทดสอบอัตราส่วนภาวะน่าจะเป็นและการทดสอบสกอร์ ใช้ค่าวิกฤตที่ Z_{α} ซึ่งเท่ากับ 1.645 นั่นคือ จะปฏิเสธสมมติฐานว่าง H_0 เมื่อค่าสถิติทดสอบที่คำนวณได้มากกว่า 1.645 ส่วนการทดสอบ Q^* ใช้ค่าวิกฤตจากตารางที่ 1 โดยจะปฏิเสธสมมติฐานว่าง H_0 เมื่อค่าสถิติทดสอบที่คำนวณได้มากกว่าค่าวิกฤต ดังนี้ เมื่อขนาดตัวอย่างเท่ากับ 30 ค่าวิกฤตคือ 42.714 และเมื่อขนาดตัวอย่างเท่ากับ 50 ค่าวิกฤตคือ 66.583

1.7 เปรียบเทียบกำลังการทดสอบที่ได้จากการทดลองของการทดสอบทั้งเจ็ด

2. กรณี Underdispersion มีขั้นตอนดังนี้

2.1 กำหนด $\alpha = 0.05$ ขนาดตัวอย่าง (n) เท่ากับ 30, 50, 100 และ 200 และค่าพารามิเตอร์การกระจาย (φ) สำหรับตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 เท่ากับ 0.0, -0.015, -0.02, -0.025, -0.03, -0.035, -0.04, -0.05, -0.06, -0.07, -0.08, -0.09 , -0.10 และ -0.2

2.2 สร้างตัวแบบการถดถอยปัวซงและตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 ที่มีตัวแปรอธิบายเหมือนกัน

2.3 คำนวณค่าของการทดสอบต่าง ๆ ดังนี้

2.3.1 การทดสอบ $\mathrm{Z}_{\bar{\hat{\mu}}}$

2.3.2 การทดสอบ $\mathrm{Z}_{\bar{\mathrm{Y}}}$

2.3.3 การทดสอบ $\mathrm{Z}_{\mathbf{o}}$

2.3.4 การทดสอบวาลด์ที

2.3.5 การทดสอบอัตราส่วนภาวะน่าจะเป็น

2.3.6 การทดสอบสกอร์

2.3.7 การทดสอบ Q^*

2.4 สรุปผลการยอมรับหรือการปฏิเสธสมมติฐานว่างของการทดสอบแต่ละตัว ซึ่งสมมติฐานที่ใช้ในการทดสอบคือ $H_0 : \varphi = 0$ vs. $H_1 : \varphi < 0$

2.5 ทำซ้ำขั้นตอนที่ 2.1-2.4 จำนวน 5,000 ครั้ง สำหรับข้อมูลที่ได้จากการจำลองแบบในแต่ละเงื่อนไข

2.6 หากำลังการทดสอบที่ได้จากการทดลองของการทดสอบทั้งเจ็ด โดยคำนวณจากสัดส่วนของจำนวนครั้งที่ S น้อยกว่าค่าวิกฤต C หารด้วยจำนวนของการทำซ้ำ คือ

$$\frac{\#(S < C)}{R}$$

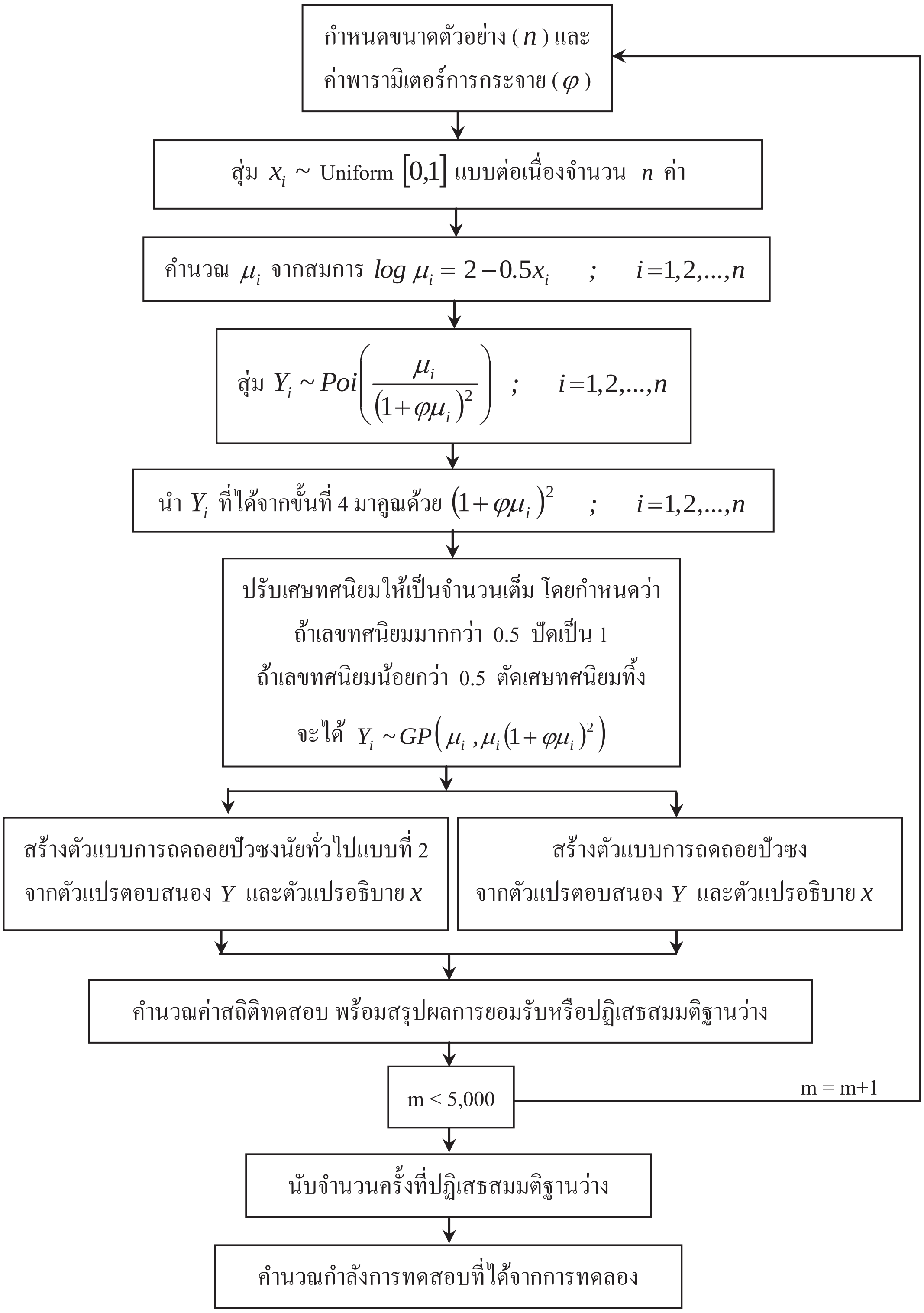
โดยที่ S แทน ค่าสถิติทดสอบที่คำนวณได้จากแต่ละการทดสอบ

□ แทน จำนวนรอบของการทำซ้ำ ในที่นี้ □= 5,000

C แทน ค่าวิกฤตที่ใช้ในการปฏิเสธสมมติฐานว่าง H_0 โดยการทดสอบ $\mathrm{Z}_{\bar{\hat{\mu}}}$ การทดสอบ $\mathrm{Z}_{\bar{\mathrm{Y}}}$ การทดสอบ $\mathrm{Z}_{\mathbf{o}}$ การทดสอบวาลด์ที การทดสอบอัตราส่วนภาวะน่าจะเป็นและการทดสอบสกอร์ ใช้ค่าวิกฤตที่ $\mathrm{Z}_{1-\alpha}$ ซึ่งเท่ากับ -1.645 นั่นคือ จะปฏิเสธสมมติฐานว่าง H_0 เมื่อค่าสถิติทดสอบที่คำนวณได้น้อยกว่า -1.645 ส่วนการทดสอบ Q^* ใช้ค่าวิกฤตจากตารางที่ 1 โดยจะปฏิเสธสมมติฐานว่าง H_0 เมื่อค่าสถิติทดสอบที่คำนวณได้น้อยกว่าค่าวิกฤต ดังนี้ เมื่อขนาดตัวอย่างเท่ากับ 30 ค่าวิกฤตคือ 18.000 และเมื่อขนาดตัวอย่างเท่ากับ 50 ค่าวิกฤตคือ 34.231

2.7 เปรียบเทียบกำลังการทดสอบที่ได้จากการทดลองของการทดสอบทั้งเจ็ด

ขั้นตอนการวิเคราะห์และจำลองแบบข้อมูลกรณี Overdispersion และ Underdispersion แสดงดังภาพที่ 1



ภาพที่ 1 ขั้นตอนการวิเคราะห์และจำลองแบบข้อมูลกรณี Overdispersion และ Underdispersion

บทที่ 4

ผลการวิจัย

ในการเปรียบเทียบการทดสอบภาวะสารูปดีสำหรับตัวแบบการถดถอยปัวซงและตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 ผู้วิจัยจะเสนอการทดสอบ 2 การทดสอบ คือการทดสอบ $Z_{\bar{\mu}}$ และการทดสอบ $Z_{\bar{Y}}$ เพื่อใช้ในการเปรียบเทียบตัวแบบการถดถอยปัวซงเทียบกับตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 และทำการศึกษาการทดสอบ $Z_{\mathbf{0}}$ การทดสอบวาลด์ที การทดสอบอัตราส่วนภาวะน่าจะเป็น การทดสอบสกอร์ และการทดสอบ Q^{*} แล้วนำการทดสอบทั้งสองที่เสนอขึ้นคือ การทดสอบ $Z_{\bar{\mu}}$ และการทดสอบ $Z_{\bar{Y}}$ มาเปรียบเทียบกำลังการทดสอบกับการทดสอบ $Z_{\mathbf{0}}$ การทดสอบวาลด์ที การทดสอบอัตราส่วนภาวะน่าจะเป็น การทดสอบสกอร์ และการทดสอบ Q^{*} ภายใต้ตัวแบบ 2 ตัวแบบคือ ตัวแบบการถดถอยปัวซงและตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2

ข้อมูลที่นำมาใช้ในการวิจัยนี้เป็นข้อมูลที่ได้จากการจำลองแบบ ซึ่งตัวแบบที่ใช้ในการวิจัยนี้คือ ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 โดยจะใช้ตัวแปรอธิบาย 1 ตัวในการจำลองแบบ ภายใต้เงื่อนไขของขนาดตัวอย่างเท่ากับ 30, 50, 100 และ 200 และกำหนดค่าพารามิเตอร์การกระจาย (φ) ดังนี้ กรณี Overdispersion กำหนดค่าพารามิเตอร์การกระจายเท่ากับ 0.0, 0.015, 0.02, 0.025, 0.03, 0.035, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.10 และ 0.2 กรณี Underdispersion กำหนดค่าพารามิเตอร์การกระจายเท่ากับ 0.0, -0.015, -0.02, -0.025, -0.03, -0.035, -0.04, -0.05, -0.06, -0.07, -0.08, -0.09 , -0.10 และ -0.2 เงื่อนไขละ 5,000 ครั้ง และนำข้อมูลที่ได้จากการจำลองแบบนี้มาทำการทดสอบ โดยทำการทดสอบภายใต้การทดสอบทั้งเจ็ดที่กล่าวมาข้างต้น และสมมติฐานว่างที่ใช้ในงานวิจัยนี้คือ ตัวแบบการถดถอยปัวซงเทียบกับสมมติฐานแย้งคือ ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2

ผลการวิเคราะห์ข้อมูลคือ กำลังการทดสอบที่ได้จากการทดลองของการทดสอบ $Z_{\bar{\mu}}$ การทดสอบ $Z_{\bar{Y}}$ การทดสอบ $Z_{\mathbf{0}}$ การทดสอบวาลด์ที การทดสอบอัตราส่วนภาวะน่าจะเป็น การทดสอบสกอร์ และการทดสอบ Q^{*} ภายใต้เงื่อนไขต่าง ๆ รวมถึงแสดงอิทธิพลของขนาดตัวอย่างที่มีผลต่อกำลังการทดสอบที่ได้จากการทดลองภายใต้กรณี Overdispersion และกรณี Underdispersion ของการทดสอบทั้งเจ็ดที่กล่าวมาข้างต้น ซึ่งผลการวิเคราะห์มีรายละเอียดดังนี้

1. กรณี Overdispersion

ตารางที่ 2 กำลังการทดสอบ (ร้อยละ) ที่ได้จากการทดลองของการทดสอบต่าง ๆ ในกรณี Overdispersion โดยทำซ้ำ 5,000 ครั้ง

n	Method	Power (%)													
		$\varphi = 0.0$	0.015	0.02	0.025	0.03	0.035	0.04	0.05	0.06	0.07	0.08	0.09	0.10	0.20
30	**Propose1**	**3.86**	**18.28**	**22.30**	**27.68**	**37.38**	**44.14**	**50.54**	**58.38**	**70.98**	**81.90**	**90.52**	**94.46**	**96.60**	**100.00**
	Propose2	**3.86**	**18.28**	**22.30**	**27.68**	**37.40**	**44.14**	**50.52**	**58.38**	**70.98**	**81.88**	**90.52**	**94.46**	**96.60**	**100.00**
	Z_o	14.48	36.18	42.00	46.52	55.40	62.80	68.80	73.66	83.68	90.00	95.04	97.42	98.36	100.00
	Wald-t	0.96	6.60	8.78	11.84	18.86	23.08	28.92	35.98	50.04	64.78	78.46	85.70	90.00	99.96
	SSR-LRT	**2.52**	**13.76**	**16.90**	**22.14**	**31.22**	**37.32**	**43.36**	**51.08**	**64.70**	**77.02**	**87.22**	**92.22**	**94.84**	**100.00**
	Score	**3.94**	**16.54**	**20.96**	**26.42**	**35.30**	**42.24**	**48.16**	**56.94**	**69.20**	**80.48**	**89.08**	**93.46**	**95.82**	**100.00**
	Q*	11.76	32.34	37.78	42.60	50.80	59.10	65.44	70.50	81.22	88.18	94.02	96.74	97.98	100.00
50	**Propose1**	**4.08**	**25.96**	**32.12**	**39.58**	**53.52**	**61.26**	**70.30**	**78.28**	**88.88**	**95.08**	**98.58**	**99.54**	**99.80**	**100.00**
	Propose2	**4.08**	**25.96**	**32.12**	**39.58**	**53.52**	**61.26**	**70.30**	**78.26**	**88.88**	**95.08**	**98.58**	**99.54**	**99.80**	**100.00**
	Z_o	17.50	48.26	55.20	61.84	72.32	79.34	84.98	88.72	95.44	98.26	99.48	99.88	99.92	100.00
	Wald-t	1.72	13.86	18.50	24.82	37.20	45.02	53.88	64.48	78.58	89.74	96.14	98.34	99.16	100.00
	SSR-LRT	**3.12**	**21.86**	**26.70**	**34.38**	**47.68**	**56.14**	**64.96**	**73.68**	**85.88**	**93.62**	**97.78**	**99.10**	**99.62**	**100.00**
	Score	**4.38**	**24.64**	**30.74**	**38.48**	**52.18**	**60.18**	**68.24**	**77.22**	**87.98**	**94.66**	**98.08**	**99.34**	**99.68**	**100.00**
	Q *	15.04	44.38	51.38	58.10	69.50	77.00	83.04	87.04	94.10	97.68	99.26	99.82	99.88	100.00

หมายเหตุ Propose1 หมายถึง การทดสอบ $Z_{\bar{\hat{\mu}}}$ Propose2 หมายถึง การทดสอบ $Z_{\bar{Y}}$ Z_o หมายถึง การทดสอบ $\mathbf{Z_o}$ Wald-t หมายถึง การทดสอบวาลด์ที

SSR-LRT หมายถึง การทดสอบอัตราส่วนภาวะน่าจะเป็น Score หมายถึง การทดสอบสกอร์ Q* หมายถึง การทดสอบ Q^*

อักษรตัวหนา หมายถึง การทดสอบที่จะนำมาใช้ในเปรียบเทียบกำลังการทดสอบที่ได้จากการทดลอง

ตารางที่ 2 (ต่อ) กำลังการทดสอบ (ร้อยละ) ที่ได้จากการทดลองของการทดสอบต่าง ๆ ในกรณี Overdispersion โดยทำซ้ำ 5,000 ครั้ง

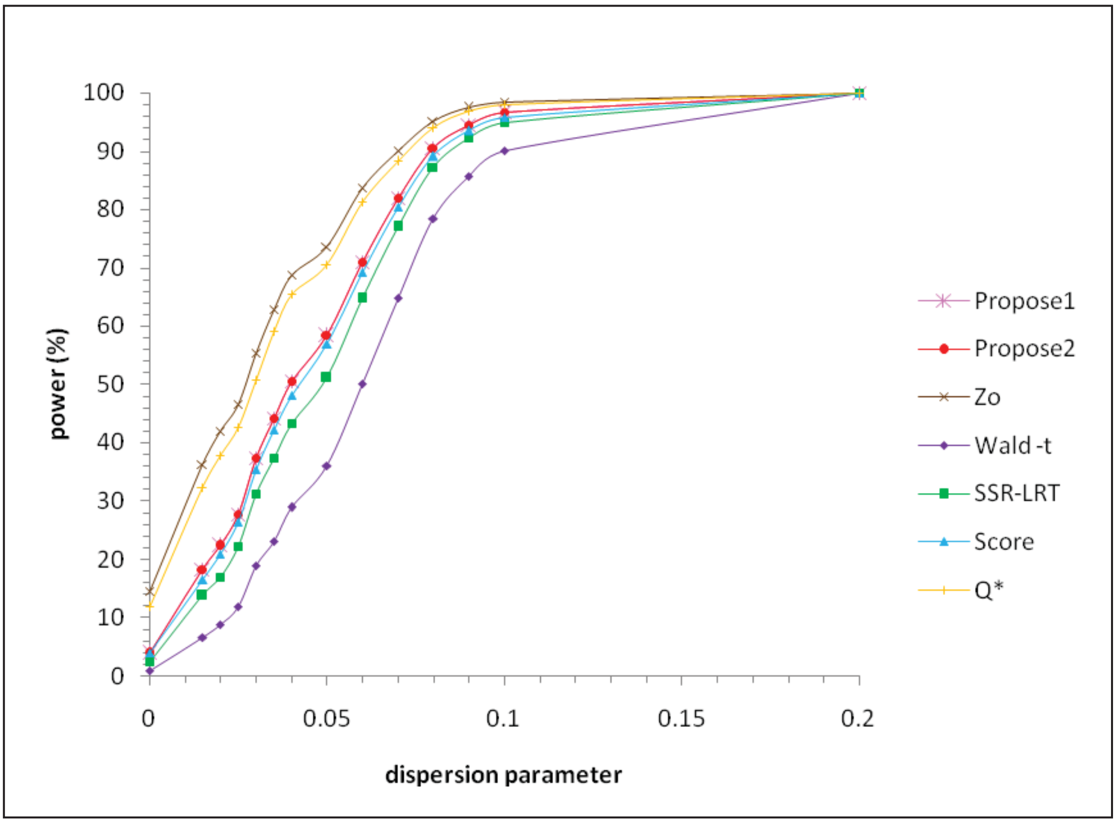
n	Method	Power (%)													
		$\varphi = 0.0$	0.015	0.02	0.025	0.03	0.035	0.04	0.05	0.06	0.07	0.08	0.09	0.10	0.20
100	**Propose1**	**4.86**	**40.42**	**50.44**	**62.00**	**78.66**	**87.00**	**92.62**	**96.32**	**98.94**	**99.86**	**100.00**	**100.00**	**100.00**	**100.00**
	Propose2	**4.86**	**40.42**	**50.44**	**62.00**	**78.66**	**87.00**	**92.62**	**96.32**	**98.94**	**99.86**	**100.00**	**100.00**	**100.00**	**100.00**
	Z_o	23.82	68.42	76.56	83.94	92.88	96.04	98.14	98.90	99.78	99.98	100.00	100.00	100.00	100.00
	Wald-t	2.28	29.66	38.66	49.86	69.76	78.80	86.86	93.04	98.04	99.50	99.96	100.00	100.00	100.00
	SSR-LRT	**4.00**	**36.08**	**46.28**	**57.60**	**74.96**	**84.08**	**90.68**	**95.22**	**98.70**	**99.78**	**100.00**	**100.00**	**100.00**	**100.00**
	Score	**5.06**	**38.98**	**49.40**	**60.80**	**77.48**	**85.80**	**91.78**	**96.04**	**98.88**	**99.80**	**100.00**	**100.00**	**100.00**	**100.00**
200	**Propose1**	**4.78**	**62.60**	**76.30**	**86.28**	**96.04**	**98.50**	**99.46**	**99.94**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
	Propose2	**4.78**	**62.60**	**76.30**	**86.28**	**96.04**	**98.50**	**99.46**	**99.94**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
	Z_o	34.34	90.30	94.40	97.64	99.52	99.88	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	Wald-t	2.98	53.60	68.42	81.30	94.34	97.58	99.10	99.86	100.00	100.00	100.00	100.00	100.00	100.00
	SSR-LRT	**4.18**	**59.10**	**73.58**	**84.54**	**95.36**	**67.34**	**99.34**	**99.92**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
	Score	**5.00**	**61.02**	**75.24**	**85.96**	**95.76**	**99.92**	**99.42**	**99.92**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**

หมายเหตุ Propose1 หมายถึง การทดสอบ $Z_{\bar{\hat{\mu}}}$ Propose2 หมายถึง การทดสอบ $Z_{\bar{Y}}$ Z_o หมายถึง การทดสอบ Z_o Wald-t หมายถึง การทดสอบวาลด์ที

SSR-LRT หมายถึง การทดสอบอัตราส่วนภาวะน่าจะเป็น Score หมายถึง การทดสอบสกอร์

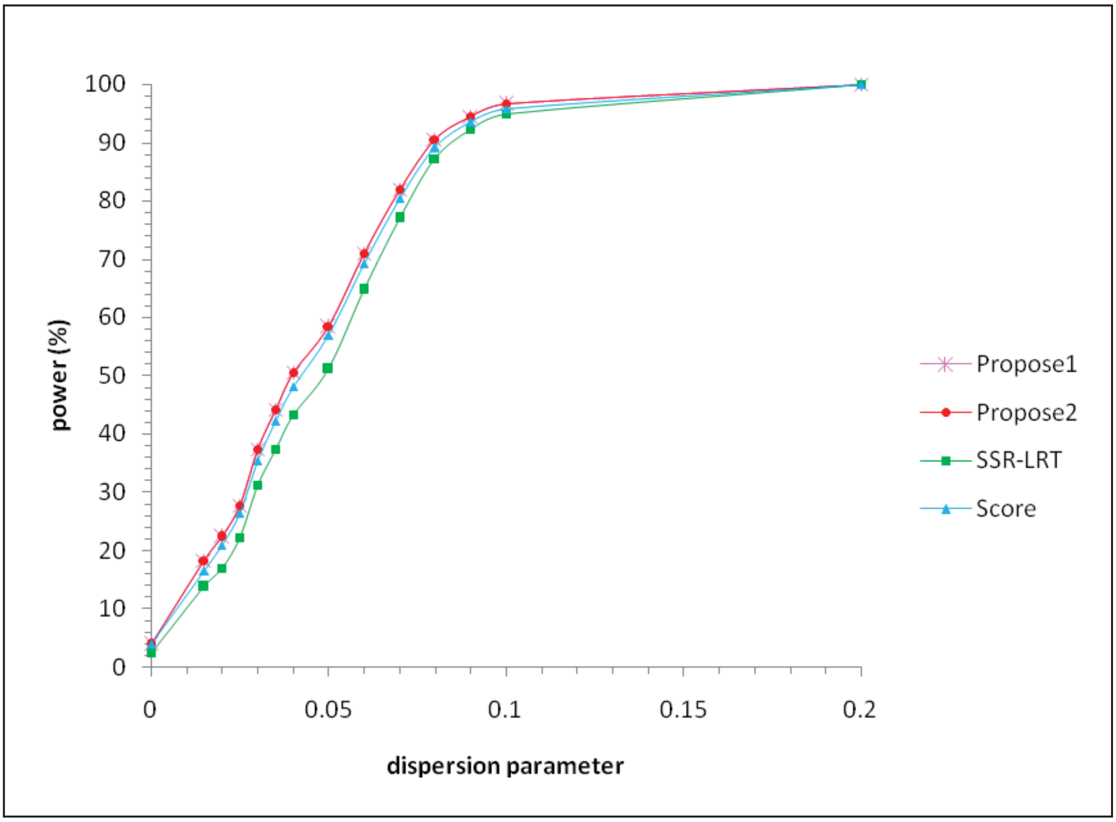
อักษรตัวหนา หมายถึง การทดสอบที่จะนำมาใช้ในเปรียบเทียบกำลังการทดสอบที่ได้จากการทดลอง

ตารางที่ 2 แสดงกำลังการทดสอบ (ร้อยละ) ที่ได้จากการทดลองของการทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) การทดสอบ $Z_{\mathbf{0}}$ (Z_o) การทดสอบวาลด์ที (Wald-t) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) การทดสอบสกอร์ (Score) และการทดสอบ Q^* (Q*) ในกรณี Overdispersion โดยทำซ้ำ 5,000 ครั้ง ซึ่งจากตารางพบว่า เมื่อพิจารณาที่ $\varphi=0$ ตัวแบบที่ถูกต้องคือ ตัวแบบการถดถอยปัวซง การตีความหมายของการทดสอบต่าง ๆ พบว่า การทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบสกอร์ (Score) มีแนวโน้มที่จะให้ค่า $\alpha=0.05$ เมื่อ $\varphi=0$ นั่นคือ การทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบสกอร์ (Score) เหมาะสมเมื่อ $\varphi=0$ ส่วนการทดสอบ $Z_{\mathbf{0}}$ (Z_o) การทดสอบวาลด์ที (Wald-t) และการทดสอบ Q^* (Q*) ให้ผลแย่ในสถานการณ์นี้ เนื่องจากการทดสอบ $Z_{\mathbf{0}}$ (Z_o) การทดสอบวาลด์ที (Wald-t) และการทดสอบ Q^* (Q*) ให้กำลังการทดสอบที่ได้จากการทดลองไม่ใกล้เคียงกับ $\alpha=0.05$ และเมื่อพิจารณากำลังการทดสอบที่ได้จากการทดลองของการทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบสกอร์ (Score) พบว่า การทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) ที่ผู้วิจัยเสนอขึ้นให้กำลังการทดสอบที่ได้จากการทดลองใกล้เคียงกันและการทดสอบทั้งสองยังให้กำลังการทดสอบที่ได้จากการทดลองมากกว่าการทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบสกอร์ (Score)



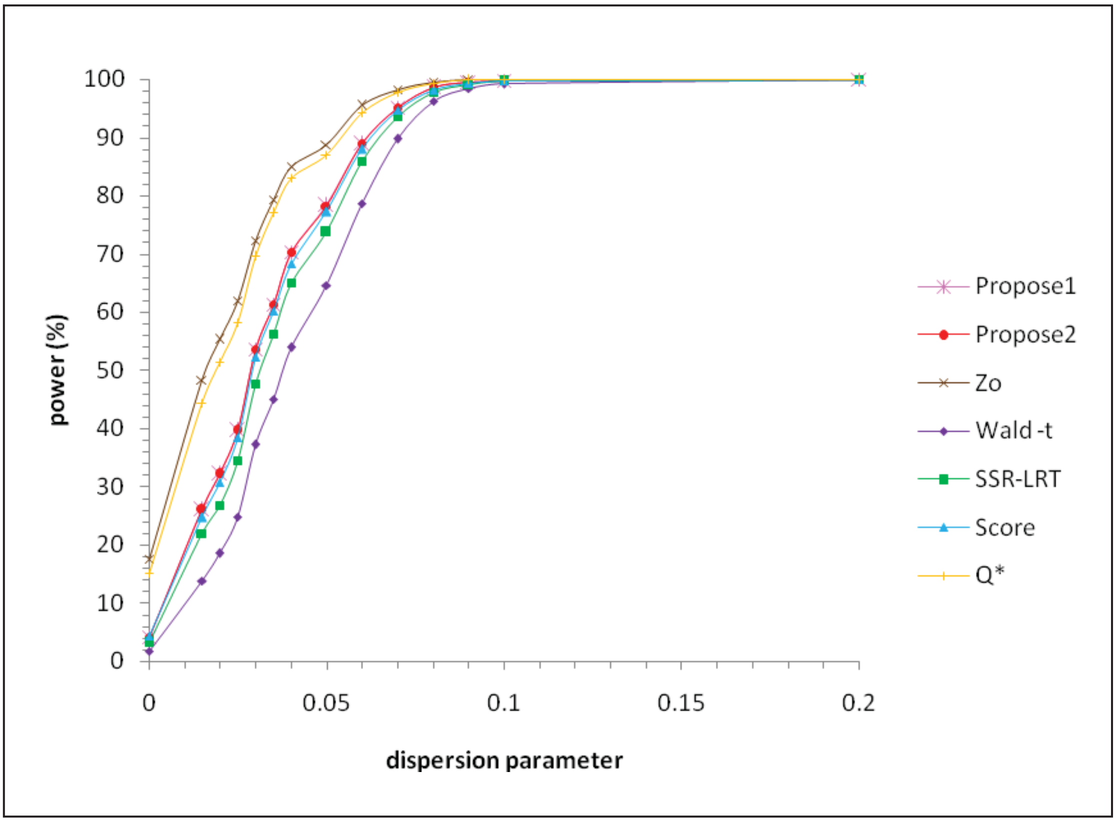
ภาพที่ 2 เส้นโค้งของกำลังการทดสอบที่ได้จากการทดลองของการทดสอบต่าง ๆ ในกรณี Overdispersion เมื่อกำหนดขนาดตัวอย่างเท่ากับ 30 โดยทำซ้ำ 5,000 ครั้ง

ภาพที่ 2 แสดงเส้นโค้งของกำลังการทดสอบที่ได้จากการทดลองของการทดสอบ $Z_{\bar{\mu}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) การทดสอบ Z_o (Z_o) การทดสอบวาลด์ที (Wald-t) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) การทดสอบสกอร์ (Score) และการทดสอบ Q^* (Q*) ในกรณี Overdispersion เมื่อกำหนดขนาดตัวอย่างเท่ากับ 30 โดยทำซ้ำ 5,000 ครั้ง ซึ่งเมื่อพิจารณาที่ $\varphi=0$ ตัวแบบที่ถูกต้องคือ ตัวแบบการถดถอยปัวซง การตีความหมายของการทดสอบต่าง ๆ พบว่า การทดสอบ $Z_{\bar{\mu}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบสกอร์ (Score) มีแนวโน้มที่จะให้ค่า $\alpha=0.05$ เมื่อ $\varphi=0$ นั่นคือ การทดสอบ $Z_{\bar{\mu}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบสกอร์ (Score) เหมาะสมเมื่อ $\varphi=0$ ส่วนการทดสอบ Z_o (Z_o) การทดสอบวาลด์ที (Wald-t) และการทดสอบ Q^* (Q*)ให้ผลแย่ในสถานการณ์นี้เนื่องจากการทดสอบ Z_o (Z_o) การทดสอบวาลด์ที (Wald-t) และการทดสอบ Q^* (Q*) ให้กำลังการทดสอบที่ได้จากการทดลองไม่ใกล้เคียงกับ $\alpha=0.05$ ดังนั้น ในกรณี Overdispersion เมื่อกำหนดขนาดตัวอย่างเท่ากับ 30 จะศึกษากำลังการทดสอบที่ได้จากการทดลองของการทดสอบ $Z_{\bar{\mu}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบสกอร์ (Score) เท่านั้น (แสดงดังภาพที่ 3)



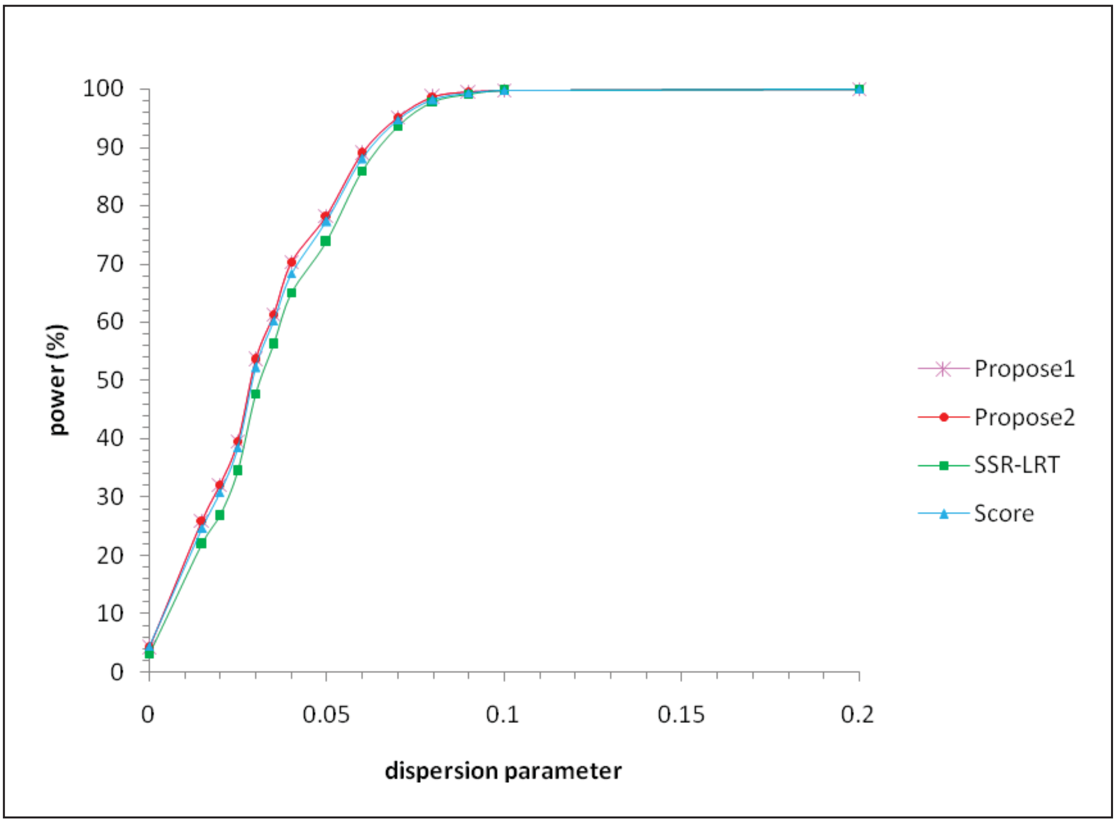
ภาพที่ 3 เส้นโค้งของกำลังการทดสอบในกรณี Overdispersion ของการทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบสกอร์ (Score) เมื่อกำหนดขนาดตัวอย่างเท่ากับ 30 โดยทำซ้ำ 5,000 ครั้ง

ภาพที่ 3 แสดงเส้นโค้งของกำลังการทดสอบที่ได้จากการทดลองของการทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบสกอร์ (Score) ในกรณี Overdispersion เมื่อกำหนดขนาดตัวอย่างเท่ากับ 30 โดยทำซ้ำ 5,000 ครั้ง เมื่อพิจารณาภาพที่ 3 พบว่า การทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) และการทดสอบ $Z_{\bar{Y}}$ (Propose2) ที่ผู้วิจัยเสนอให้กำลังการทดสอบที่ได้จากการทดลองใกล้เคียงกันและการทดสอบทั้งสองยังให้กำลังการทดสอบที่ได้จากการทดลองมากกว่าการทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบสกอร์ (Score) และเมื่อพิจารณาอิทธิพลของขนาดตัวอย่างที่มีผลต่อกำลังการทดสอบที่ได้จากการทดลองพบว่า เมื่อตัวอย่างมีขนาดเล็กคือ $n = 30$ กำลังการทดสอบที่ได้จากการทดลองจะเพิ่มขึ้นอย่างช้า ๆ ในช่วงแรก และเมื่อ $\varphi > 0.02$ กำลังการทดสอบจะเพิ่มขึ้นอย่างรวดเร็วและเข้าใกล้ 1.0 เมื่อ $\varphi = 0.2$



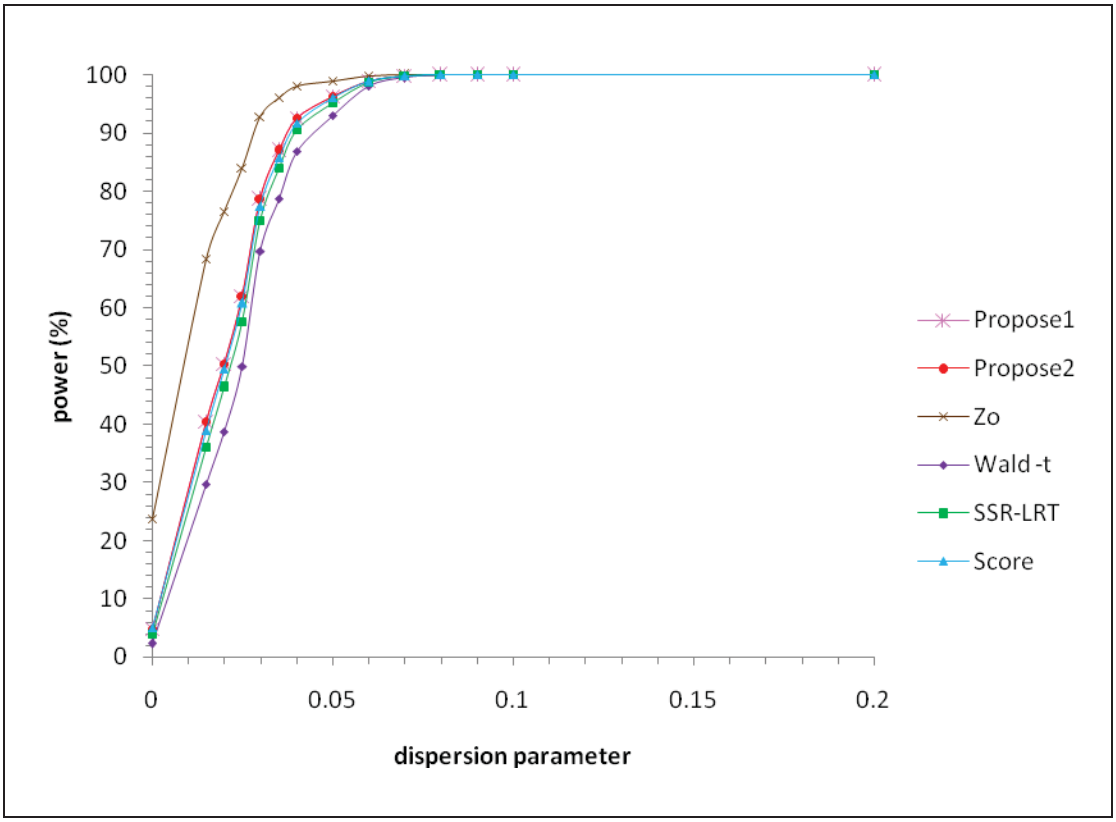
ภาพที่ 4 เส้นโค้งของกำลังการทดสอบที่ได้จากการทดลองของการทดสอบต่าง ๆ ในกรณี Overdispersion เมื่อกำหนดขนาดตัวอย่างเท่ากับ 50 โดยทำซ้ำ 5,000 ครั้ง

ภาพที่ 4 แสดงเส้นโค้งของกำลังการทดสอบที่ได้จากการทดลองของการทดสอบ $Z_{\bar{\mu}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) การทดสอบ $Z_{\mathbf{o}}$ (Z_o) การทดสอบวาลด์ที (Wald-t) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) การทดสอบสกอร์ (Score) และการทดสอบ Q^* (Q*)ในกรณี Overdispersion เมื่อกำหนดขนาดตัวอย่างเท่ากับ 50 โดยทำซ้ำ 5,000 ครั้ง ซึ่งเมื่อพิจารณาที่ $\varphi=0$ ตัวแบบที่ถูกต้องคือ ตัวแบบการถดถอยปัวซง การตีความหมายของการทดสอบต่าง ๆ พบว่า การทดสอบ $Z_{\bar{\mu}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบสกอร์ (Score) มีแนวโน้มที่จะให้ค่า $\alpha=0.05$ เมื่อ $\varphi=0$ นั่นคือ การทดสอบ $Z_{\bar{\mu}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบสกอร์ (Score) เหมาะสมเมื่อ $\varphi=0$ ส่วนการทดสอบ $Z_{\mathbf{o}}$ (Z_o) การทดสอบวาลด์ที (Wald-t) และการทดสอบ Q^* (Q*) ให้ผลแย่ในสถานการณ์นี้ เนื่องจากการทดสอบ $Z_{\mathbf{o}}$ (Z_o) การทดสอบวาลด์ที (Wald-t) และการทดสอบ Q^* (Q*) ให้กำลังการทดสอบที่ได้จากการทดลองไม่ใกล้เคียงกับ $\alpha=0.05$ ดังนั้นในกรณี Overdispersion เมื่อกำหนดขนาดตัวอย่างเท่ากับ 50 จะศึกษากำลังการทดสอบที่ได้จากการทดลองของการทดสอบ $Z_{\bar{\mu}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบสกอร์ (Score) เท่านั้น (แสดงดังภาพที่ 5)



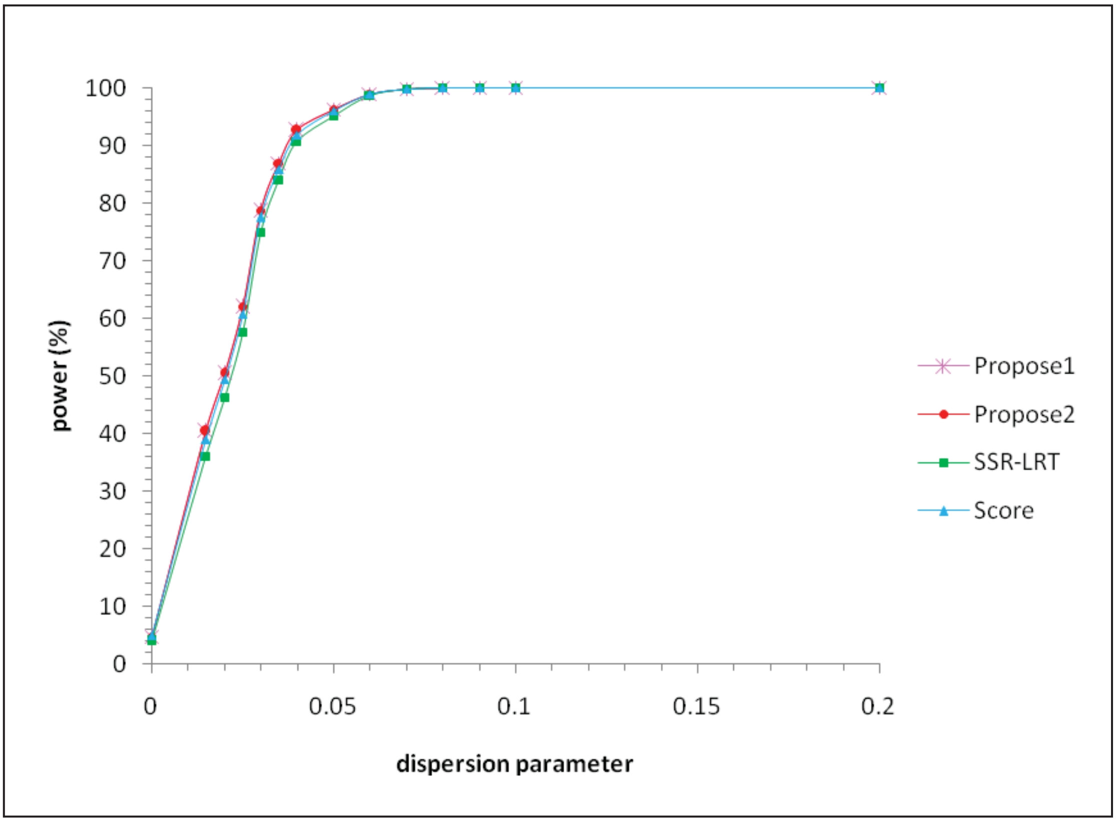
ภาพที่ 5 เส้นโค้งของกำลังการทดสอบในกรณี Overdispersion ของการทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบสกอร์ (Score) เมื่อกำหนดขนาดตัวอย่างเท่ากับ 50 โดยทำซ้ำ 5,000 ครั้ง

ภาพที่ 5 แสดงเส้นโค้งของกำลังการทดสอบที่ได้จากการทดลองของการทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบสกอร์ (Score) ในกรณี Overdispersion เมื่อกำหนดขนาดตัวอย่างเท่ากับ 50 โดยทำซ้ำ 5,000 ครั้ง เมื่อพิจารณาภาพที่ 5 พบว่า การทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) และการทดสอบ $Z_{\bar{Y}}$ (Propose2) ที่ผู้วิจัยเสนอให้กำลังการทดสอบที่ได้จากการทดลองใกล้เคียงกันและการทดสอบทั้งสองยังให้กำลังการทดสอบที่ได้จากการทดลองมากกว่าการทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบสกอร์ (Score) และเมื่อพิจารณาอิทธิพลของขนาดตัวอย่างที่มีผลต่อกำลังการทดสอบที่ได้จากการทดลองพบว่า เมื่อ $n = 50$ กำลังการทดสอบที่ได้จากการทดลองจะเพิ่มขึ้นอย่างรวดเร็วและเข้าใกล้ 1.0 เมื่อ $\varphi = 0.2$



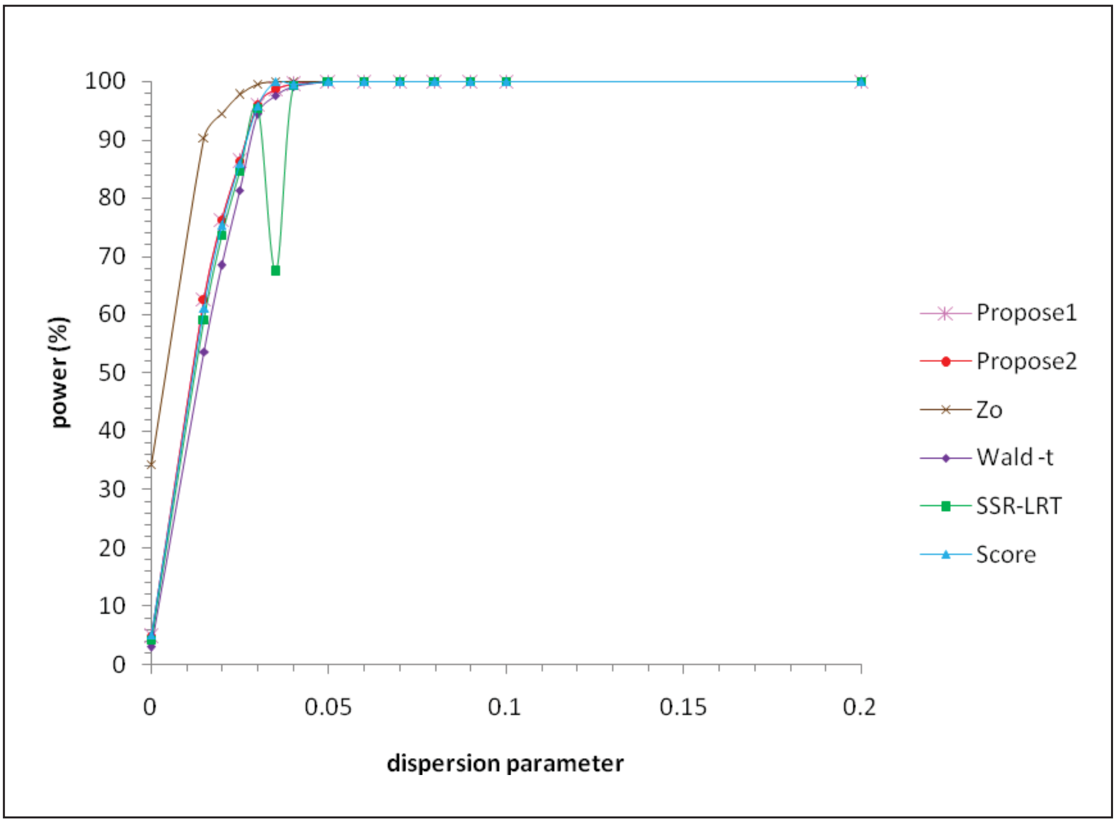
ภาพที่ 6 เส้นโค้งของกำลังการทดสอบที่ได้จากการทดลองของการทดสอบต่าง ๆ ในกรณี Overdispersion เมื่อกำหนดขนาดตัวอย่างเท่ากับ 100 โดยทำซ้ำ 5,000 ครั้ง

ภาพที่ 6 แสดงเส้นโค้งของกำลังการทดสอบที่ได้จากการทดลองของการทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) การทดสอบ $Z_{\mathbf{o}}$ (Z_o) การทดสอบวาลด์ที (Wald-t) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบสกอร์ (Score) ในกรณี Overdispersion เมื่อกำหนดขนาดตัวอย่างเท่ากับ 100 โดยทำซ้ำ 5,000 ครั้ง ซึ่งเมื่อพิจารณาที่ $\varphi=0$ ตัวแบบที่ถูกต้องคือ ตัวแบบการถดถอยปัวซง การตีความหมายของการทดสอบต่าง ๆ พบว่า การทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบสกอร์ (Score) มีแนวโน้มที่จะให้ค่า $\alpha=0.05$ เมื่อ $\varphi=0$ นั่นคือ การทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบสกอร์ (Score) เหมาะสมเมื่อ $\varphi=0$ ส่วนการทดสอบ $Z_{\mathbf{o}}$ (Z_o) และการทดสอบวาลด์ที (Wald-t) ให้ผลแย่ในสถานการณ์นี้ เนื่องจากการทดสอบ $Z_{\mathbf{o}}$ (Z_o) และการทดสอบวาลด์ที (Wald-t) ให้กำลังการทดสอบที่ได้จากการทดลองไม่ใกล้เคียงกับ $\alpha=0.05$ ดังนั้น ในกรณี Overdispersion เมื่อกำหนดขนาดตัวอย่างเท่ากับ 100 จะศึกษากำลังการทดสอบที่ได้จากการทดลองของการทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบสกอร์ (Score) เท่านั้น (แสดงดังภาพที่ 7)



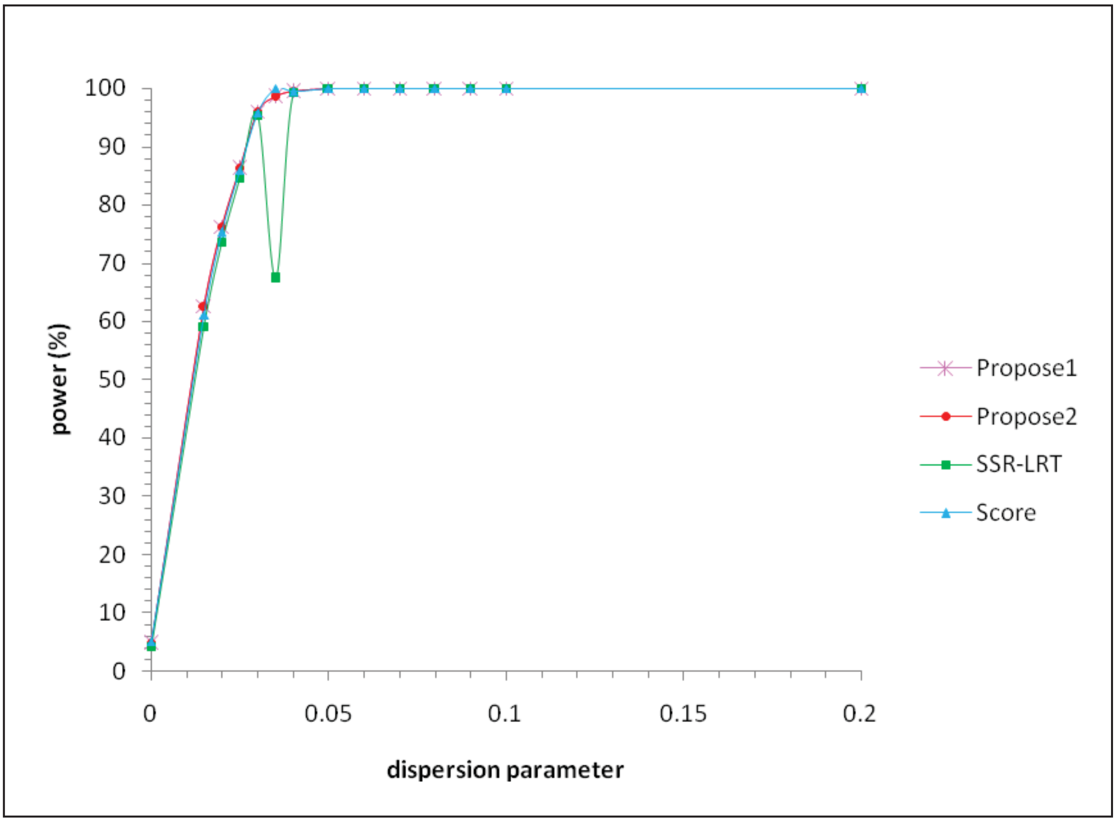
ภาพที่ 7 เส้นโค้งของกำลังการทดสอบในกรณี Overdispersion ของการทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบสกอร์ (Score) เมื่อกำหนดขนาดตัวอย่างเท่ากับ 100 โดยทำซ้ำ 5,000 ครั้ง

ภาพที่ 7 แสดงเส้นโค้งของกำลังการทดสอบที่ได้จากการทดลองของการทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบสกอร์ (Score) ในกรณี Overdispersion เมื่อกำหนดขนาดตัวอย่างเท่ากับ 100 โดยทำซ้ำ 5,000 ครั้ง เมื่อพิจารณาภาพที่ 7 พบว่า การทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) และการทดสอบ $Z_{\bar{Y}}$ (Propose2) ที่ผู้วิจัยเสนอให้กำลังการทดสอบที่ได้จากการทดลองใกล้เคียงกันและการทดสอบทั้งสองยังให้กำลังการทดสอบที่ได้จากการทดลองมากกว่าการทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบสกอร์ (Score) และเมื่อพิจารณาอิทธิพลของขนาดตัวอย่างที่มีผลต่อกำลังการทดสอบที่ได้จากการทดลองพบว่า เมื่อตัวอย่างมีขนาดใหญ่คือ $n = 100$ กำลังการทดสอบจะเพิ่มขึ้นอย่างรวดเร็วและเข้าใกล้ 1.0 เมื่อ $\varphi = 0.09$



ภาพที่ 8 เส้นโค้งของกำลังการทดสอบที่ได้จากการทดลองของการทดสอบต่าง ๆ ในกรณี Overdispersion เมื่อกำหนดขนาดตัวอย่างเท่ากับ 200 โดยทำซ้ำ 5,000 ครั้ง

ภาพที่ 8 แสดงเส้นโค้งของกำลังการทดสอบที่ได้จากการทดลองของการทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) การทดสอบ $Z_{\mathbf{o}}$ (Z_o) การทดสอบวาลด์ที (Wald-t) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบสกอร์ (Score) ในกรณี Overdispersion เมื่อกำหนดขนาดตัวอย่างเท่ากับ 200 โดยทำซ้ำ 5,000 ครั้ง ซึ่งเมื่อพิจารณาที่ $\varphi=0$ ตัวแบบที่ถูกต้องคือ ตัวแบบการถดถอยปัวซง การตีความหมายของการทดสอบต่าง ๆ พบว่า การทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบสกอร์ (Score) มีแนวโน้มที่จะให้ค่า $\alpha=0.05$ เมื่อ $\varphi=0$ นั่นคือ การทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบสกอร์ (Score) เหมาะสมเมื่อ $\varphi=0$ ส่วนการทดสอบ $Z_{\mathbf{o}}$ (Z_o) และการทดสอบวาลด์ที (Wald-t) ให้ผลแย่ในสถานการณ์นี้ เนื่องจากการทดสอบ $Z_{\mathbf{o}}$ (Z_o) และการทดสอบวาลด์ที (Wald-t) ให้กำลังการทดสอบที่ได้จากการทดลองไม่ใกล้เคียงกับ $\alpha=0.05$ ดังนั้น ในกรณี Overdispersion เมื่อกำหนดขนาดตัวอย่างเท่ากับ 200 จะศึกษากำลังการทดสอบที่ได้จากการทดลองของการทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบสกอร์ (Score) เท่านั้น (แสดงดังภาพที่ 9)



ภาพที่ 9 เส้นโค้งของกำลังการทดสอบในกรณี Overdispersion ของการทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบสกอร์ (Score) เมื่อกำหนดขนาดตัวอย่างเท่ากับ 200 โดยทำซ้ำ 5,000 ครั้ง

ภาพที่ 9 แสดงเส้นโค้งของกำลังการทดสอบที่ได้จากการทดลองของการทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบสกอร์ (Score) ในกรณี Overdispersion เมื่อกำหนดขนาดตัวอย่างเท่ากับ 200 โดยทำซ้ำ 5,000 ครั้ง เมื่อพิจารณาภาพที่ 9 พบว่า การทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) และการทดสอบ $Z_{\bar{Y}}$ (Propose2) ที่ผู้วิจัยเสนอให้กำลังการทดสอบที่ได้จากการทดลองใกล้เคียงกันและการทดสอบทั้งสองยังให้กำลังการทดสอบที่ได้จากการทดลองมากกว่าการทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบสกอร์ (Score) นอกจากนั้นยังพบว่า เมื่อตัวอย่างมีขนาดใหญ่ ไม่ควรเลือกใช้การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) เนื่องจากการทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) ให้กำลังการทดสอบที่ได้จากการทดลองต่ำแบบผิดปกติ เมื่อค่าพารามิเตอร์การกระจายเท่ากับ 0.035 และเมื่อพิจารณาอิทธิพลของขนาดตัวอย่างที่มีผลต่อกำลังการทดสอบที่ได้จากการทดลองพบว่า เมื่อตัวอย่างมีขนาดใหญ่คือ $n = 200$ กำลังการทดสอบจะเพิ่มขึ้นอย่างรวดเร็วและเข้าใกล้ 1.0 เมื่อ $\varphi = 0.06$

ในกรณี Overdispersion นี้สามารถวิเคราะห์ผลลัพธ์ต่าง ๆ ได้ดังนี้ ผลลัพธ์ในตารางที่ 2 และภาพที่ 2-9 พบว่า เมื่อ $\varphi=0$ ตัวแบบที่ถูกต้องคือ ตัวแบบการถดถอยปัวซง การตีความหมายของการทดสอบต่าง ๆ พบว่า การทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบสกอร์ (Score) เหมาะสม ส่วนการทดสอบ $\mathbf{Z_0}$ (Z_o) การทดสอบวาลด์ที (Wald-t) และการทดสอบ Q^* (Q*) ให้ผลแย่ในสถานการณ์นี้ เมื่อตัวอย่างมีขนาดใหญ่ การทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบสกอร์ (Score) ยังมีแนวโน้มที่จะให้ค่าประมาณของ α ได้อย่างใกล้เคียงกับค่า $\alpha=0.05$ เมื่อ $\varphi=0$ ดังนั้น การทดสอบที่สามารถนำมาใช้ในการทดสอบ Overdispersion ได้แก่ การทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบสกอร์ (Score) แต่เมื่อพิจารณากำลังการทดสอบที่ได้จากการทดลองของการทดสอบทั้งสี่พบว่า กำลังการทดสอบที่ได้จากการทดลองของการทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) และการทดสอบ $Z_{\bar{Y}}$ (Propose2) ที่ผู้วิจัยเสนอให้กำลังการทดสอบที่ได้จากการทดลองใกล้เคียงกันและการทดสอบทั้งสองยังให้กำลังการทดสอบที่ได้จากการทดลองมากกว่าการทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบสกอร์ (Score) ในทุกขนาดตัวอย่าง นอกจากนั้นยังพบว่า เมื่อตัวอย่างมีขนาดใหญ่ ไม่ควรเลือกใช้การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) เนื่องจากการทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) ให้กำลังการทดสอบที่ได้จากการทดลองต่ำแบบผิดปกติ เมื่อค่าพารามิเตอร์การกระจายเท่ากับ 0.035 และเมื่อพิจารณาอิทธิพลของขนาดตัวอย่างที่มีผลต่อกำลังการทดสอบที่ได้จากการทดลองจะพบว่า เมื่อตัวอย่างมีขนาดใหญ่คือ $n=100$ และ $n=200$ กำลังการทดสอบจะเพิ่มขึ้นอย่างรวดเร็วและเข้าใกล้ 1.0 เมื่อ $\varphi=0.09$ และ $\varphi=0.06$ ตามลำดับ และเมื่อ $n=50$ กำลังการทดสอบจะเข้าใกล้ 1.0 เมื่อ $\varphi=0.2$ และเมื่อตัวอย่างมีขนาดเล็กคือ $n=30$ กำลังการทดสอบจะเพิ่มขึ้นอย่างช้า ๆ ในช่วงแรกและเมื่อ $\varphi>0.02$ กำลังการทดสอบจะเพิ่มขึ้นอย่างรวดเร็วและเข้าใกล้ 1.0 เมื่อ $\varphi=0.2$

2. กรณี Underdispersion

ตารางที่ 3 กำลังการทดสอบ (ร้อยละ) ที่ได้จากการทดลองของการทดสอบต่าง ๆ ในกรณี Underdispersion โดยทำซ้ำ 5,000 ครั้ง

n	Method	Power (%)													
		$\varphi = 0.0$	-0.015	-0.02	-0.025	-0.03	-0.035	-0.04	-0.05	-0.06	-0.07	-0.08	-0.09	-0.10	-0.20
30	**Propose1**	**4.16**	**11.12**	**17.30**	**24.80**	**32.02**	**42.08**	**53.66**	**73.06**	**90.08**	**97.92**	**99.76**	**99.96**	**100.00**	**100.00**
	Propose2	**4.14**	**11.10**	**17.30**	**24.82**	**32.02**	**42.08**	**53.68**	**73.06**	**90.08**	**97.92**	**99.76**	**99.98**	**100.00**	**100.00**
	Z_o	1.06	2.38	4.90	7.88	9.78	15.36	21.80	36.88	57.78	77.64	91.48	97.86	99.70	99.98
	Wald-t	15.74	32.20	40.82	51.76	61.74	72.70	81.20	92.54	98.60	99.86	100.00	100.00	100.00	100.00
	SSR-LRT	10.08	22.26	29.74	40.46	49.70	60.84	71.56	86.84	97.02	99.44	99.96	99.86	100.00	99.86
	Score	**4.70**	**12.06**	**18.26**	**26.94**	**34.32**	**45.16**	**57.06**	**76.74**	**92.82**	**98.52**	**99.86**	**99.98**	**100.00**	**100.00**
	Q*	2.24	4.66	8.48	13.16	16.72	24.22	32.56	50.08	71.30	86.84	96.20	99.20	99.92	100.00
50	**Propose1**	**4.96**	**16.88**	**25.40**	**39.08**	**48.96**	**64.70**	**77.72**	**92.72**	**99.20**	**99.98**	**100.00**	**100.00**	**100.00**	**100.00**
	Propose2	**4.96**	**16.86**	**25.40**	**39.10**	**48.98**	**64.68**	**77.72**	**92.72**	**99.20**	**99.98**	**100.00**	**100.00**	**100.00**	**100.00**
	Z_o	0.88	3.40	6.26	12.74	18.06	27.40	39.04	62.70	86.70	96.92	99.68	100.00	100.00	100.00
	Wald-t	12.36	33.94	45.36	60.04	71.70	83.12	91.06	98.20	99.86	100.00	100.00	100.00	100.00	100.00
	SSR-LRT	8.60	26.24	35.68	51.02	62.38	76.76	86.64	96.52	99.72	100.00	100.00	100.00	100.00	100.00
	Score	**5.30**	**17.86**	**26.08**	**41.48**	**52.04**	**67.30**	**80.14**	**94.30**	**99.44**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
	Q*	1.74	5.20	9.24	17.98	24.04	35.88	48.58	71.70	91.42	98.38	99.88	100.00	100.00	100.00

หมายเหตุ Propose1 หมายถึง การทดสอบ $Z_{\bar{\hat{\mu}}}$ Propose2 หมายถึง การทดสอบ $Z_{\bar{Y}}$ Z_o หมายถึง การทดสอบ $\mathbf{Z_o}$ Wald-t หมายถึง การทดสอบวาลด์ที

SSR-LRT หมายถึง การทดสอบอัตราส่วนภาวะน่าจะเป็น Score หมายถึง การทดสอบสกอร์ Q* หมายถึง การทดสอบ Q^*

อักษรตัวหนา หมายถึง การทดสอบที่จะนำมาใช้ในเปรียบเทียบกำลังการทดสอบที่ได้จากการทดลอง

ตารางที่ 3 (ต่อ) กำลังการทดสอบ (ร้อยละ) ที่ได้จากการทดลองของการทดสอบต่าง ๆ ในกรณี Underdispersion โดยทำซ้ำ 5,000 ครั้ง

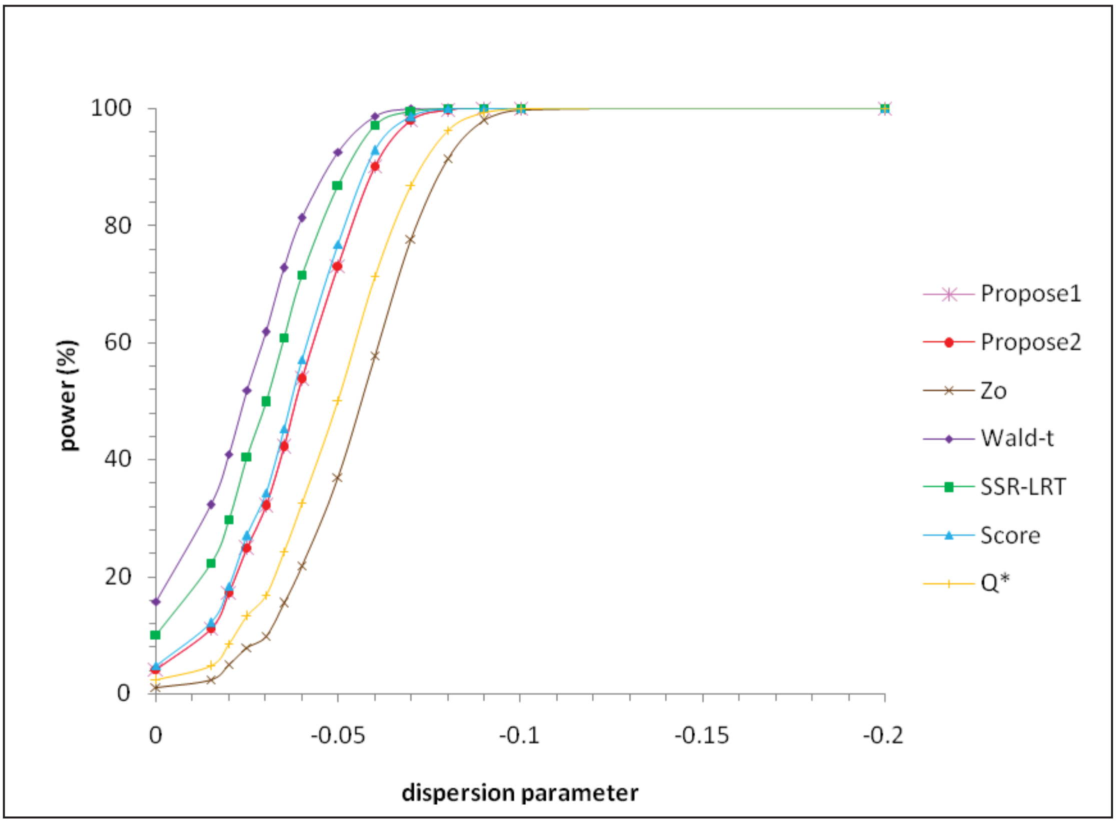
n	Method	Power (%)													
		$\varphi = 0.0$	-0.015	-0.02	-0.025	-0.03	-0.035	-0.04	-0.05	-0.06	-0.07	-0.08	-0.09	-0.10	-0.20
100	**Propose1**	**4.88**	**26.24**	**42.20**	**65.14**	**79.16**	**91.06**	**97.32**	**99.90**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
	Propose2	**4.88**	**26.24**	**42.20**	**65.14**	**79.12**	**91.06**	**97.32**	**99.90**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
	Z_o	0.58	4.32	9.84	22.42	33.52	52.86	70.20	92.70	99.50	99.98	100.00	100.00	100.00	100.00
	Wald-t	9.72	40.00	57.68	78.58	89.24	96.26	99.26	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	SSR-LRT	7.28	32.90	50.58	72.96	85.42	94.40	98.68	99.94	100.00	100.00	100.00	100.00	100.00	100.00
	Score	**5.02**	**27.42**	**43.50**	**66.68**	**80.70**	**92.04**	**97.72**	**99.92**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
200	**Propose1**	**5.24**	**44.24**	**68.28**	**89.60**	**97.08**	**99.64**	**99.96**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
	Propose2	**5.24**	**44.24**	**68.28**	**89.60**	**97.08**	**99.64**	**99.96**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
	Z_o	0.28	5.36	16.26	41.04	58.82	82.18	94.80	99.76	100.00	100.00	100.00	100.00	100.00	100.00
	Wald-t	8.18	54.64	77.38	94.58	98.66	99.84	99.98	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	SSR-LRT	6.70	50.10	72.74	92.62	98.26	99.82	83.60	98.06	100.00	100.00	100.00	100.00	100.00	100.00
	Score	**5.26**	**46.12**	**69.26**	**90.66**	**97.50**	**99.76**	**51.62**	**94.10**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**

หมายเหตุ Propose1 หมายถึง การทดสอบ $Z_{\bar{\hat{\mu}}}$ Propose2 หมายถึง การทดสอบ $Z_{\overline{Y}}$ Z_o หมายถึง การทดสอบ Z_o Wald-t หมายถึง การทดสอบวาลด์ที

SSR-LRT หมายถึง การทดสอบอัตราส่วนภาวะน่าจะเป็น Score หมายถึง การทดสอบสกอร์

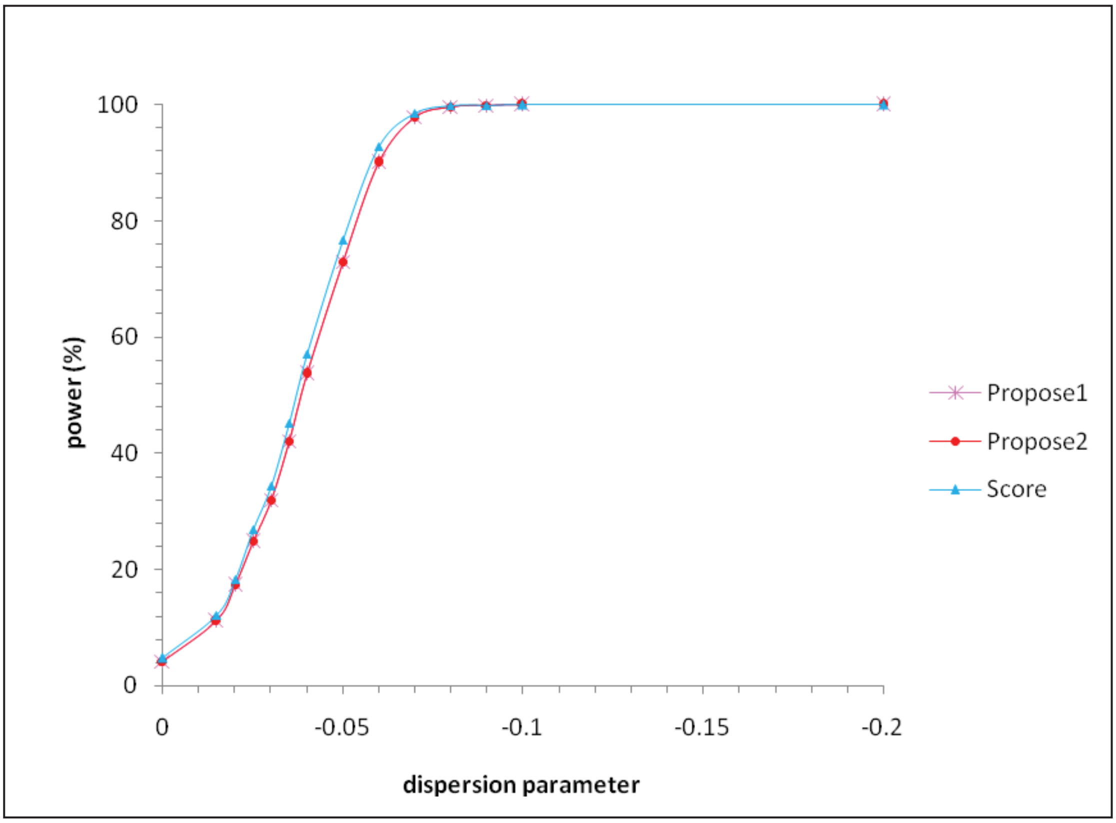
อักษรตัวหนา หมายถึง การทดสอบที่จะนำมาใช้ในเปรียบเทียบกำลังการทดสอบที่ได้จากการทดลอง

ตารางที่ 3 แสดงกำลังการทดสอบ (ร้อยละ) ที่ได้จากการทดลองของการทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) การทดสอบ $Z_{\mathbf{0}}$ (Z_o) การทดสอบวาลด์ที (Wald-t) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) การทดสอบสกอร์(Score) และการทดสอบ Q^* (Q*) ในกรณี Underdispersion โดยทำซ้ำ 5,000 ครั้ง ซึ่งจากตารางพบว่า เมื่อพิจารณาที่ $\varphi=0$ ตัวแบบที่ถูกต้องคือ ตัวแบบการถดถอยปัวซง การตีความหมายของการทดสอบต่าง ๆ พบว่า การทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) และการทดสอบสกอร์ (Score) มีแนวโน้มที่จะให้ค่า $\alpha=0.05$ เมื่อ $\varphi=0$ นั่นคือการทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) และการทดสอบสกอร์ (Score) เหมาะสมเมื่อ $\varphi=0$ ส่วนการทดสอบ $Z_{\mathbf{0}}$ (Z_o) การทดสอบวาลด์ที (Wald-t) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบ Q^* (Q*) ให้ผลแย่ในสถานการณ์นี้ เนื่องจากการทดสอบ $Z_{\mathbf{0}}$ (Z_o) การทดสอบวาลด์ที (Wald-t) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบ Q^* (Q*) ให้กำลังการทดสอบที่ได้จากการทดลองไม่ใกล้เคียงกับ $\alpha=0.05$ และเมื่อพิจารณากำลังการทดสอบที่ได้จากการทดลองของการทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) และการทดสอบสกอร์ (Score) พบว่า เมื่อตัวอย่างมีขนาดเล็ก การทดสอบสกอร์ (Score) จะให้กำลังการทดสอบที่ได้จากการทดลองมากกว่าการทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) และการทดสอบ $Z_{\bar{Y}}$ (Propose2) แต่เมื่อตัวอย่างมีขนาดใหญ่ การทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) และการทดสอบ $Z_{\bar{Y}}$ (Propose2) ที่ผู้วิจัยเสนอขึ้นให้กำลังการทดสอบที่ได้จากการทดลองใกล้เคียงกับการทดสอบสกอร์ (Score) นอกจากนั้นยังพบว่า การทดสอบสกอร์ (Score) ให้กำลังการทดสอบที่ได้จากการทดลองต่ำแบบผิดปกติในบางกรณี ดังนั้น เมื่อตัวอย่างมีขนาดใหญ่ การทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) หรือการทดสอบ $Z_{\bar{Y}}$ (Propose2) ที่ผู้วิจัยเสนอขึ้น เหมาะสมกว่าการทดสอบสกอร์ (Score)



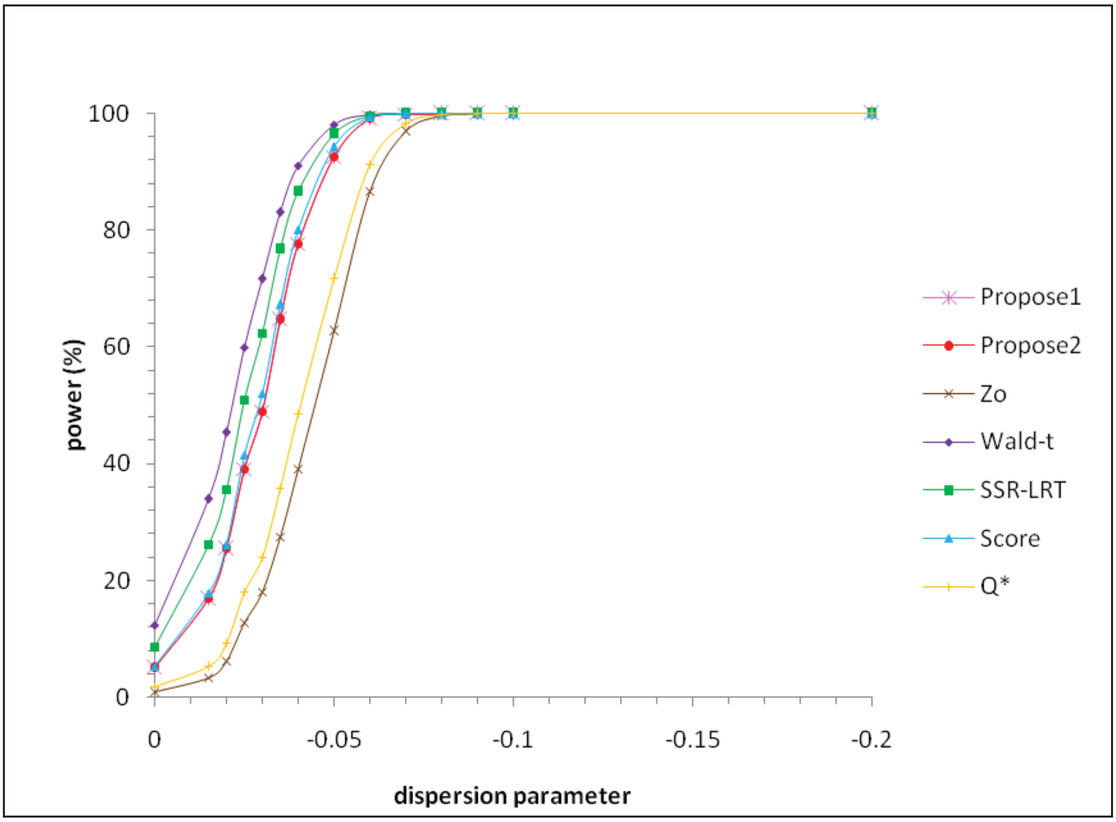
ภาพที่10 เส้นโค้งของกำลังการทดสอบที่ได้จากการทดลองของการทดสอบต่าง ๆ ในกรณี Underdispersion เมื่อกำหนดขนาดตัวอย่างเท่ากับ 30 โดยทำซ้ำ 5,000 ครั้ง

ภาพที่ 10 แสดงเส้นโค้งของกำลังการทดสอบที่ได้จากการทดลองของการทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) การทดสอบ $\mathbf{Z_o}$ (Z_o) การทดสอบวาลด์ที (Wald-t) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) การทดสอบสกอร์ (Score) และการทดสอบ Q^* (Q*) ในกรณี Underdispersion เมื่อกำหนดขนาดตัวอย่างเท่ากับ 30 โดยทำซ้ำ 5,000 ครั้ง ซึ่งเมื่อพิจารณาที่ $\varphi=0$ ตัวแบบที่ถูกต้องคือ ตัวแบบการถดถอยปัวซง การตีความหมายของการทดสอบต่าง ๆ พบว่า การทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) และการทดสอบสกอร์ (Score) มีแนวโน้มที่จะให้ค่า $\alpha=0.05$ เมื่อ $\varphi=0$ นั่นคือ การทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) และการทดสอบสกอร์ (Score) เหมาะสมเมื่อ $\varphi=0$ ส่วนการทดสอบ $\mathbf{Z_o}$ (Z_o) การทดสอบวาลด์ที (Wald-t) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบ Q^* (Q*) ให้ผลแย่ในสถานการณ์นี้เนื่องจาก การทดสอบ $\mathbf{Z_o}$ (Z_o) การทดสอบวาลด์ที (Wald-t) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบ Q^* (Q*) ให้กำลังการทดสอบที่ได้จากการทดลองไม่ใกล้เคียงกับ $\alpha=0.05$ ดังนั้น ในกรณี Underdispersion เมื่อกำหนดขนาดตัวอย่างเท่ากับ 30 จะศึกษากำลังการทดสอบที่ได้จากการทดลองของการทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) และการทดสอบสกอร์ (Score) เท่านั้น (แสดงดังภาพที่ 11)



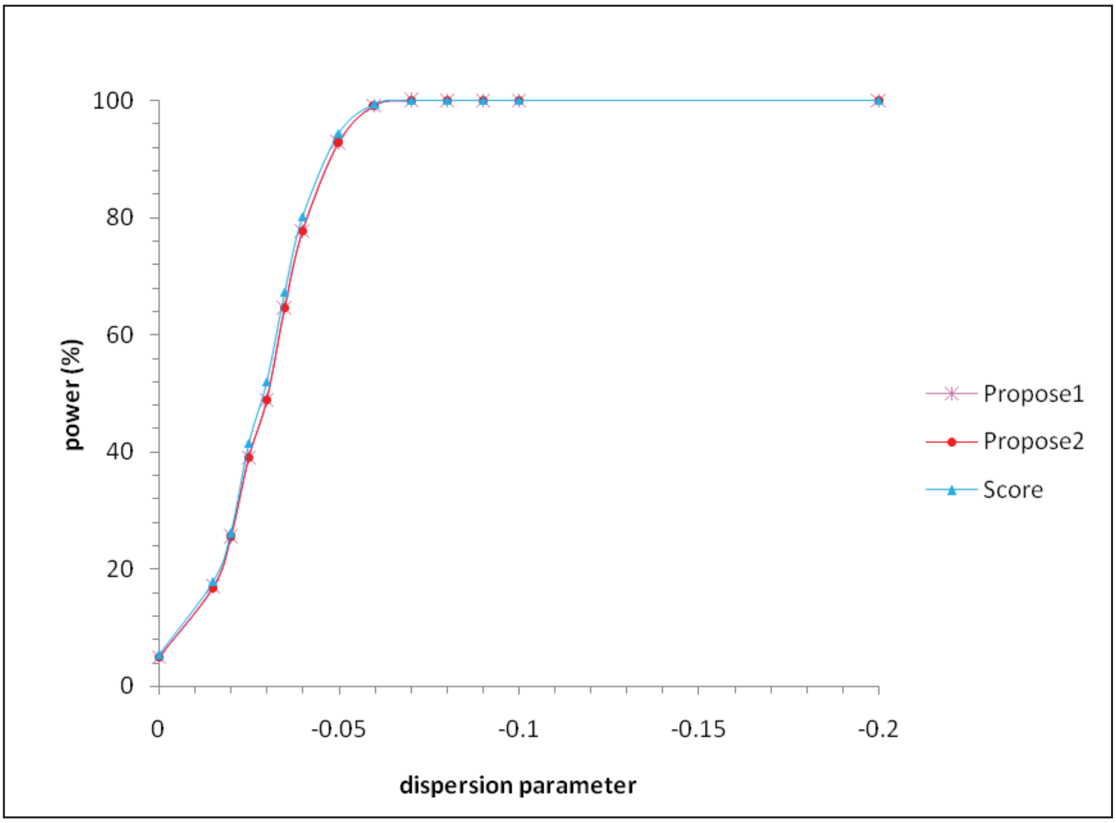
ภาพที่ 11 เส้นโค้งของกำลังการทดสอบของการทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) และการทดสอบสกอร์ (Score) ในกรณี Underdispersion เมื่อกำหนดขนาดตัวอย่างเท่ากับ 30 โดยทำซ้ำ 5,000 ครั้ง

ภาพที่ 11 แสดงเส้นโค้งของกำลังการทดสอบที่ได้จากการทดลองของการทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) และการทดสอบสกอร์(Score) ในกรณี Underdispersion เมื่อกำหนดขนาดตัวอย่างเท่ากับ 30 โดยทำซ้ำ 5,000 ครั้ง เมื่อพิจารณาภาพที่ 11 พบว่า การทดสอบสกอร์ (Score) ให้กำลังการทดสอบที่ได้จากการทดลองมากกว่าการทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) และการทดสอบ $Z_{\bar{Y}}$ (Propose2) และเมื่อพิจารณาอิทธิพลของขนาดตัวอย่างที่มีผลต่อกำลังการทดสอบที่ได้จากการทดลองพบว่า เมื่อตัวอย่างมีขนาดเล็กคือ $n = 30$ กำลังการทดสอบที่ได้จากการทดลองจะเพิ่มขึ้นอย่างช้า ๆ ในช่วงแรกและเมื่อ $\varphi < -0.02$ กำลังการทดสอบจะเพิ่มขึ้นอย่างรวดเร็วและเข้าใกล้ 1.0 เมื่อ $\varphi = -0.1$



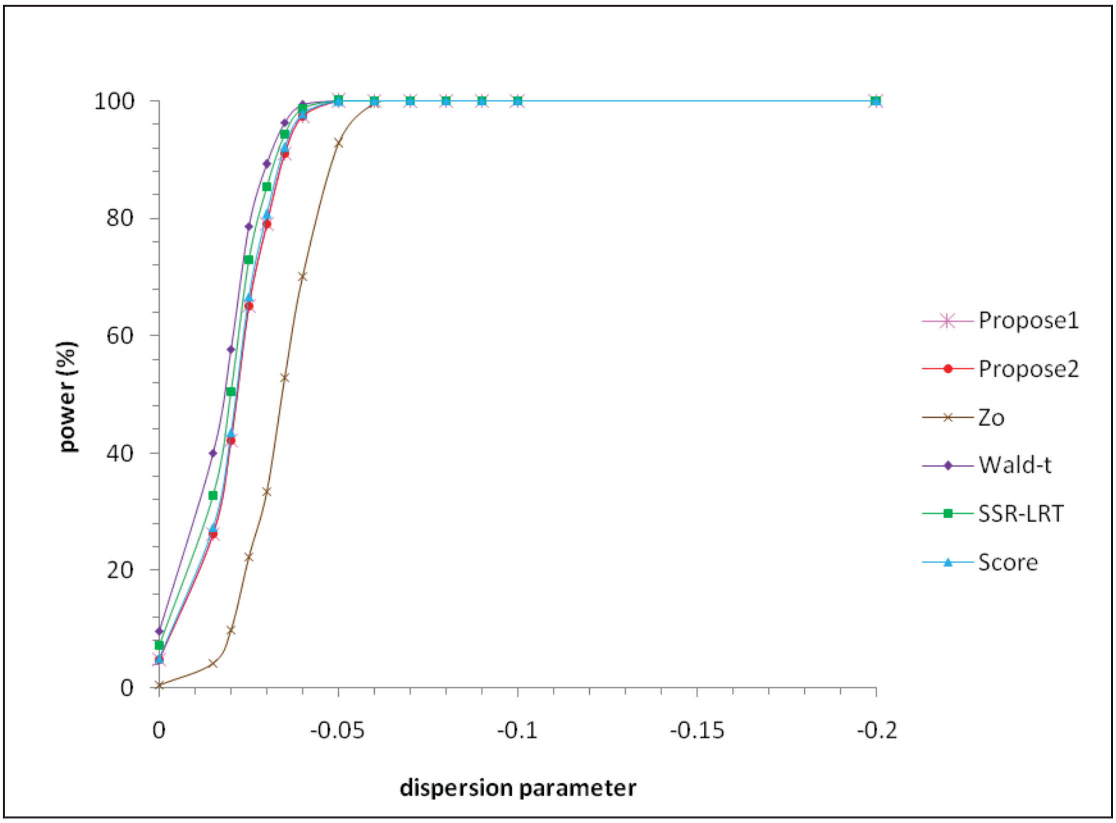
ภาพที่12 เส้นโค้งของกำลังการทดสอบที่ได้จากการทดลองของการทดสอบต่าง ๆ ในกรณี Underdispersion เมื่อกำหนดขนาดตัวอย่างเท่ากับ 50 โดยทำซ้ำ 5,000 ครั้ง

ภาพที่ 12 แสดงเส้นโค้งของกำลังการทดสอบที่ได้จากการทดลองของการทดสอบ $Z_{\bar{\mu}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) การทดสอบ $Z_{\mathbf{o}}$ (Z_o) การทดสอบวาลด์ที (Wald-t) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) การทดสอบสกอร์ (Score) และการทดสอบ Q^* (Q*) ในกรณี Underdispersion เมื่อกำหนดขนาดตัวอย่างเท่ากับ 50 โดยทำซ้ำ 5,000 ครั้ง ซึ่งเมื่อพิจารณาที่ $\varphi=0$ ตัวแบบที่ถูกต้องคือ ตัวแบบการถดถอยปัวซง การตีความหมายของการทดสอบต่าง ๆ พบว่า การทดสอบ $Z_{\bar{\mu}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) และการทดสอบสกอร์ (Score) มีแนวโน้มที่จะให้ค่า $\alpha=0.05$ เมื่อ $\varphi=0$ นั่นคือ การทดสอบ $Z_{\bar{\mu}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) และการทดสอบสกอร์ (Score) เหมาะสมเมื่อ $\varphi=0$ ส่วนการทดสอบ $Z_{\mathbf{o}}$ (Z_o) การทดสอบวาลด์ที (Wald-t) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบ Q^* (Q*) ให้ผลแย่ในสถานการณ์นี้เนื่องจากการทดสอบ $Z_{\mathbf{o}}$ (Z_o) การทดสอบวาลด์ที (Wald-t) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบ Q^* (Q*) ให้กำลังการทดสอบที่ได้จากการทดลองไม่ใกล้เคียงกับ $\alpha=0.05$ ดังนั้นในกรณี Underdispersion เมื่อกำหนดขนาดตัวอย่างเท่ากับ 50 จะศึกษากำลังการทดสอบที่ได้จากการทดลองของการทดสอบ $Z_{\bar{\mu}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) และการทดสอบสกอร์ (Score) เท่านั้น (แสดงดังภาพที่ 13)



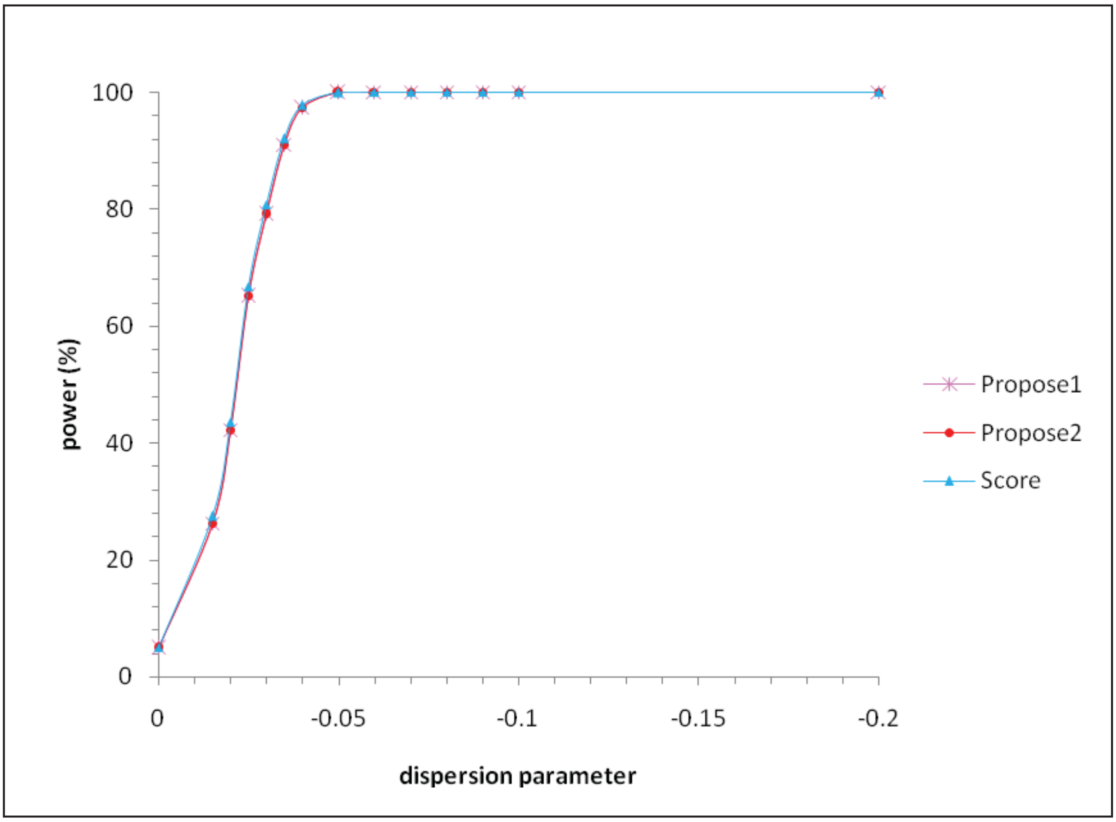
ภาพที่ 13 เส้นโค้งของกำลังการทดสอบของการทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\overline{Y}}$ (Propose2) และการทดสอบสกอร์ (Score) ในกรณี Underdispersion เมื่อกำหนดขนาดตัวอย่างเท่ากับ 50 โดยทำซ้ำ 5,000 ครั้ง

ภาพที่ 13 แสดงเส้นโค้งของกำลังการทดสอบที่ได้จากการทดลองของการทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\overline{Y}}$ (Propose2) และการทดสอบสกอร์ (Score) ในกรณี Underdispersion เมื่อกำหนดขนาดตัวอย่างเท่ากับ 50 โดยทำซ้ำ 5,000 ครั้ง เมื่อพิจารณาภาพที่ 13 พบว่า การทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) และ การทดสอบ $Z_{\overline{Y}}$ (Propose2) ให้กำลังการทดสอบที่ได้จากการทดลองใกล้เคียงกับการทดสอบสกอร์ (Score) และเมื่อพิจารณาอิทธิพลของขนาดตัวอย่างที่มีผลต่อกำลังการทดสอบที่ได้จากการทดลองพบว่า เมื่อ $n = 50$ กำลังการทดสอบที่ได้จากการทดลองจะเพิ่มขึ้นอย่างรวดเร็วและเข้าใกล้ 1.0 เมื่อ $\varphi = -0.08$



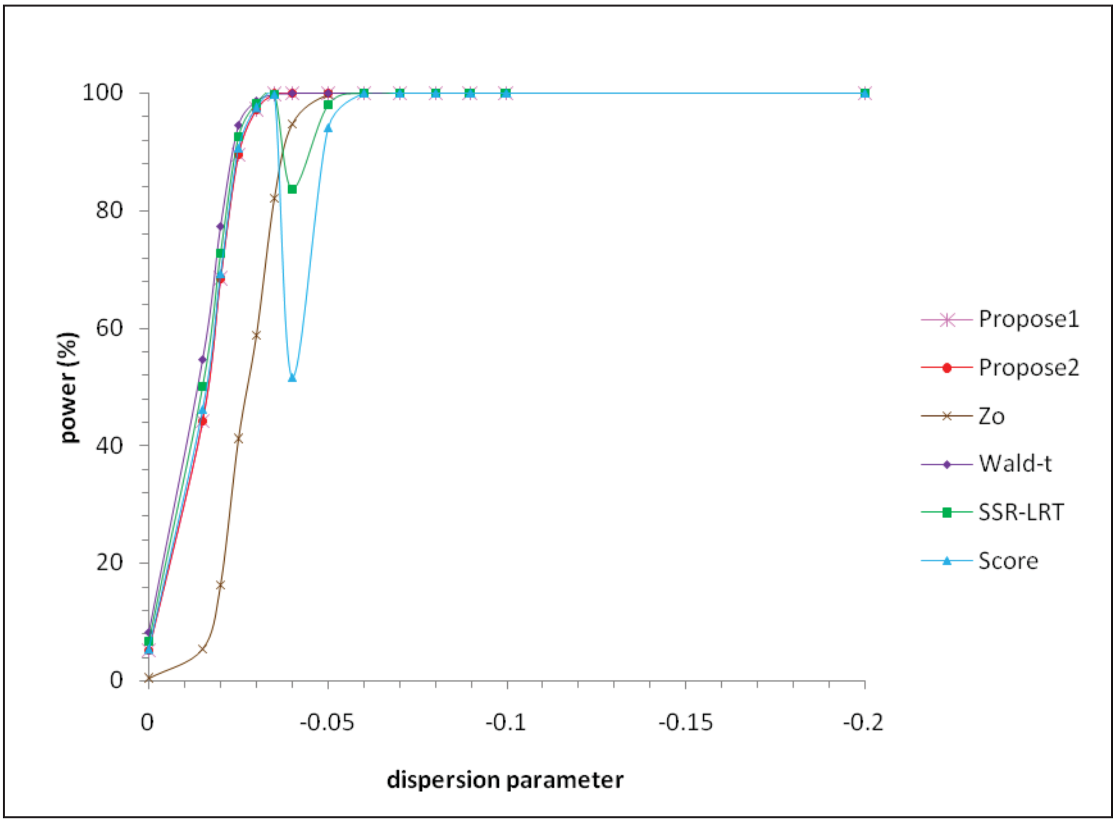
ภาพที่ 14 เส้นโค้งของกำลังการทดสอบที่ได้จากการทดลองของการทดสอบต่าง ๆ ในกรณี Underdispersion เมื่อกำหนดขนาดตัวอย่างเท่ากับ 100 โดยทำซ้ำ 5,000 ครั้ง

ภาพที่ 14 แสดงเส้นโค้งของกำลังการทดสอบที่ได้จากการทดลองของการทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) การทดสอบ $Z_{\mathbf{o}}$ (Z_o) การทดสอบวาลด์ที (Wald-t) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบสกอร์ (Score) ในกรณี Underdispersion เมื่อกำหนดขนาดตัวอย่างเท่ากับ 100 โดยทำซ้ำ 5,000 ครั้ง ซึ่งเมื่อพิจารณาที่ $\varphi=0$ ตัวแบบที่ถูกต้องคือตัวแบบการถดถอยปัวซง การตีความหมายของการทดสอบต่าง ๆ พบว่า การทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) และการทดสอบสกอร์ (Score) มีแนวโน้มที่จะให้ค่า $\alpha=0.05$ เมื่อ $\varphi=0$ นั่นคือ การทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) และการทดสอบสกอร์ (Score) เหมาะสมเมื่อ $\varphi=0$ ส่วนการทดสอบ $Z_{\mathbf{o}}$ (Z_o) การทดสอบวาลด์ที (Wald-t) และการทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) ให้ผลแย่ในสถานการณ์นี้เนื่องจากการทดสอบ $Z_{\mathbf{o}}$ (Z_o) การทดสอบวาลด์ที (Wald-t) และการทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) ให้กำลังการทดสอบที่ได้จากการทดลองไม่ใกล้เคียงกับ $\alpha=0.05$ ดังนั้น ในกรณี Underdispersion เมื่อกำหนดขนาดตัวอย่างเท่ากับ 100 จะศึกษากำลังการทดสอบที่ได้จากการทดลองของการทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) และการทดสอบสกอร์ (Score) เท่านั้น (แสดงดังภาพที่ 15)



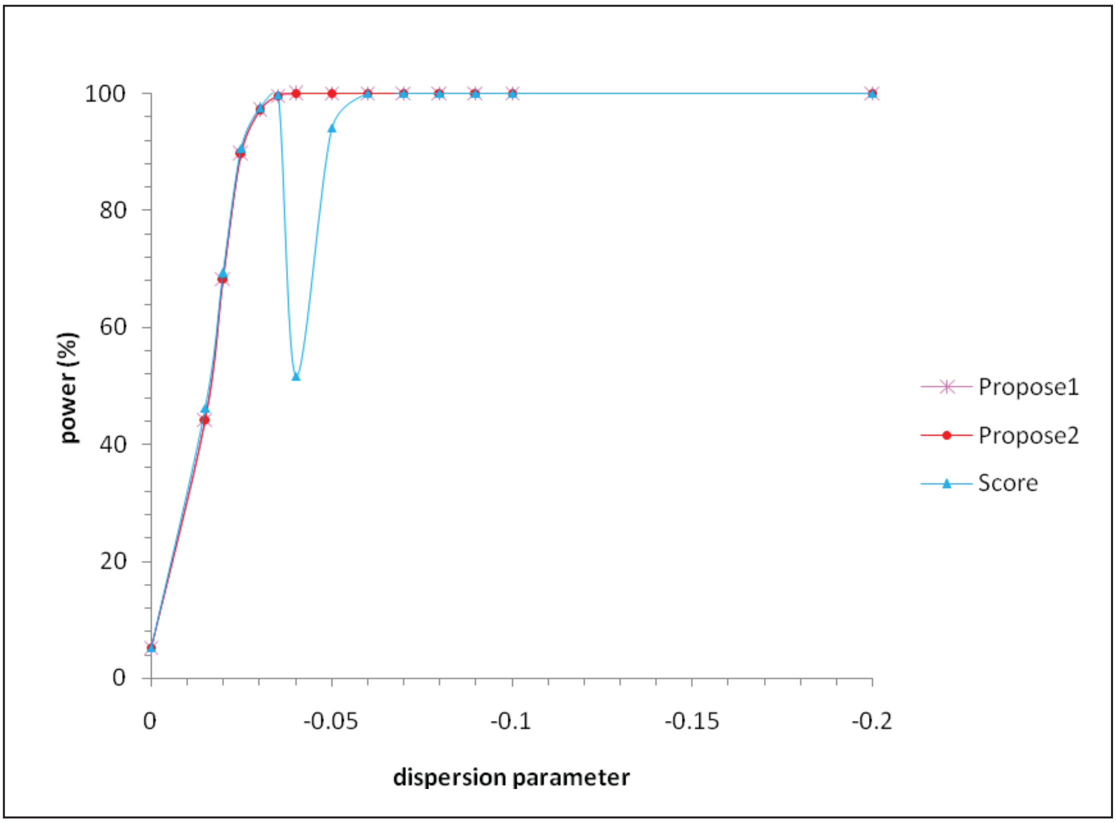
ภาพที่ 15 เส้นโค้งของกำลังการทดสอบของการทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) และการทดสอบสกอร์ (Score) ในกรณี Underdispersion เมื่อกำหนดขนาดตัวอย่างเท่ากับ 100 โดยทำซ้ำ 5,000 ครั้ง

ภาพที่ 15 แสดงเส้นโค้งของกำลังการทดสอบที่ได้จากการทดลองของการทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) และการทดสอบสกอร์ (Score) ในกรณี Underdispersion เมื่อกำหนดขนาดตัวอย่างเท่ากับ 100 โดยทำซ้ำ 5,000 ครั้ง เมื่อพิจารณาภาพที่ 15 พบว่า การทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) และ การทดสอบ $Z_{\bar{Y}}$ (Propose2) ให้กำลังการทดสอบที่ได้จากการทดลองใกล้เคียงกับการทดสอบสกอร์ (Score) และเมื่อพิจารณาอิทธิพลของขนาดตัวอย่างที่มีผลต่อกำลังการทดสอบที่ได้จากการทดลองพบว่า เมื่อตัวอย่างมีขนาดใหญ่คือ $n = 100$ กำลังการทดสอบจะเพิ่มขึ้นอย่างรวดเร็วและเข้าใกล้ 1.0 เมื่อ $\varphi = -0.06$



ภาพที่ 16 เส้นโค้งของกำลังการทดสอบที่ได้จากการทดลองของการทดสอบต่าง ๆ ในกรณี Underdispersion เมื่อกำหนดขนาดตัวอย่างเท่ากับ 200 โดยทำซ้ำ 5,000 ครั้ง

ภาพที่ 16 แสดงเส้นโค้งของกำลังการทดสอบที่ได้จากการทดลองของการทดสอบ $Z_{\bar{\mu}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) การทดสอบ Z_o (Z_o) การทดสอบวาลด์ที (Wald-t) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบสกอร์ (Score) ในกรณี Underdispersion เมื่อกำหนดขนาดตัวอย่างเท่ากับ 200 โดยทำซ้ำ 5,000 ครั้ง ซึ่งเมื่อพิจารณาที่ $\varphi=0$ ตัวแบบที่ถูกต้องคือตัวแบบการถดถอยปัวซง การตีความหมายของการทดสอบต่าง ๆ พบว่า การทดสอบ $Z_{\bar{\mu}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) และการทดสอบสกอร์ (Score) มีแนวโน้มที่จะให้ค่า $\alpha=0.05$ เมื่อ $\varphi=0$ นั่นคือ การทดสอบ $Z_{\bar{\mu}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) และการทดสอบสกอร์ (Score) เหมาะสมเมื่อ $\varphi=0$ ส่วนการทดสอบ Z_o (Z_o) การทดสอบวาลด์ที (Wald-t) และการทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) ให้ผลแย่ในสถานการณ์นี้เนื่องจากการทดสอบ Z_o (Z_o) การทดสอบวาลด์ที (Wald-t) และการทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) ให้กำลังการทดสอบที่ได้จากการทดลองไม่ใกล้เคียงกับ $\alpha=0.05$ ดังนั้นในกรณี Underdispersion เมื่อกำหนดขนาดตัวอย่างเท่ากับ 100 จะศึกษากำลังการทดสอบที่ได้จากการทดลองของการทดสอบ $Z_{\bar{\mu}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) และการทดสอบสกอร์ (Score) เท่านั้น (แสดงดังภาพที่ 17)

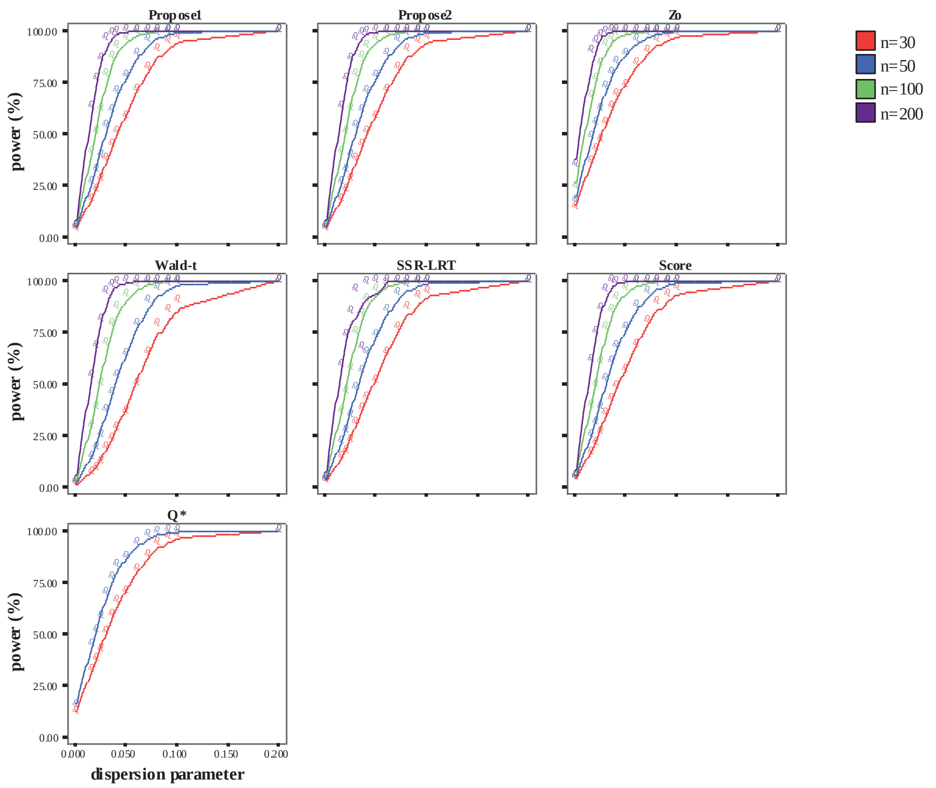


ภาพที่ 17 เส้นโค้งของกำลังการทดสอบของการทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) และการทดสอบสกอร์ (Score) ในกรณี Underdispersion เมื่อกำหนดขนาดตัวอย่างเท่ากับ 200 โดยทำซ้ำ 5,000 ครั้ง

ภาพที่ 17 แสดงเส้นโค้งของกำลังการทดสอบที่ได้จากการทดลองของการทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) และการทดสอบสกอร์ (Score) ในกรณี Underdispersion เมื่อกำหนดขนาดตัวอย่างเท่ากับ 200 โดยทำซ้ำ 5,000 ครั้ง เมื่อพิจารณาภาพที่ 17 พบว่า การทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) และ การทดสอบ $Z_{\bar{Y}}$ (Propose2) ให้กำลังการทดสอบที่ได้จากการทดลองใกล้เคียงกับการทดสอบสกอร์ (Score) นอกจากนั้นยังพบว่า เมื่อตัวอย่างมีขนาดใหญ่ ไม่ควรเลือกใช้การทดสอบสกอร์ (Score) เนื่องจากการทดสอบสกอร์ (Score) ให้กำลังการทดสอบที่ได้จากการทดลองต่ำแบบผิดปกติ เมื่อค่าพารามิเตอร์การกระจายเท่ากับ -0.04 และ -0.05 และเมื่อพิจารณาอิทธิพลของขนาดตัวอย่างที่มีผลต่อกำลังการทดสอบที่ได้จากการทดลองพบว่า เมื่อตัวอย่างมีขนาดใหญ่คือ $n = 200$ กำลังการทดสอบจะเพิ่มขึ้นอย่างรวดเร็วและเข้าใกล้ 1.0 เมื่อ $\varphi = -0.06$

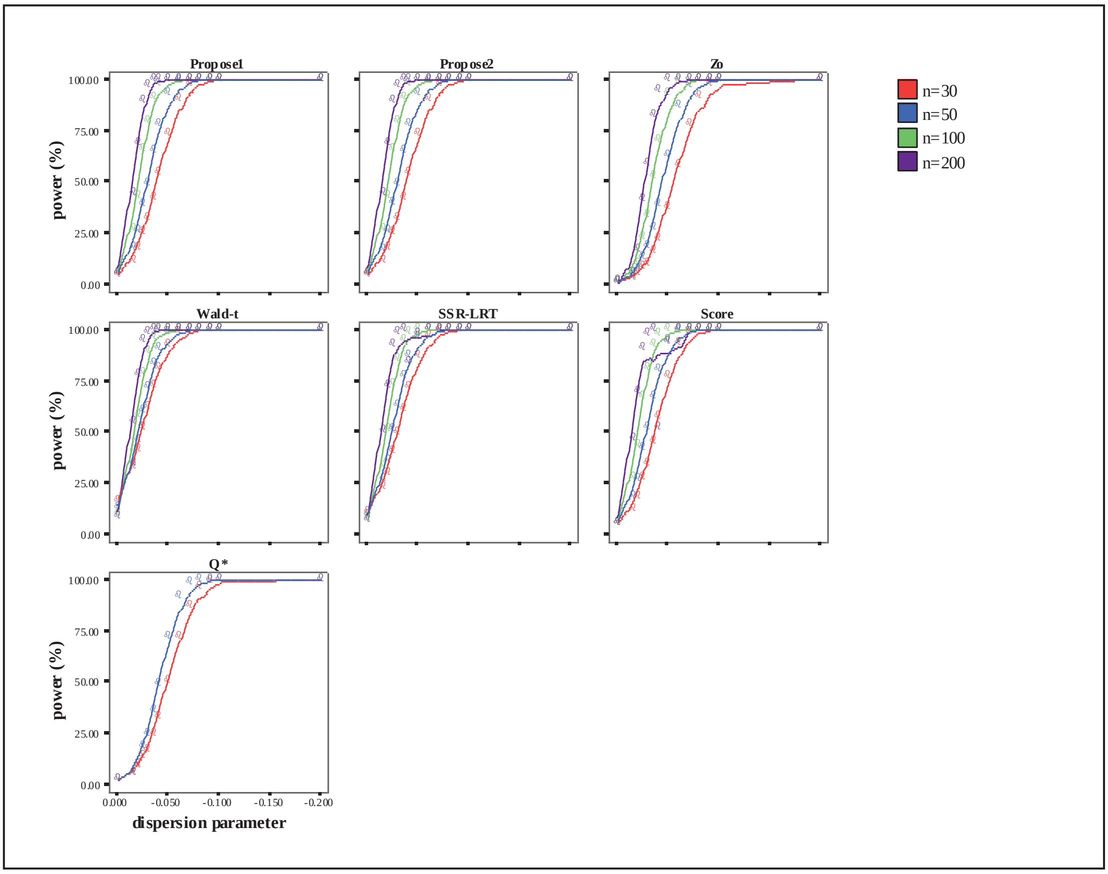
ในกรณี Underdispersion นี้สามารถวิเคราะห์ผลลัพธ์ต่าง ๆ ได้ดังนี้ ผลลัพธ์ในตารางที่ 3 และภาพที่ 10-17 พบว่า เมื่อ $\varphi=0$ ตัวแบบที่ถูกต้องคือ ตัวแบบการถดถอยปัวซง การตีความหมายของการทดสอบต่าง ๆ พบว่า การทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) และการทดสอบสกอร์ (Score) เหมาะสม ส่วนการทดสอบ $Z_{\mathbf{o}}$ (Z_o) การทดสอบวาลด์ที (Wald-t) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) และการทดสอบ Q^* (Q*) ให้ผลแย่ในสถานการณ์นี้ เมื่อตัวอย่างมีขนาดใหญ่ การทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) และการทดสอบสกอร์ (Score) ยังมีแนวโน้มที่จะให้ค่าประมาณของ α ได้อย่างใกล้เคียงกับค่า $\alpha=0.05$ เมื่อ $\varphi=0$ ดังนั้น การทดสอบที่สามารถนำมาใช้ในการทดสอบ Underdispersion ได้แก่ การทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) และการทดสอบสกอร์ (Score) แต่เมื่อพิจารณากำลังการทดสอบที่ได้จากการทดลองของการทดสอบทั้งสามพบว่า กำลังการทดสอบที่ได้จากการทดลองของการทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) และการทดสอบ $Z_{\bar{Y}}$ (Propose2) ที่ผู้วิจัยเสนอให้กำลังการทดสอบที่ได้จากการทดลองใกล้เคียงกันและการทดสอบทั้งสองยังให้กำลังการทดสอบที่ได้จากการทดลองใกล้เคียงกับการทดสอบสกอร์ (Score) เมื่อตัวอย่างมีขนาดใหญ่ นอกจากนั้นยังพบว่า เมื่อตัวอย่างมีขนาดใหญ่ ไม่ควรเลือกใช้การทดสอบสกอร์ (Score) เนื่องจากการทดสอบสกอร์ (Score) ให้กำลังการทดสอบที่ได้จากการทดลองต่ำแบบผิดปกติ เมื่อค่าพารามิเตอร์การกระจายเท่ากับ -0.04 และ -0.05 แต่เมื่อตัวอย่างมีขนาดเล็ก การทดสอบสกอร์ (Score) จะให้กำลังการทดสอบที่ได้จากการทดลองมากกว่าการทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) และการทดสอบ $Z_{\bar{Y}}$ (Propose2) และเมื่อพิจารณาอิทธิพลของขนาดตัวอย่างที่มีผลต่อกำลังการทดสอบที่ได้จากการทดลอง พบว่า เมื่อตัวอย่างมีขนาดใหญ่คือ $n=100$ และ $n=200$ กำลังการทดสอบจะเพิ่มขึ้นอย่างรวดเร็วและเข้าใกล้ 1.0 เมื่อ $\varphi=-0.06$ ทั้งสองขนาดตัวอย่าง และเมื่อ $n=50$ กำลังการทดสอบจะเข้าใกล้ 1.0 เมื่อ $\varphi=-0.08$ และเมื่อตัวอย่างมีขนาดเล็กคือ $n=30$ กำลังการทดสอบจะเพิ่มขึ้นอย่างช้า ๆ ในช่วงแรกและเมื่อ $\varphi<-0.02$ กำลังการทดสอบจะเพิ่มขึ้นอย่างรวดเร็วและเข้าใกล้ 1.0 เมื่อ $\varphi=-0.1$

3. อิทธิพลของขนาดตัวอย่างที่มีผลต่อกำลังการทดสอบที่ได้จากการทดลองกรณี Overdispersion และกรณี Underdispersion



ภาพที่ 18 อิทธิพลของขนาดตัวอย่างที่มีผลต่อกำลังการทดสอบที่ได้จากการทดลอง ภายใต้กรณี Overdispersion ของการทดสอบต่าง ๆ

ภาพที่ 18 แสดงอิทธิพลของขนาดตัวอย่างที่มีผลต่อกำลังการทดสอบที่ได้จากการทดลอง ภายใต้กรณี Overdispersion ของการทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) การทดสอบ Z_{o} (Z_o) การทดสอบวาลด์ที (Wald-t) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) การทดสอบสกอร์ (Score) และการทดสอบ Q^* (Q*) จากภาพพบว่า ในกรณี Overdispersion เมื่อตัวอย่างมีขนาดใหญ่ขึ้นกำลังการทดสอบที่ได้จากการทดลองจะเพิ่มขึ้นด้วย



ภาพที่ 19 อิทธิพลของขนาดตัวอย่างที่มีผลต่อกำลังการทดสอบที่ได้จากการทดลอง ภายใต้กรณี Underdispersion ของการทดสอบต่าง ๆ

ภาพที่ 19 แสดงอิทธิพลของขนาดตัวอย่างที่มีผลต่อกำลังการทดสอบที่ได้จากการทดลอง ภายใต้กรณี Overdispersion ของการทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1) การทดสอบ $Z_{\bar{Y}}$ (Propose2) การทดสอบ $Z_{\mathbf{o}}$ (Z_o) การทดสอบวาลด์ที (Wald-t) การทดสอบอัตราส่วนภาวะน่าจะเป็น (SSR-LRT) การทดสอบสกอร์ (Score) และการทดสอบ Q^* (Q*) จากภาพพบว่า ในกรณี Underdispersion เมื่อตัวอย่างมีขนาดใหญ่ขึ้นกำลังการทดสอบที่ได้จากการทดลองจะเพิ่มขึ้นด้วย

จากภาพที่ 18 และภาพที่ 19 ทำให้สรุปได้ว่า ขนาดตัวอย่างมีอิทธิพลต่อกำลังการทดสอบที่ได้จากการทดลองทั้งกรณี Overdispersion และกรณี Underdispersion คือ ไม่ว่าจะเกิด Overdispersion และ Underdispersion กำลังการทดสอบที่ได้จากการทดลองของการทดสอบต่าง ๆ จะเพิ่มขึ้นเมื่อตัวอย่างมีขนาดใหญ่ขึ้น

บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

ในงานวิจัยนี้ผู้วิจัยเสนอการทดสอบ $Z_{\bar{\hat{\mu}}}$ และการทดสอบ $Z_{\bar{Y}}$ เพื่อใช้ในการทดสอบ Overdispersion และ Underdispersion ในตัวแบบการถดถอยปัวซงและทำการศึกษาการทดสอบต่าง ๆ ดังนี้ การทดสอบ Z_0 การทดสอบวาลด์ที การทดสอบอัตราส่วนภาวะน่าจะเป็น การทดสอบสกอร์ และการทดสอบ Q^* โดยทำการเปรียบเทียบกำลังการทดสอบที่ได้จากการทดลองของการทดสอบต่าง ๆ ซึ่งข้อมูลที่ใช้ในงานวิจัยเป็นข้อมูลที่ได้จากการจำลองแบบ ซึ่งจำลองแบบข้อมูลจากตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 จำนวน 5,000 ครั้ง ภายใต้เงื่อนไขของขนาดตัวอย่างและค่าพารามิเตอร์การกระจาย นอกจากนั้นยังศึกษาอิทธิพลของขนาดตัวอย่างที่มีผลต่อกำลังการทดสอบที่ได้จากการทดลองภายใต้กรณี Overdispersion และกรณี Underdispersion อีกด้วย จากผลการวิจัย สามารถสรุปผลและอภิปรายผลการวิจัย ดังนี้

1. สรุปและอภิปรายผลการวิจัย

จากผลการวิจัย สามารถสรุปและอภิปรายผลการวิจัย ได้ดังนี้

1.1 การเปรียบเทียบกำลังการทดสอบของการทดสอบ $Z_{\bar{\hat{\mu}}}$ การทดสอบ $Z_{\bar{Y}}$ การทดสอบ Z_0 การทดสอบวาลด์ที การทดสอบอัตราส่วนภาวะน่าจะเป็น การทดสอบสกอร์ และการทดสอบ Q^* แบ่งเป็น 2 กรณี ดังนี้

1.1.1 กรณี Overdispersion

เมื่อ $\varphi=0$ ตัวแบบที่ถูกต้องคือ ตัวแบบการถดถอยปัวซง การตีความหมายของการทดสอบต่าง ๆ พบว่า การทดสอบ $Z_{\bar{\hat{\mu}}}$ การทดสอบ $Z_{\bar{Y}}$ การทดสอบอัตราส่วนภาวะน่าจะเป็น และการทดสอบสกอร์ เหมาะสม ส่วนการทดสอบ Z_0 การทดสอบวาลด์ที และการทดสอบ Q^* ให้ผลแย่ในสถานการณ์นี้ นอกจากนั้น เมื่อตัวอย่างมีขนาดใหญ่ การทดสอบ $Z_{\bar{\hat{\mu}}}$ การทดสอบ $Z_{\bar{Y}}$ การทดสอบอัตราส่วนภาวะน่าจะเป็น และการทดสอบสกอร์ ยังมีแนวโน้มที่จะให้ค่าประมาณได้อย่างใกล้เคียงกับค่า $\alpha=0.05$ มากกว่าวิธีการทดสอบอื่น ๆ เมื่อ $\varphi=0$ ดังนั้น ในกรณี Overdispersion ควรเลือกใช้การทดสอบ $Z_{\bar{\hat{\mu}}}$ การทดสอบ $Z_{\bar{Y}}$ การทดสอบอัตราส่วนภาวะน่าจะเป็น หรือการทดสอบสกอร์ ซึ่ง

เมื่อพิจารณากำลังการทดสอบที่ได้จากการทดลองพบว่า กำลังการทดสอบที่ได้จากการทดลองของการทดสอบ $Z_{\bar{\hat{\mu}}}$ และการทดสอบ $Z_{\bar{Y}}$ ที่ผู้วิจัยเสนอให้กำลังการทดสอบที่ได้จากการทดลองใกล้เคียงกัน และการทดสอบทั้งสองยังให้กำลังการทดสอบที่ได้จากการทดลองมากกว่าการทดสอบอัตราส่วนภาวะน่าจะเป็นและการทดสอบสกอร์ ในทุกขนาดตัวอย่าง นอกจากนั้นยังพบว่า เมื่อตัวอย่างมีขนาดใหญ่ ไม่ควรเลือกใช้การทดสอบอัตราส่วนภาวะน่าจะเป็น เนื่องจากการทดสอบอัตราส่วนภาวะน่าจะเป็นให้กำลังการทดสอบที่ได้จากการทดลองต่ำแบบผิดปกติในบางกรณี

ดังนั้น ในกรณี Overdispersion การทดสอบ $Z_{\bar{\hat{\mu}}}$ และการทดสอบ $Z_{\bar{Y}}$ ที่ผู้วิจัยเสนอให้ผลการทดสอบที่เด่นกว่าการทดสอบอื่น ๆ เมื่อพิจารณาในเทอมของกำลังการทดสอบและการแจกแจงค่อนข้างสมมาตรตลอดจนการทดสอบ $Z_{\bar{Y}}$ มีรูปแบบที่ง่ายกว่าการทดสอบ $Z_{\bar{\hat{\mu}}}$ ดังนั้น การทดสอบ $Z_{\bar{Y}}$ จึงเหมาะสมมากกว่าเมื่อนำไปใช้ในการทดสอบ Overdispersion ในตัวแบบการถดถอยปัวซงโดยเทียบกับตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2

1.1.2 กรณี Underdispersion

เมื่อ $\varphi=0$ ตัวแบบที่ถูกต้องคือ ตัวแบบการถดถอยปัวซง การตีความหมายของการทดสอบต่าง ๆ พบว่า การทดสอบ $Z_{\bar{\hat{\mu}}}$ การทดสอบ $Z_{\bar{Y}}$ และการทดสอบสกอร์ เหมาะสม ส่วนการทดสอบ $Z_{\mathbf{0}}$ การทดสอบวาลด์ที การทดสอบอัตราส่วนภาวะน่าจะเป็น และการทดสอบ Q^{*} ให้ผลแย่ในสถานการณ์นี้ นอกจากนี้ เมื่อตัวอย่างมีขนาดใหญ่ การทดสอบ $Z_{\bar{\hat{\mu}}}$ การทดสอบ $Z_{\bar{Y}}$ และการทดสอบสกอร์ ยังมีแนวโน้มที่จะให้ค่า $\alpha=0.05$ เมื่อ $\varphi=0$ ดังนั้น ในกรณี Underdispersion ควรเลือกใช้การทดสอบ $Z_{\bar{\hat{\mu}}}$ การทดสอบ $Z_{\bar{Y}}$ หรือการทดสอบสกอร์ แต่เมื่อพิจารณากำลังการทดสอบที่ได้จากการทดลองพบว่า เมื่อตัวอย่างมีขนาดเล็ก การทดสอบสกอร์จะให้กำลังการทดสอบที่ได้จากการทดลองมากกว่าการทดสอบ $Z_{\bar{\hat{\mu}}}$ และการทดสอบ $Z_{\bar{Y}}$ แต่เมื่อตัวอย่างมีขนาดใหญ่ การทดสอบ $Z_{\bar{\hat{\mu}}}$ และการทดสอบ $Z_{\bar{Y}}$ ที่ผู้วิจัยเสนอจะให้กำลังการทดสอบที่ได้จากการทดลองใกล้เคียงกับการทดสอบสกอร์และการแจกแจงค่อนข้างสมมาตร นอกจากนั้นยังพบว่า การทดสอบสกอร์ให้กำลังการทดสอบที่ได้จากการทดลองต่ำแบบผิดปกติ ในบางกรณี ดังนั้น เมื่อตัวอย่างมีขนาดใหญ่ควรเลือกใช้การทดสอบ $Z_{\bar{\hat{\mu}}}$ หรือการทดสอบ $Z_{\bar{Y}}$ ที่ผู้วิจัยเสนอขึ้น

ดังนั้น ในกรณี Underdispersion เมื่อตัวอย่างมีขนาดเล็ก ควรเลือกใช้การทดสอบสกอร์ในการเปรียบเทียบตัวแบบการถดถอยปัวซงเทียบกับตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 แต่เมื่อตัวอย่างมีขนาดใหญ่ควรเลือกใช้การทดสอบ $Z_{\bar{\hat{\mu}}}$ หรือการทดสอบ $Z_{\bar{Y}}$ ที่ผู้วิจัยเสนอขึ้น แต่การทดสอบ $Z_{\bar{Y}}$ มีรูปแบบที่ง่ายกว่าการทดสอบ $Z_{\bar{\hat{\mu}}}$ ดังนั้น เมื่อตัวอย่างมีขนาดใหญ่การทดสอบ $Z_{\bar{Y}}$ จึงเหมาะสมมากกว่าเมื่อนำไปใช้ในการทดสอบ Underdispersion ในตัวแบบการถดถอยปัวซงโดยเทียบกับตัวแบบปัวซงนัยทั่วไปแบบที่ 2

1.2 การศึกษาอิทธิพลของขนาดตัวอย่างที่มีผลต่อกำลังการทดสอบที่ได้จากการทดลองภายใต้กรณี Overdispersion และกรณี Underdispersion

ในการศึกษาอิทธิพลของขนาดตัวอย่างที่มีผลต่อกำลังการทดสอบที่ได้จากการทดลองภายใต้กรณี Overdispersion และกรณี Underdispersion พบว่า เมื่อตัวอย่างมีขนาดเล็ก กำลังการทดสอบจะเพิ่มขึ้นอย่างช้า ๆ ในช่วงแรกและต่อมากำลังการทดสอบจะเพิ่มขึ้นอย่างรวดเร็วและเข้าใกล้ 1.0 แต่เมื่อตัวอย่างมีขนาดใหญ่ กำลังการทดสอบจะเพิ่มขึ้นอย่างรวดเร็วและเข้าใกล้ 1.0 นอกจากนั้นยังพบว่า เมื่อเกิด Overdispersion หรือ Underdispersion กำลังการทดสอบที่ได้จากการทดลองของการทดสอบต่าง ๆ จะเพิ่มขึ้นเมื่อตัวอย่างมีขนาดใหญ่ขึ้น

2. ข้อเสนอแนะของงานวิจัย

จากผลการวิจัยพบว่า เมื่อเกิด Overdispersion ควรเลือกใช้การทดสอบ $Z_{\bar{\hat{\mu}}}$ หรือการทดสอบ $Z_{\bar{Y}}$ ที่ผู้วิจัยเสนอขึ้น เนื่องจากการทดสอบ $Z_{\bar{\hat{\mu}}}$ และการทดสอบ $Z_{\bar{Y}}$ ให้กำลังการทดสอบที่ได้จากการทดลองมากกว่าการทดสอบอื่น ๆ และการแจกแจงค่อนข้างสมมาตรตลอดจนการทดสอบ $Z_{\bar{Y}}$ มีรูปแบบที่ง่ายกว่าการทดสอบ $Z_{\bar{\hat{\mu}}}$ ดังนั้น การทดสอบ $Z_{\bar{Y}}$ จึงเหมาะสมมากกว่าเมื่อนำไปใช้ในการทดสอบ Overdispersion ในตัวแบบการถดถอยปัวซงโดยเทียบกับตัวแบบปัวซงนัยทั่วไปแบบที่ 2 และในกรณี Underdispersion เมื่อตัวอย่าง มีขนาดใหญ่ ควรเลือกใช้การทดสอบ $Z_{\bar{\hat{\mu}}}$ หรือการทดสอบ $Z_{\bar{Y}}$ ที่ผู้วิจัยเสนอขึ้น เนื่องจากการทดสอบ $Z_{\bar{\hat{\mu}}}$ และการทดสอบ $Z_{\bar{Y}}$ ให้กำลังการทดสอบที่ได้จากการทดลองใกล้เคียงกับการทดสอบสกอร์และการแจกแจงค่อนข้างสมมาตรตลอดจนการทดสอบ $Z_{\bar{Y}}$ มีรูปแบบที่ง่ายกว่าการทดสอบ $Z_{\bar{\hat{\mu}}}$ ดังนั้นการทดสอบ $Z_{\bar{Y}}$ จึงเหมาะสมมากกว่าเมื่อนำไปใช้ในการทดสอบ Underdispersion ในตัวแบบการถดถอยปัวซงโดยเทียบกับตัวแบบปัวซงนัยทั่วไปแบบที่ 2 แต่เมื่อตัวอย่างมีขนาดเล็ก ควรเลือกใช้การทดสอบสกอร์ เนื่องจากการทดสอบสกอร์ยังให้กำลังการทดสอบที่ได้จากการทดลองมากกว่าการทดสอบอื่น ๆ

3. ข้อเสนอแนะของงานวิจัยครั้งต่อไป

เนื่องจากในงานวิจัยนี้เสนอการทดสอบที่เหมาะที่ใช้สำหรับเปรียบเทียบตัวแบบการถดถอยปัวซงเทียบกับตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 และทำการศึกษาการทดสอบอื่น ๆ สำหรับเปรียบเทียบตัวแบบปัวซงเทียบกับตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 ซึ่งในงานวิจัยนี้ทำการเปรียบเทียบกำลังการทดสอบที่ได้จากการทดลองของการทดสอบต่าง ๆ ในงานวิจัยครั้งต่อไปอาจศึกษาการทดสอบใหม่ ๆ และอาจตรวจสอบสมบัติของการทดสอบที่เหมาะสำหรับเปรียบเทียบตัวแบบการถดถอยปัวซงเทียบกับตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 นอกจากนี้อาจศึกษาเกณฑ์อื่น ๆ ที่อาจใช้ในการตรวจสอบความเหมาะสมของการทดสอบที่นอกเหนือจากการใช้กำลังการทดสอบที่ได้จากการทดลองและอาจศึกษาในกรณีที่มีตัวแปรอธิบายจำนวนหลายตัวด้วย

บทที่ 6

การนำการทดสอบที่ผู้วิจัยเสนอไปประยุกต์กับข้อมูลจริง

ในบทนี้ผู้วิจัยจะแสดงการนำไปใช้ของการทดสอบ $Z_{\overline{Y}}$ (Propose2) เพื่อใช้ในการเปรียบเทียบตัวแบบการถดถอยปัวซงเทียบกับตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 โดยข้อมูลที่นำมาศึกษาเป็นข้อมูลจริงเกี่ยวกับจำนวนปูเพศผู้ ซึ่งเก็บตัวอย่างจากปูเพศเมีย จำนวน 173 ตัว (Agresti 2002 : 127, อ้างถึงใน Data courtesy of Jane Brockmann 1996) โดยตัวแปรตอบสนองคือ จำนวนปูเพศผู้ และตัวแปรอธิบายคือ น้ำหนักของปูเพศเมีย ซึ่งข้อมูลต่าง ๆ แสดงดังตารางที่ 4

ตารางที่ 4 จำนวนปูเพศผู้จำแนกตามคุณลักษณะของปูเพศเมีย

y	x	y	x	y	x	y	x	y	x	y	x	y	x	y	x	y	x
8	3.05	4	2.25	0	1.60	5	2.00	6	1.95	0	1.85	6	2.50	1	2.00	0	2.45
4	2.60	0	2.90	2	1.85	0	2.75	7	3.05	6	2.80	0	1.80	4	1.95	0	2.25
0	2.15	3	2.25	3	2.28	3	2.45	6	2.25	5	3.30	6	2.50	3	2.00	1	2.87
0	1.85	0	1.70	0	2.20	10	3.20	3	2.92	4	2.10	2	1.65	0	2.60	1	2.00
1	3.00	0	3.20	4	3.28	7	2.80	4	3.73	5	2.90	4	1.47	0	2.00	2	1.90
3	2.30	1	1.97	0	2.35	0	1.90	4	2.85	15	3.00	0	1.80	0	2.65	0	2.10
0	1.30	1	1.60	0	1.55	0	1.20	0	1.90	0	2.25	0	2.20	3	3.10	12	3.23
0	2.10	1	2.90	0	2.10	0	1.65	0	1.80	5	2.15	6	2.63	9	3.25	6	1.80
8	2.00	4	2.30	0	2.15	0	3.05	8	3.05	0	2.40	0	2.00	3	3.00	3	2.90
6	3.15	2	2.10	14	2.30	5	3.85	0	1.80	1	1.65	4	3.02	6	2.70	2	2.02
5	2.80	0	1.40	0	2.20	0	1.55	0	2.62	0	1.60	0	2.30	3	2.70	4	2.90
4	2.80	2	3.28	1	1.60	1	2.20	9	2.30	5	2.10	4	1.95	0	2.55	5	3.10
3	3.60	0	2.30	3	3.15	1	2.55	0	1.90	4	2.55	4	3.50	1	2.80	7	5.20
4	1.60	6	2.30	4	3.20	1	2.40	0	2.65	0	2.75	0	2.15	0	1.30	0	2.40
3	2.30	10	2.25	5	2.70	3	3.25	8	2.95	0	2.55	2	2.17	0	1.80	10	1.90
5	2.05	5	2.40	0	1.90	2	3.33	5	2.70	1	2.80	0	2.63	3	2.20	0	2.00
8	3.05	3	3.32	6	2.50	5	2.40	2	2.60	1	3.00	0	2.10	0	2.25	0	3.20
3	2.40	8	2.10	6	2.60	0	2.22	5	2.70	4	2.55	0	1.95	0	2.30	4	2.35
6	2.25	9	3.00	5	2.10	3	3.20	0	2.60	1	3.10	11	3.05	0	1.90	7	2.75
3	2.75	0	2.00														

กำหนดให้ ตัวแปรตอบสนอง (y) คือ จำนวนปูเพศผู้ และตัวแปรอธิบาย (x) แทน น้ำหนักของปูเพศเมีย (หน่วยเป็นกิโลกรัม)

การประมาณค่าสัมประสิทธิ์สำหรับตัวแบบการถดถอยปัวซงแสดงดังตารางที่ 5

ตารางที่ 5 ค่าประมาณสัมประสิทธิ์สำหรับตัวแบบการถดถอยปัวซง

พารามิเตอร์	องศาอิสระ	$\hat{\beta}$	$S.E.(\hat{\beta})$	ค่าสถิติ t	P - value
b0	173	-0.4282	0.1789	-2.39	0.0178
b1	173	0.5892	0.06500	9.06	<.0001

จากตารางสามารถสร้างตัวแบบการถดถอยปัวซง ได้ดังนี้

$$\hat{\mu} = -0.4282 + 0.5892x$$

โดยที่ $\hat{\mu}$ แทน ค่าเฉลี่ยของจำนวนปูเพศผู้ และ x แทน น้ำหนักของปูเพศเมีย (หน่วยเป็นกิโลกรัม)

การประมาณค่าสัมประสิทธิ์สำหรับตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 แสดงดังตารางที่ 6

ตารางที่ 6 ค่าประมาณสัมประสิทธิ์สำหรับตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2

พารามิเตอร์	องศาอิสระ	$\hat{\beta}$	$S.E.(\hat{\beta})$	ค่าสถิติ t	P - value
b0	173	-1.0481	0.4856	-2.16	0.0323
b1	173	0.8379	0.1980	4.23	<.0001
$\hat{\varphi}$	173	0.3525	0.05580	6.32	<.0001

จากตารางสามารถสร้างตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 ได้ดังนี้

$$\hat{\mu} = -1.0481 + 0.8379x$$

โดยที่ $\hat{\mu}$ แทน ค่าเฉลี่ยของจำนวนปูเพศผู้ และ x แทน น้ำหนักของปูเพศเมีย (หน่วยเป็นกิโลกรัม)

ข้อมูลชุดนี้ประกอบด้วยค่าสังเกต 173 ค่า ซึ่งจำนวนนับที่น้อยที่สุดคือ 0 และจำนวนนับที่มากที่สุดคือ 15 โดยมีค่าเฉลี่ยเท่ากับ 2.91908 และความแปรปรวนเท่ากับ 9.91204 ดัชนีการกระจาย (อัตราส่วนระหว่างความแปรปรวนและค่าเฉลี่ย) คือ 3.39560 ดังนั้นข้อมูลชุดนี้เกิด Overdispersion เพราะฉะนั้นควรทำการเปรียบเทียบตัวแบบการถดถอยปัวซงเทียบกับตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 ดังนั้น สมมติฐานที่ใช้ในการทดสอบคือ $H_0 : \varphi = 0$ เทียบกับ $H_1 : \varphi > 0$ การทดสอบที่จะใช้ในการทดสอบสมมติฐานคือ การทดสอบ $Z_{\bar{Y}}$ (Propose2) ซึ่งมีรูปแบบคือ

$$Z_{\bar{Y}} = \sqrt{\frac{n-1}{2}}\left(\left(1+\hat{\varphi}\bar{Y}\right)^2 - 1\right)$$

เมื่อประมวลผลด้วยโปรแกรม SAS (แสดงรายละเอียดการเขียนโปรแกรมในภาคผนวก ค) พบว่า ค่าประมาณพารามิเตอร์การกระจายของตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 คือ $\hat{\varphi} = 0.3525$ ดังนั้น เราสามารถคำนวณค่าสถิติทดสอบ $Z_{\overline{Y}}$ (Propose2) ได้ดังนี้

$$Z_{\overline{Y}} = \sqrt{\frac{173-1}{2}} \left\{ \left[1+\left(0.3525 * 2.91908\right)\right]^2 - 1 \right\} = 28.9069$$

เนื่องจากค่าสถิติทดสอบ $Z_{\overline{Y}}$ เท่ากับ 28.9069 มากกว่าค่าวิกฤต $Z_{\alpha=0.05} = 1.645$ ดังนั้นมีหลักฐานเพียงพอที่จะปฏิเสธสมมติฐานว่างคือ ตัวแบบการถดถอยปัวซงเหมาะสมกับข้อมูล ภายใต้ระดับนัยสำคัญ 0.05

ดังนั้น ตัวแบบที่เหมาะสำหรับข้อมูลชุดนี้คือ ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 ซึ่งมีรูปแบบคือ

$$\hat{\mu} = -1.0481 + 0.8379x$$

โดยที่ $\hat{\mu}$ แทน ค่าเฉลี่ยของจำนวนปูเพศผู้

x แทน น้ำหนักของปูเพศเมีย (หน่วยเป็นกิโลกรัม)

บรรณานุกรม

Agresti, A. Categorical data analysis. 2. NewYork : John Wiley & Sons, 2002.

Annafari, M.T. "An Empirical Analysis of the Factors Determining Multiple Subscriptions in the Swedish Mobile Phone Market." IEEE computer society (2010) : 25-32.

Böhning, D. "A note on test for Poisson overdispersion." Biometika, 81 (1994) : 418-419.

Cameron, A.C., Trivedi, P.K. "Essentials of Count Data Regression [Online]. Accessed 5 July 2011. Available from http://www.google.co.th File *Essentials of Count Data Regression* - Pennsylvania State University

Consul, P.C. and Jain, G.C. "A generalization of the Poisson distribution." Technometrics 15 (1973) : 791-799.

Dean, C.B. "Testing for overdispersion in Poisson and binomial regression models." J. Amer. Statist. Assoc, 87 (1992) : 451-457.

Dean, C., Lawless, J.F. "Test for detecting overdispersion in Poisson regression models." J. Amer. Statist. Assoc, 84 (1989) : 467-471.

Famoye, F., Wulu, J.T., Singh, Jr. and Singh, K.P. "On the Generalized Poisson Regression Model with an Application to Accident Data." Journal of Data Science, 2 (2004) : 287-295.

He, B., Xie, M., Goh, T.N. and Tsui, K.L. "On Control Charts Based on the Generalized Poisson Model." Quality Technology& Quantitative Management, 3 (2006) : 383 -400.

Heinzl, H. and Mittlböck, M. "Pseudo R-squared measures for Poisson regression models with over- or underdispersion." Computational Statistics & Data Analysis, 44 (2003) : 253-271.

Jani, P.N., Shanubhogue, A. and Muralidharan, K. A method of constructing test based on the quadratic form [Online]. Accessed 1 July 2011. Available from ftp://metron.sta.uniroma1.it/RePEc/articoli/2000-LVIII-1_2-11.pdf

Oliveira, T.D. Some elementary tests for mixtures of discrete distribution [Online]. Accessed 1 July 2011. Available from http://www.dtic.mil/cgi-bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=AD0405910

Özmen, I. and Famoye, F. "Count Regression Models with an Application to Zoological Data Containing Structural Zeros." Journal of Data Science, 5 (2007) : 491-502.

Rashwan, N.A. and Kamel, M.M. “Using Generalized Poisson Log Linear Regression Models in Analyzing Two-Way Contingency Tables.” Applied Mathematical Sciences, 5 (2011) : 213-222.

Singh, K.P., Wulu, J.T. and Bartolucci, A.A. “A Note on Generalized Poisson Regression Models.” Statistical Modeling in Health and Medical Science, Fred Ghassemi (ed.). Modeling and Simulation Society of Australia,4 (2001) : 2029-2034.

Singh, K.P., Wulu, J.T., Bartolucci, A.A. and Valappil, T. “Analyzing Frequencies of Gay Men’s Sexual Events Using A generalized Poisson regression Approach.” In : Epidemiology, Health and Medical Research, L Oxley, F Scrimgour and M McAleer (eds). Modeling and Simulation Society of Australia, 2 (1999) : 531-536.

Singh, K.P., Wulu, J.T., Jr., Bae, S., Bartolucci, A.A. and Trevino, F. “Modeling Hospital Discharge Counts Across Zip Code Area.” Statistical Modeling in Health and Biomedical Sciences, David Post (ed.). Modeling and Simulation Society of Australia, 4 (2003) : 2029-2034.

Wang, W. and Famoye, F. “Modeling household fertility decisions with generalized Poisson regression.” Journal of Population Economics, 10 (1997) : 273-283.

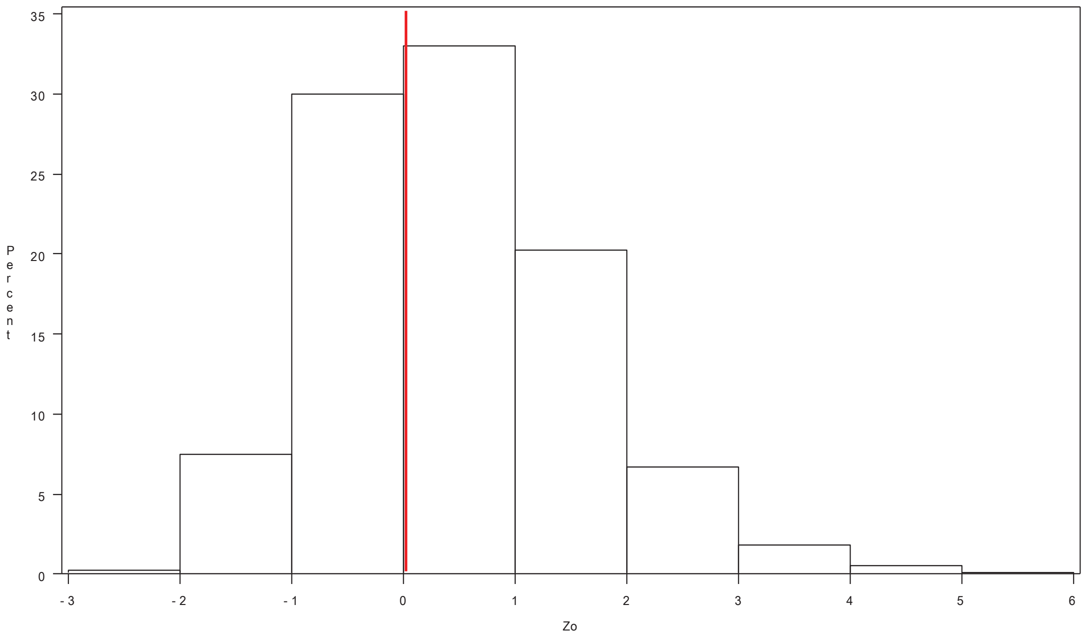
Xie, M., He, B. and Goh, T.N. “Zero-inflated Poisson model in statistical process control.” Computational Statistics & Data Analysis, 38 (2001) : 191-201.

Yang, Z., Hardin, J.W. and Addy, C.L. “A score test for overdispersion in Poisson regression based on the generalized Poisson-2 model.” Journal of Statistical Planning and Inference, 139 (2009) : 1514-1521.

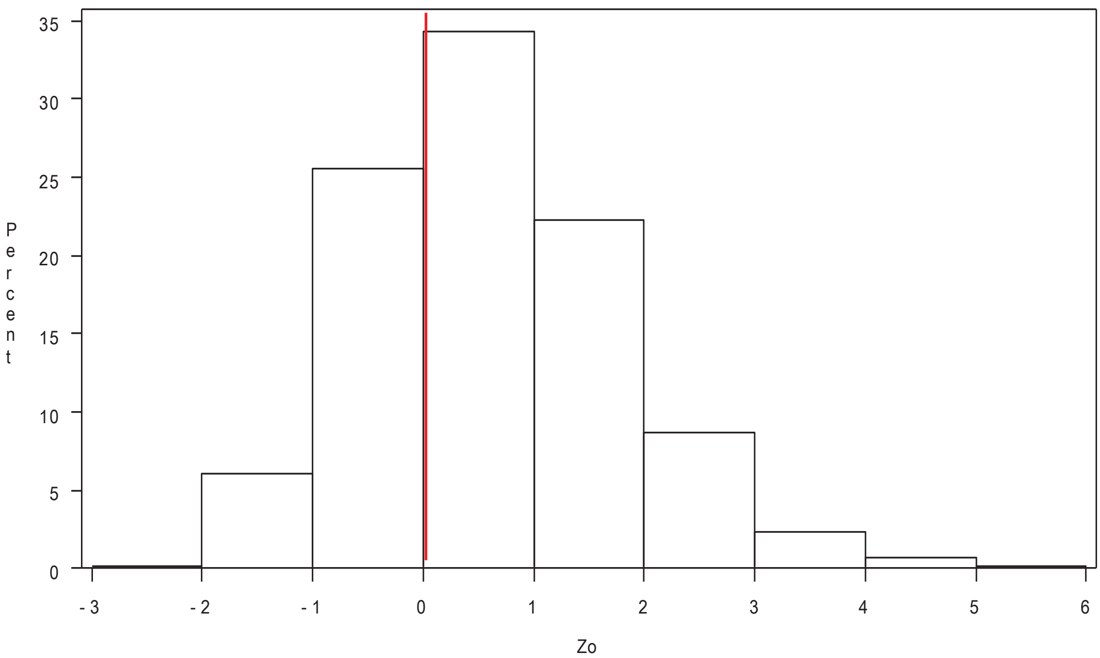
ภาคผนวก

ภาคผนวก ก

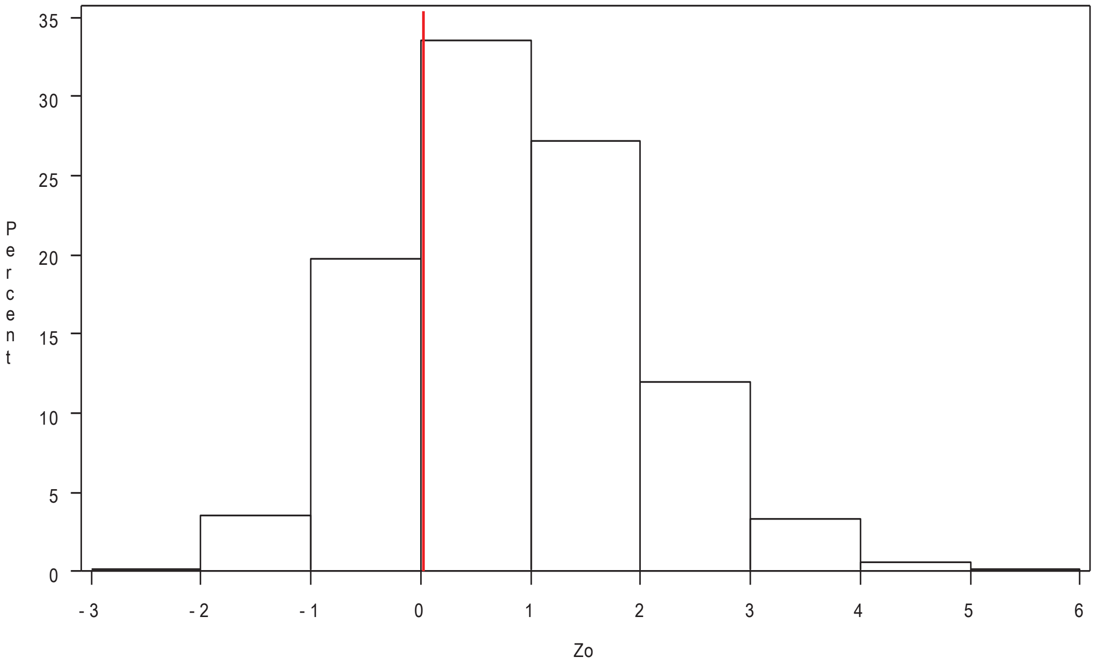
การแจกแจงของการทดสอบ Z_o



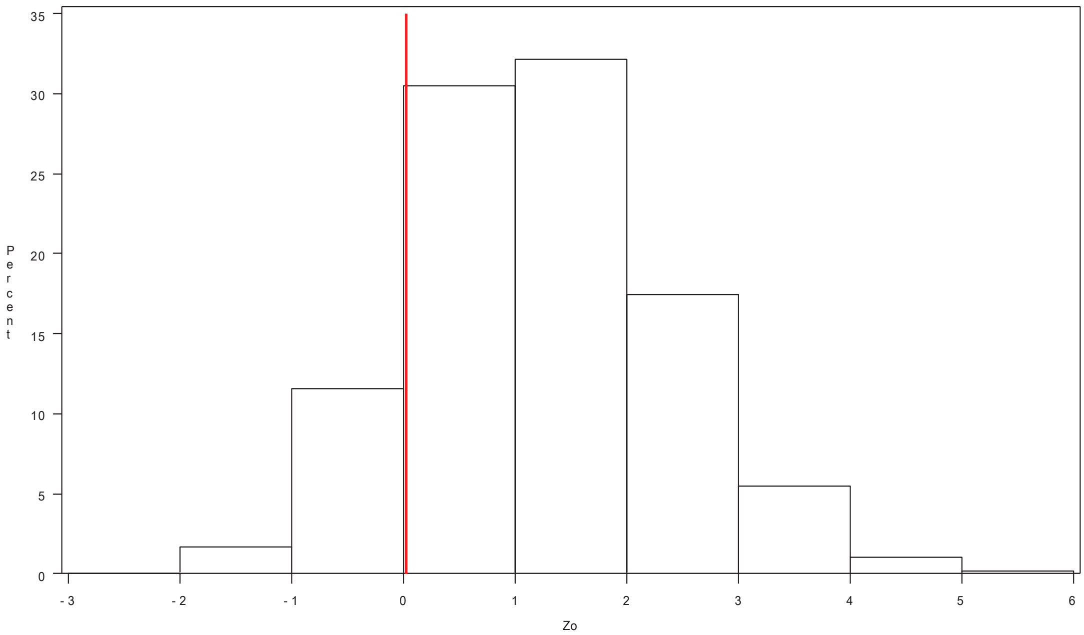
ภาพที่ 20 การแจกแจงของการทดสอบ Z_o เมื่อกำหนดค่าพารามิเตอร์การกระจาย (φ) เท่ากับ 0.0 และขนาดตัวอย่างเท่ากับ 30



ภาพที่ 21 การแจกแจงของการทดสอบ Z_o เมื่อกำหนดค่าพารามิเตอร์การกระจาย (φ) เท่ากับ 0.0 และขนาดตัวอย่างเท่ากับ 50

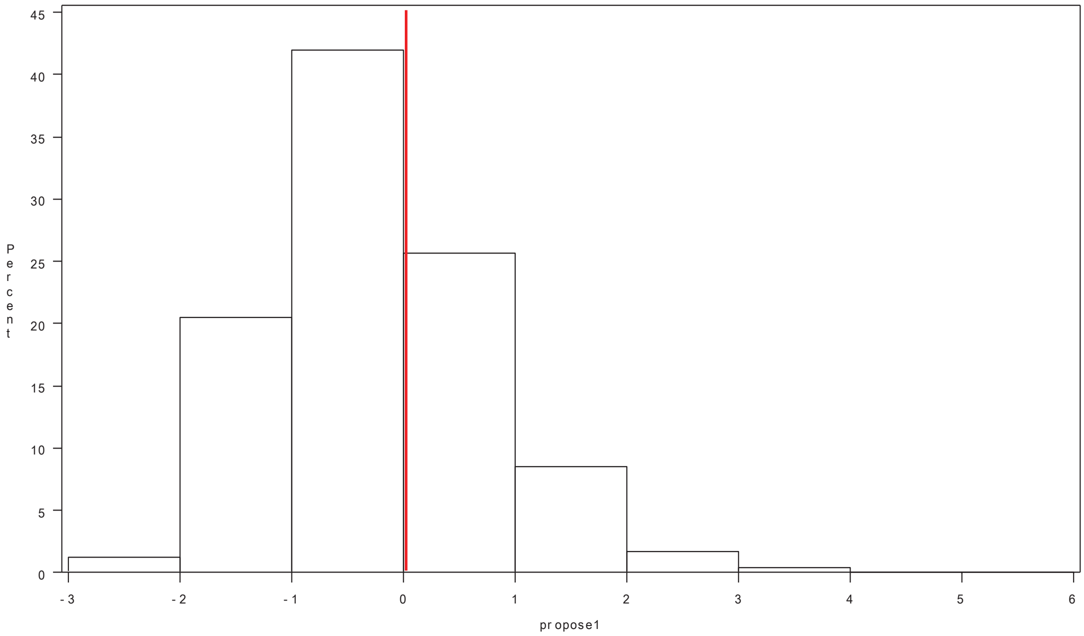


ภาพที่ 22 การแจกแจงของการทดสอบ Z_0 เมื่อกำหนดค่าพารามิเตอร์การกระจาย (φ) เท่ากับ 0.0 และขนาดตัวอย่างเท่ากับ 100

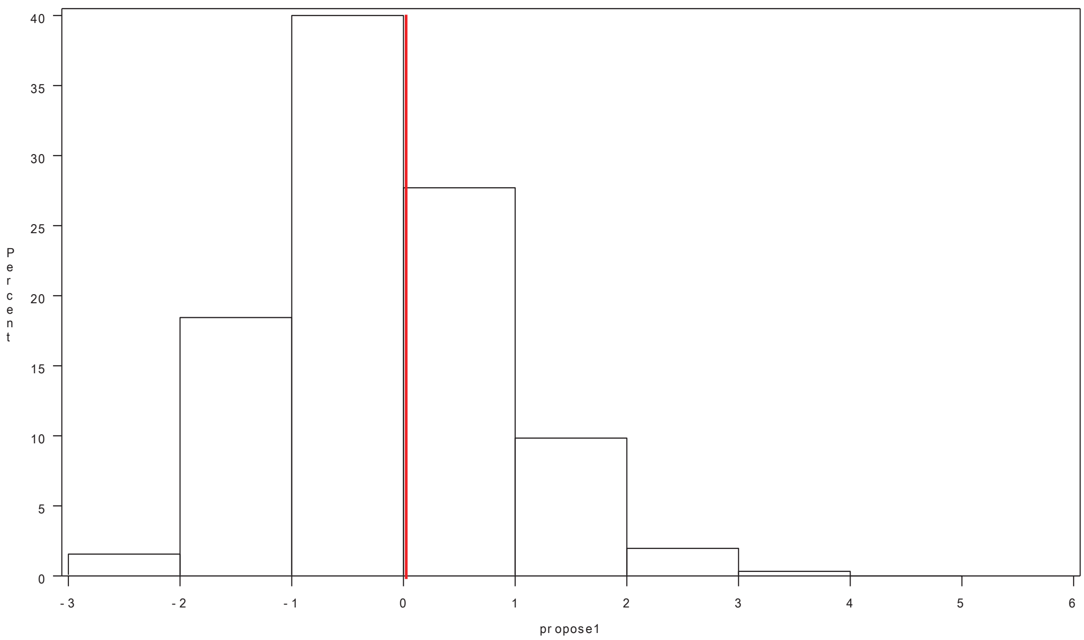


ภาพที่ 23 การแจกแจงของการทดสอบ Z_0 เมื่อกำหนดค่าพารามิเตอร์การกระจาย (φ) เท่ากับ 0.0 และขนาดตัวอย่างเท่ากับ 200

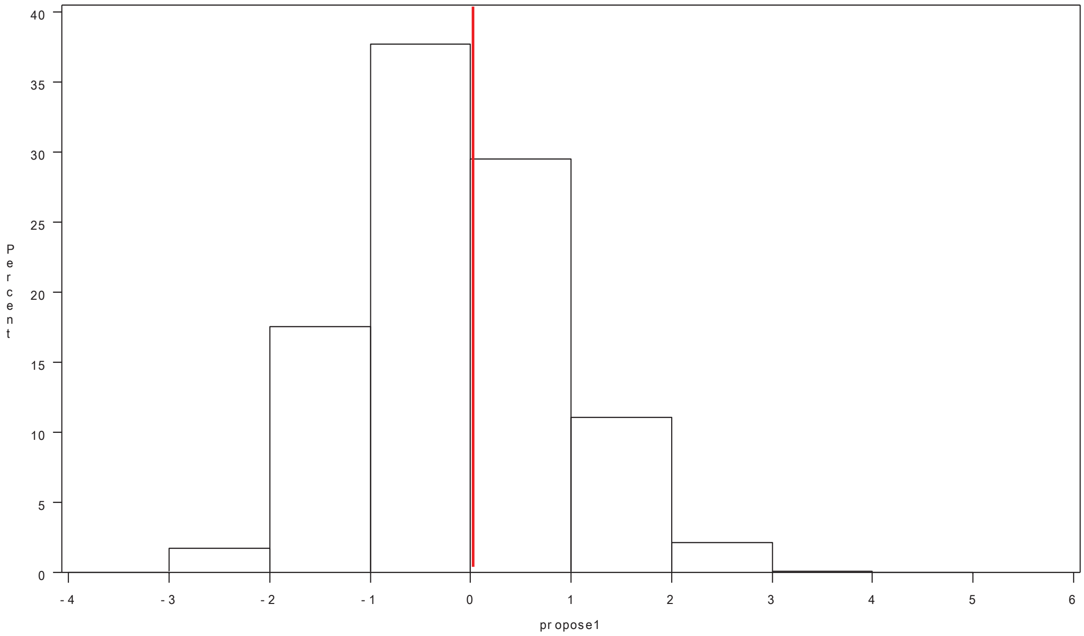
การแจกแจงของการทดสอบ $Z_{\bar{\hat{\mu}}}$ (Propose1)



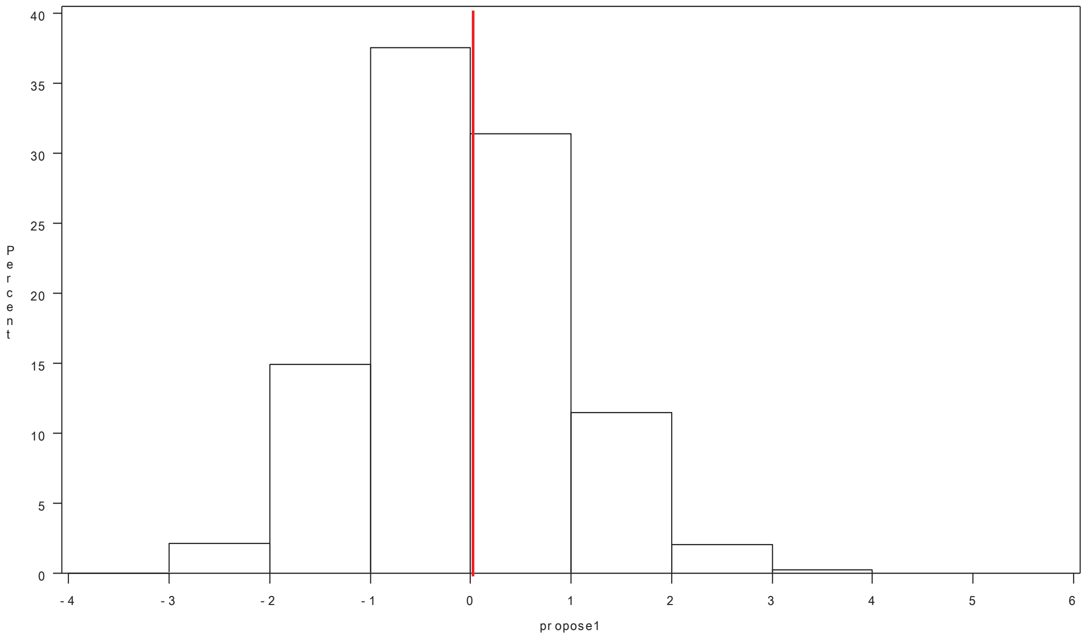
ภาพที่ 24 การแจกแจงของการทดสอบ $Z_{\bar{\hat{\mu}}}$ เมื่อกำหนดค่าพารามิเตอร์การกระจาย (φ) เท่ากับ 0.0 และขนาดตัวอย่างเท่ากับ 30



ภาพที่ 25 การแจกแจงของการทดสอบ $Z_{\bar{\hat{\mu}}}$ เมื่อกำหนดค่าพารามิเตอร์การกระจาย (φ) เท่ากับ 0.0 และขนาดตัวอย่างเท่ากับ 50

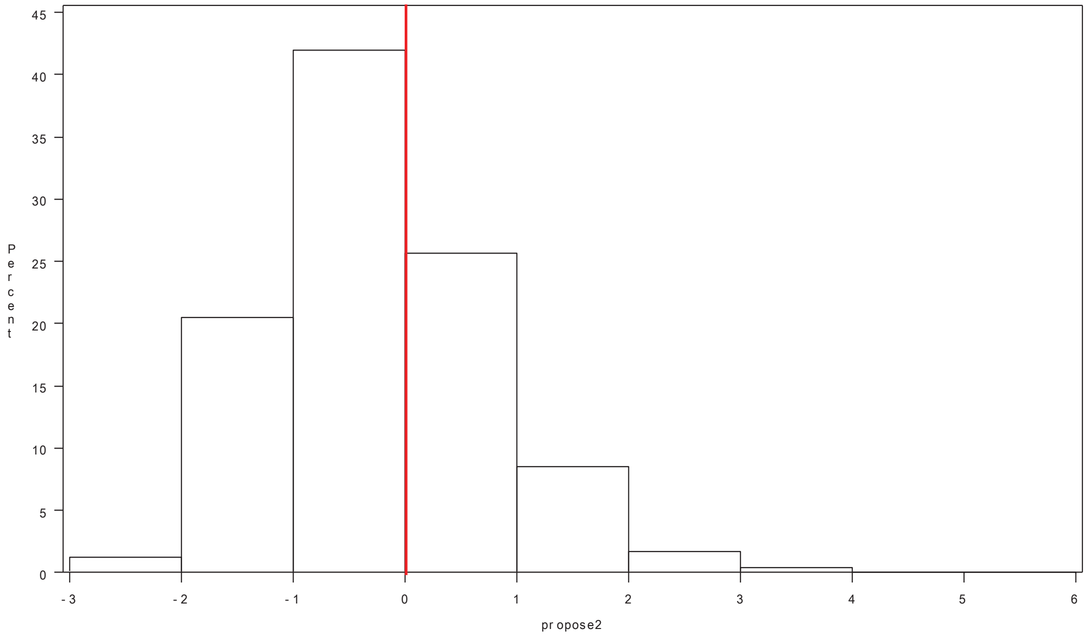


ภาพที่ 26 การแจกแจงของการทดสอบ $Z_{\bar{\hat{\mu}}}$ เมื่อกำหนดค่าพารามิเตอร์การกระจาย (φ) เท่ากับ 0.0 และขนาดตัวอย่างเท่ากับ 100

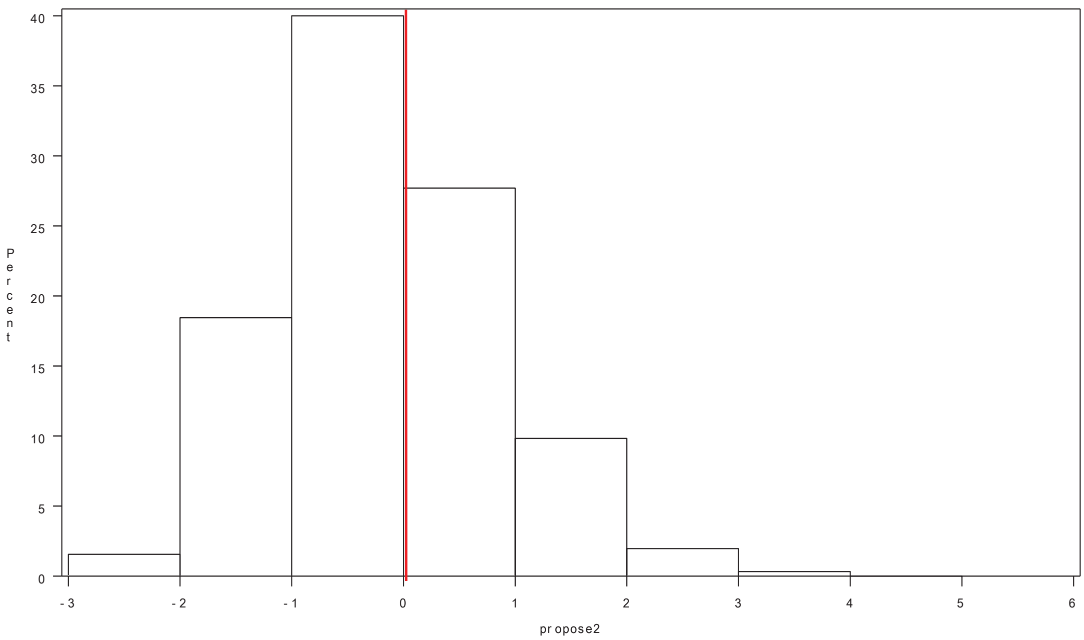


ภาพที่ 27 การแจกแจงของการทดสอบ $Z_{\bar{\hat{\mu}}}$ เมื่อกำหนดค่าพารามิเตอร์การกระจาย (φ) เท่ากับ 0.0 และขนาดตัวอย่างเท่ากับ 200

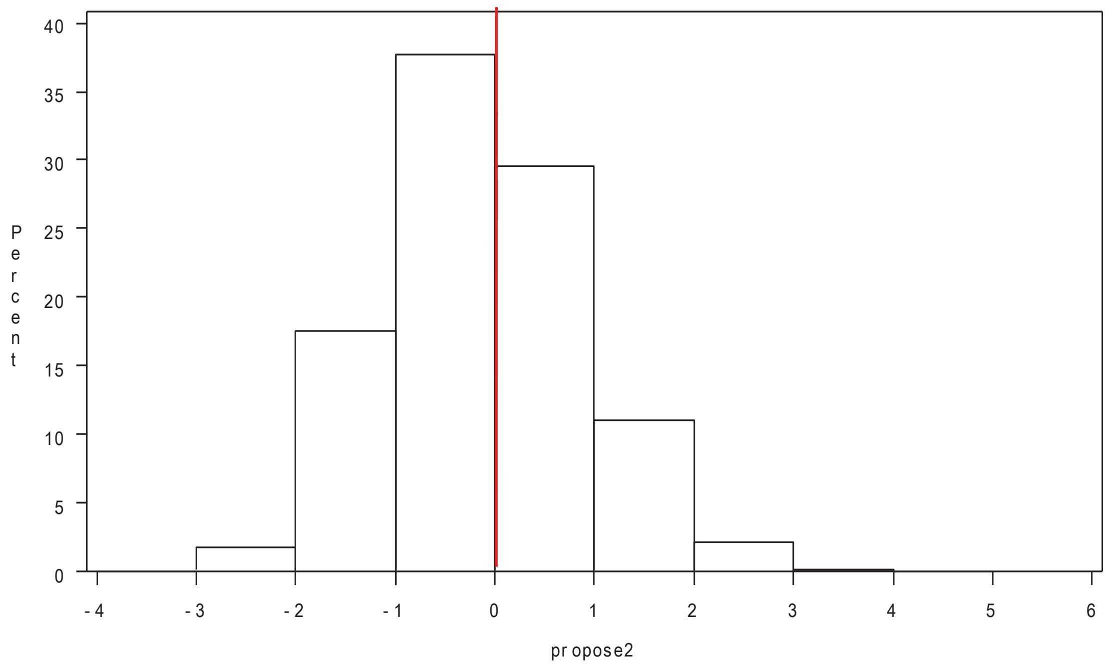
การแจกแจงของการทดสอบ $Z_{\overline{Y}}$ (Propose2)



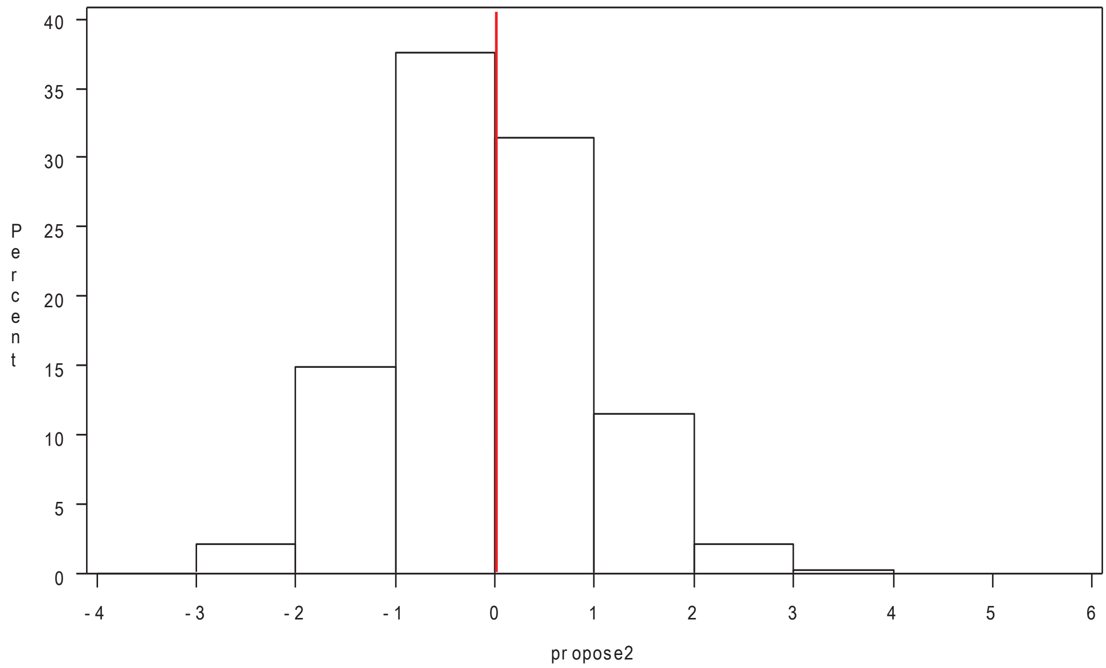
ภาพที่ 28 การแจกแจงของการทดสอบ $Z_{\overline{Y}}$ เมื่อกำหนดค่าพารามิเตอร์การกระจาย (φ) เท่ากับ 0.0 และขนาดตัวอย่างเท่ากับ 30



ภาพที่ 29 การแจกแจงของการทดสอบ $Z_{\overline{Y}}$ เมื่อกำหนดค่าพารามิเตอร์การกระจาย (φ) เท่ากับ 0.0 และขนาดตัวอย่างเท่ากับ 50



ภาพที่ 30 การแจกแจงของการทดสอบ $Z_{\overline{Y}}$ เมื่อกำหนดค่าพารามิเตอร์การกระจาย (φ) เท่ากับ 0.0 และขนาดตัวอย่างเท่ากับ 100



ภาพที่ 31 การแจกแจงของการทดสอบ $Z_{\overline{Y}}$ เมื่อกำหนดค่าพารามิเตอร์การกระจาย (φ) เท่ากับ 0.0 และขนาดตัวอย่างเท่ากับ 200

ภาคผนวก ข

โปรแกรมคอมพิวเตอร์ที่ใช้ในงานวิจัย

งานวิจัยนี้ทำการจำลองแบบข้อมูล ประมวลผลข้อมูลโดยใช้ตัวแบบการถดถอยปัวซงและตัวแบบถดถอยปัวซงนัยทั่วไปแบบที่ 2 รวมถึงคำนวณค่าสถิติทดสอบต่าง ๆ โดยใช้โปรแกรมสำเร็จรูป SAS® version 9.1 ซึ่งในที่นี้จะแสดงโปรแกรมสำหรับการจำลองแบบข้อมูลและการคำนวณค่าสถิติทดสอบต่าง ๆ ตามลำดับ โดยสัญลักษณ์ต่าง ๆ ที่ใช้ในการเขียนโปรแกรมแมคโคร ได้แก่

nrep	หมายถึง	จำนวนครั้งของการจำลองแบบข้อมูล
n	หมายถึง	ขนาดตัวอย่าง
y	หมายถึง	ค่าของตัวแปรตอบสนอง
x	หมายถึง	ค่าของตัวแปรอธิบาย
k1	หมายถึง	ค่าพารามิเตอร์การกระจาย
muS	หมายถึง	ค่าพยากรณ์ที่ได้จากสมการ $\log \mu_i = 2 - 0.5x_i$
mu_gp	หมายถึง	ค่าพยากรณ์จากตัวแบบถดถอยปัวซงนัยทั่วไปแบบที่ 2
mu_poi	หมายถึง	ค่าพยากรณ์จากตัวแบบถดถอยปัวซง
ybar	หมายถึง	ค่าเฉลี่ยของตัวแปรตอบสนอง
s2	หมายถึง	ความแปรปรวนของตัวแปรตอบสนอง

1. โปรแกรมสำหรับการจำลองแบบข้อมูล

ข้อมูลที่นำมาใช้ในการวิจัยนี้เป็นข้อมูลที่ได้จากการจำลองแบบ ซึ่งตัวแบบที่ใช้ในการจำลองแบบคือ ตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 โดยจะใช้ตัวแปรอธิบาย 1 ตัวในการจำลองแบบ ซึ่งมีรายละเอียดดังนี้ ตัวแปรตอบสนอง (Y) จำลองมาจากตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 ซึ่งมีค่าเฉลี่ยเท่ากับ μ_i และความแปรปรวนเท่ากับ $\mu_i\left(1+\varphi\mu_i\right)^2$ โดยคำนวณ μ_i จากเซตของพารามิเตอร์ $\{\beta_0, \beta_1\}$ ดังนี้ $\{2, -0.5\}$ โดยสมการที่ใช้ในการคำนวณคือ $\log \mu_i = 2 - 0.5x_i$ โดย x_i มีการแจกแจงแบบ Uniform [0,1] แบบต่อเนื่อง ภายใต้เงื่อนไขของขนาดตัวอย่างเท่ากับ 30, 50, 100, 200 และค่าพารามิเตอร์การกระจาย (φ) ดังนี้ กรณี Overdispersion กำหนดค่าพารามิเตอร์การกระจายเท่ากับ 0.0, 0.015, 0.02, 0.025, 0.03, 0.035, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.10 และ 0.2 กรณี Underdispersion กำหนดค่าพารามิเตอร์การกระจายเท่ากับ 0.0, -0.015, -0.02, -0.025, -0.03, -0.035, -0.04, -0.05, -0.06, -0.07, -0.08, -0.09 , -0.10 และ -0.2 จำนวน 5,000 ครั้ง ในแต่ละเงื่อนไข โดยโปรแกรมที่ใช้จำลองแบบข้อมูล มีรายละเอียดดังนี้

```
data sim;n=&n;nrep=&nrep;k1=&k1;seed=12345;
      do nrep=1 to nrep;
            do obs=1 to n;
                  x=ranuni(seed);
                  muS=exp(2-0.5*x);
                  lamda=muS/(1+k1*muS)**2;
                  Y_k=ranpoi(seed,lamda);
                  Y1=Y_k*(1+k1*muS)**2;
                  y=round(Y1);
            output;end;
      end;
run;
```

2. โปรแกรมสำหรับการคำนวณค่าสถิติทดสอบต่าง ๆ

2.1 กรณี Overdispersion

2.1.1 การเขียนโปรแกรมสำหรับคำนวณค่าสถิติทดสอบ $Z_{\bar{\hat{\mu}}}$

```
proc nlmixed data=sim;by nrep;
      parms b0=0 b1=0 k=0;
      eta= b0 + b1*x;
      mu=exp(eta);
      loglike=(y*log(mu/(1+k*mu)))-log(fact(y))+((y-1)*log(1+(k*y)))-
(mu*(1+(k*y)))/(1+(k*mu));
prob=exp(loglike);
model y~general(loglike);
predict mu out=mu_gp;
ods output parameterestimates=para_est;
run;
data para_est;set sim;set para_est;
      para=parameter;
      est=estimate;
      se=standarderror;
keep para est se;run;
data k_hat;set para_est;
      if para='k' then do;
      k_hat=est;
      output;end;
keep k_hat;run;
proc print data= k_hat;run;
data mu11;set mu_gp;
      mu_gp=pred;
keep mu_gp nrep;run;
proc print data=mu11;run;
ods output simplestats=simpout3;
proc corr data=mu11;by nrep;
      VAR mu_gp;run;
data mu;set simpout3;
      para1=variable;
      mu_gp =mean;
      if para1='mu_gp' then do;
      mu_gp=mu_gp;
      output;end;
keep mu_gp nrep;run;
proc print data=mu;run;
```

```
data propose1;set k_hat;set mu;n=&n;
   propose1=(sqrt((n-1)/2))*(((mu_gp*(1+(k_hat*mu_gp))**2)/mu_gp)-1);
keep propose1;run;
proc print data=propose1;run;
data order;nrep=&nrep;
      do nrep=1 to nrep;
      output;end;
run;
data w5;set propose1;set order;by nrep;
retain w5;
      if (first.nrep) then do;
            w5=0;end;
            if propose1>1.645 then w5=w5+1;
            else;
      if (last.nrep) then output;
run;
proc sql;
      create table t5 as
      select sum(w5) as w5
      from w5;
quit;
data z5;set t5;nrep=&nrep;
      propose1=(w5*100)/nrep;
keep propose1;run;
proc print data=z5;run;
```

2.1.2 การเขียนโปรแกรมสำหรับคำนวณค่าสถิติทดสอบ $Z_{\overline{Y}}$

```
proc nlmixed data=sim;by nrep;
      parms b0=0 b1=0 k=0;
      eta= b0 + b1*x;
      mu=exp(eta);
      loglike=(y*log(mu/(1+k*mu)))-log(fact(y))+((y-1)*log(1+(k*y)))-
(mu*(1+(k*y)))/(1+(k*mu));
prob=exp(loglike);
model y~general(loglike);
ods output parameterestimates=para_est;
run;
data para_est;set sim;set para_est;
      para=parameter;
      est=estimate;
      se=standarderror;
keep para est se;run;
data k_hat;set para_est;
      if para='k' then do;
      k_hat=est;
      output;end;
keep k_hat;run;
proc print data= k_hat;run;
ods output simplestats=simpout4;
proc corr data=sim;by nrep;
      VAR y;run;
data ybar;set simpout4;
      para1=variable;
      ybar=mean;
      if para1='y' then do;
      ybar=ybar;
      output;end;
```

```
keep ybar nrep;run;
proc print data=ybar;run;
data propose2;set k_hat;set ybar;n=&n;
      propose2=(sqrt((n-1)/2))*(((ybar*(1+(k_hat*ybar))**2)/ybar)-1);
keep propose2;run;
proc print data=propose2;run;
data order;nrep=&nrep;
      do nrep=1 to nrep;
      output;end;
run;
data w6;set propose2;set order;by nrep;
retain w6;
      if (first.nrep) then do;
            w6=0;end;
            if propose2>1.645 then w6=w6+1;
            else;
      if (last.nrep) then output;
run;
proc sql;
      create table t6 as
      select sum(w6) as w6
      from w6;
quit;
data z6;set t6;nrep=&nrep;
      propose2=(w6*100)/nrep;
keep propose2;run;
proc print data=z6;run;
```

2.1.3 การเขียนโปรแกรมสำหรับคำนวณค่าสถิติทดสอบ Z_0

```
ods output simplestats=simpout2;
proc corr data=sim;by nrep;
      VAR y;run;
data ybar;set simpout2;
      para1=variable;
      ybar=mean;
      if para1='y' then do;
      ybar=ybar;
      output;end;
keep ybar nrep;run;
proc print data=ybar;run;
data s;set simpout2;
      para1=variable;
      s=stddev;
      if para1='y' then do;
      s=s;
      output;end;
keep s nrep;run;
proc print data=s;run;
data s2;set s;
      s2=s**2;
keep s2 nrep;run;
proc print data=s2;run;
data O_new;set ybar;set s2;n=&n;
      O_new=(sqrt((n-1)/2))*((s2/ybar)-1);
keep O_new nrep;run;
proc print data=O_new;run;
data order;nrep=&nrep;
```

```
      do nrep=1 to nrep;
      output;end;
run;
data w4;set O_new;set order;by nrep;
retain w4;
      if (first.nrep) then do;
            w4=0;end;
            if O_new>1.645 then w4=w4+1;
            else;
      if (last.nrep) then output;
run;
proc sql;
      create table t4 as
      select sum(w4) as w4
      from w4;
quit;
data z4;set t4;nrep=&nrep;
      O_new=(w4*100)/nrep;
keep O_new;run;
proc print data=z4;run;
```

2.1.4 การเขียนโปรแกรมสำหรับคำนวณค่าสถิติทดสอบวาลด์ที

```
proc nlmixed data=sim;by nrep;
      parms b0=0 b1=0 k=0;
      eta= b0 + b1*x;
      mu=exp(eta);
      loglike=(y*log(mu/(1+k*mu))-log(fact(y))+((y-1)*log(1+(k*y)))-
(mu*(1+(k*y)))/(1+(k*mu));
prob=exp(loglike);
model y~general(loglike);
ods output parameterestimates=para_est;
run;
data para_est;set sim;set para_est;
      para=parameter;
      est=estimate;
      se=standarderror;
keep para est se;run;
data wald1;set para_est;
      if para='k' then do;
      k_hat=est;
      se_k=se;
      output;end;
keep k_hat se_k;run;
proc print data=wald1;run;
data wald;set wald1;
      t=k_hat/se_k;
keep t;run;
proc print data=wald;run;
data order;nrep=&nrep;
      do nrep=1 to nrep;
      output;end;
run;
data w1;set wald;set order;by nrep;
retain w1;
      if (first.nrep) then do;
            w1=0;end;
            if t>1.645 then w1=w1+1;
```

```
            else;
      if (last.nrep) then output;
run;
proc sql;
      create table t1 as
      select sum(w1) as w1
      from w1;
quit;
data z1;set t1;nrep=&nrep;
      wald=(w1*100)/nrep;
keep wald;run;
proc print data=z1;run;
```

2.1.5 การเขียนโปรแกรมสำหรับคำนวณค่าสถิติทดสอบอัตราส่วนภาวะน่าจะเป็น

```
proc nlmixed data=sim;by nrep;
      parms b0=0 b1=0 k=0;
      eta= b0 + b1*x;
      mu=exp(eta);
      loglike=(y*log(mu/(1+k*mu)))-log(fact(y))+((y-1)*log(1+(k*y)))-
(mu*(1+(k*y)))/(1+(k*mu));
prob=exp(loglike);
model y~general(loglike);
ods output parameterestimates=para_est;
ods output fitstatistics=fitstat_GP;
run;
proc nlmixed data=sim;by nrep;
      parms b0=0 b1=0;
      eta1= b0 + b1*x;
      mu1= exp(eta1);
      model y~poisson(mu1);
ods output fitstatistics=fitstat_poi;
run;
data fitstat_GP;set fitstat_GP;
      Descri=Descr;
      if Descri='-2 Log Likelihood' then do;
      neg2logL_GP=value;
      output;end;
keep neg2logL_GP;run;
proc print data=fitstat_GP;run;
data fitstat_poi;set fitstat_poi;
      Descri=Descr;
      if Descri='-2 Log Likelihood' then do;
      neg2logL_poi=value;
      output;end;
keep neg2logL_poi;run;
proc print data=fitstat_poi;run;
data para_est;set sim;set para_est;
      para=parameter;
      est=estimate;
      se=standarderror;
keep para est se;run;
data k_hat;set para_est;
      if para='k' then do;
      k_hat=est;
      output;end;
keep k_hat;run;
proc print data=k_hat;run;
```

```
data sign;set k_hat;
      if k_hat>=0 then sign=1;
      if k_hat<0 then sign=-1;
run;
proc print data=sign;run;
data LRT;set fitstat_poi;set fitstat_GP;set sign;
      if neg2logL_GP>=neg2logL_poi then LRT=0;
      if neg2logL_GP<neg2logL_poi then LRT=sign*(SQRT(neg2logL_poi-
neg2logL_GP));
keep LRT;run;
proc print data=LRT;run;
data order;nrep=&nrep;
      do nrep=1 to nrep;
      output;end;
run;
data w2;set LRT;set order;by nrep;
retain w2;
      if (first.nrep) then do;
            w2=0;end;
            if LRT>1.645 then w2=w2+1;
            else;
      if (last.nrep) then output;
run;
proc sql;
      create table t2 as
      select sum(w2) as w2
      from w2;
quit;
data z2;set t2;nrep=&nrep;
      LRT=(w2*100)/nrep;
keep LRT;run;
proc print data=z2;run;
```

2.1.6 การเขียนโปรแกรมสำหรับคำนวณค่าสถิติทดสอบสกอร์

```
proc nlmixed data=sim;by nrep;
      parms b0=0 b1=0;
      eta1= b0 + b1*x;
      mu1= exp(eta1);
      model y~poisson(mu1);
predict mu1 out=mu_p;
ods output fitstatistics=fitstat_poi;
run;
data myObs;set mu_p;set sim;by nrep;
      p=pred;
      a=((y-p)**2)-y;
      b=p**2;
keep nrep y x p a b;run;
proc print data=myObs;run;
ods output simplestats=simpout;
ods output simplestats=simpout1;
proc corr data=myObS;by nrep;
      VAR a b;run;
data simpout;set simpout;
      para1=variable;
      sum1=sum;
      if para1='a' then do;
      sum_a=sum1;
```

```
      output;end;
keep sum_a;run;
proc print data=simpout;run;
data simpout1;set simpout1;
      para1=variable;
      sum1=sum;
      if para1='b' then do;
      sum_b=sum1;
      output;end;
keep sum_b;run;
proc print data=simpout1;run;
data score;set simpout;set simpout1;
      score=sum_a/sqrt(2*sum_b);
run;
proc print data=score;run;
data order;nrep=&nrep;
      do nrep=1 to nrep;
      output;end;
run;
data w3;set score;set order;by nrep;
retain w3;
      if (first.nrep) then do;
            w3=0;end;
            if score>1.645 then w3=w3+1;
            else;
      if (last.nrep) then output;
run;
proc sql;
      create table t3 as
      select sum(w3) as w3
      from w3;
quit;
data z3;set t3;nrep=&nrep;
      score=(w3*100)/nrep;
keep score;run;
proc print data=z3;run;
```

2.1.7 การเขียนโปรแกรมสำหรับคำนวณค่าสถิติทดสอบ Q^*

2.1.7.1 กรณีขนาดตัวอย่างเท่ากับ 30

```
ods output simplestats=simpout5;
proc corr data=sim;by nrep;
      VAR y;run;
data sumY;set simpout5;
      para1=variable;
      sumY=sum;
      if para1='y' then do;
      sumY=sumY;
      output;end;
keep sumY nrep;run;
proc print data=sumY;run;
data Y2;set sim;
      y2=y**2;
keep y y2 nrep;run;
proc print data=Y2;run;
ods output simplestats=simpout6;
```

```
proc corr data=Y2;by nrep;
      VAR y2;run;
data sumY2;set simpout6;
      para1=variable;
      sumY2=sum;
      if para1='y2' then do;
      sumY2=sumY2;
      output;end;
keep sumY2 nrep;run;
proc print data=sumY2;run;
data Q;set sumY2;set sumY;n=&n;
      Q=((n*sumY2)/sumY)-sumY;
keep Q nrep;run;
proc print data=Q;run;
data order;nrep=&nrep;
      do nrep=1 to nrep;
      output;end;
run;
data w7n30;set Q;set order;by nrep;
retain w7n30;
      if (first.nrep) then do;
            w7n30=0;end;
            if Q>42.714 then w7n30=w7n30+1;
            else;
      if (last.nrep) then output;
run;
proc sql;
      create table t7n30 as
      select sum(w7n30) as w7n30
      from w7n30;
quit;
data z7n30;set t7n30;nrep=&nrep;
      Q=(w7n30*100)/nrep;
keep Q;run;
proc print data=z7n30;run;
```

2.1.7.2 กรณีขนาดตัวอย่างเท่ากับ 50

```
ods output simplestats=simpout5;
proc corr data=sim;by nrep;
      VAR y;run;
data sumY;set simpout5;
      para1=variable;
      sumY=sum;
      if para1='y' then do;
      sumY=sumY;
      output;end;
keep sumY nrep;run;
proc print data=sumY;run;
data Y2;set sim;
      y2=y**2;
keep y y2 nrep;run;
proc print data=Y2;run;
ods output simplestats=simpout6;
proc corr data=Y2;by nrep;
      VAR y2;run;
data sumY2;set simpout6;
```

```
       para1=variable;
       sumY2=sum;
       if para1='y2' then do;
       sumY2=sumY2;
       output;end;
keep sumY2 nrep;run;
proc print data=sumY2;run;
data Q;set sumY2;set sumY;n=&n;
       Q=((n*sumY2)/sumY)-sumY;
keep Q nrep;run;
proc print data=Q;run;
data order;nrep=&nrep;
       do nrep=1 to nrep;
       output;end;
run;
data w7n50;set Q;set order;by nrep;
retain w7n50;
       if (first.nrep) then do;
             w7n50=0;end;
             if Q>66.583 then w7n50=w7n50+1;
             else;
       if (last.nrep) then output;
run;
proc sql;
       create table t7n50 as
       select sum(w7n50) as w7n50
       from w7n50;
quit;
data z7n50;set t7n50;nrep=&nrep;
       Q=(w7n50*100)/nrep;
keep Q;run;
proc print data=z7n50;run;
```

2.2 กรณี Underdispersion

2.2.1 การเขียนโปรแกรมสำหรับคำนวณค่าสถิติทดสอบ $Z_{\bar{\hat{\mu}}}$

```
proc nlmixed data=sim;by nrep;
       parms b0=0 b1=0 k=0;
       eta= b0 + b1*x;
       mu=exp(eta);
       loglike=(y*log(mu/(1+k*mu)))-log(fact(y))+((y-1)*log(1+(k*y)))-
(mu*(1+(k*y)))/(1+(k*mu));
prob=exp(loglike);
model y~general(loglike);
predict mu out=mu_gp;
ods output parameterestimates=para_est;
run;
data para_est;set sim;set para_est;
       para=parameter;
       est=estimate;
       se=standarderror;
keep para est se;run;
data k_hat;set para_est;
       if para='k' then do;
       k_hat=est;
       output;end;
keep k_hat;run;
```

```
proc print data= k_hat;run;
data mu11;set mu_gp;
      mu_gp=pred;
keep mu_gp nrep;run;
proc print data=mu11;run;
ods output simplestats=simpout3;
proc corr data=mu11;by nrep;
      VAR mu_gp;run;
data mu;set simpout3;
      para1=variable;
      mu_gp =mean;
      if para1='mu_gp' then do;
      mu_gp=mu_gp;
      output;end;
keep mu_gp nrep;run;
proc print data=mu;run;

data propose1;set k_hat;set mu;n=&n;
   propose1=(sqrt((n-1)/2))*(((mu_gp*(1+(k_hat*mu_gp))**2)/mu_gp)-1);
keep propose1;run;
proc print data=propose1;run;
data order;nrep=&nrep;
      do nrep=1 to nrep;
      output;end;
run;
data w5;set propose1;set order;by nrep;
retain w5;
      if (first.nrep) then do;
            w5=0;end;
            if propose1<-1.645 then w5=w5+1;
            else;
      if (last.nrep) then output;
run;
proc sql;
      create table t5 as
      select sum(w5) as w5
      from w5;
quit;
data z5;set t5;nrep=&nrep;
      propose1=(w5*100)/nrep;
keep propose1;run;
proc print data=z5;run;
```

2.2.2 การเขียนโปรแกรมสำหรับคำนวณค่าสถิติทดสอบ $Z_{\overline{Y}}$

```
proc nlmixed data=sim;by nrep;
      parms b0=0 b1=0 k=0;
      eta= b0 + b1*x;
      mu=exp(eta);
      loglike=(y*log(mu/(1+k*mu)))-log(fact(y))+((y-1)*log(1+(k*y)))-
(mu*(1+(k*y)))/(1+(k*mu));
prob=exp(loglike);
model y~general(loglike);
ods output parameterestimates=para_est;
run;
data para_est;set sim;set para_est;
      para=parameter;
      est=estimate;
```

```
      se=standarderror;
keep para est se;run;
data k_hat;set para_est;
      if para='k' then do;
      k_hat=est;
      output;end;
keep k_hat;run;
proc print data= k_hat;run;
ods output simplestats=simpout4;
proc corr data=sim;by nrep;
      VAR y;run;
data ybar;set simpout4;
      para1=variable;
      ybar=mean;
      if para1='y' then do;
      ybar=ybar;
      output;end;
keep ybar nrep;run;
proc print data=ybar;run;
data propose2;set k_hat;set ybar;n=&n;
      propose2=(sqrt((n-1)/2))*(((ybar*(1+(k_hat*ybar))**2)/ybar)-1);
keep propose2;run;
proc print data=propose2;run;
data order;nrep=&nrep;
      do nrep=1 to nrep;
      output;end;
run;
data w6;set propose2;set order;by nrep;
retain w6;
      if (first.nrep) then do;
            w6=0;end;
            if propose2<-1.645 then w6=w6+1;
            else;
      if (last.nrep) then output;
run;
proc sql;
      create table t6 as
      select sum(w6) as w6
      from w6;
quit;
data z6;set t6;nrep=&nrep;
      propose2=(w6*100)/nrep;
keep propose2;run;
proc print data=z6;run;
```

2.2.3 การเขียนโปรแกรมสำหรับคำนวณค่าสถิติทดสอบ Z_0

```
ods output simplestats=simpout2;
proc corr data=sim;by nrep;
      VAR y;run;
data ybar;set simpout2;
      para1=variable;
      ybar=mean;
      if para1='y' then do;
      ybar=ybar;
      output;end;
keep ybar nrep;run;
proc print data=ybar;run;
```

```
data s;set simpout2;
      para1=variable;
      s=stddev;
      if para1='y' then do;
      s=s;
      output;end;
keep s nrep;run;
proc print data=s;run;
data s2;set s;
      s2=s**2;
keep s2 nrep;run;
proc print data=s2;run;
data O_new;set ybar;set s2;n=&n;
      O_new=(sqrt((n-1)/2))*((s2/ybar)-1);
keep O_new nrep;run;
proc print data=O_new;run;
data order;nrep=&nrep;
      do nrep=1 to nrep;
      output;end;
run;
data w4;set O_new;set order;by nrep;
retain w4;
      if (first.nrep) then do;
            w4=0;end;
            if O_new<-1.645 then w4=w4+1;
            else;
      if (last.nrep) then output;
run;
proc sql;
      create table t4 as
      select sum(w4) as w4
      from w4;
quit;
data z4;set t4;nrep=&nrep;
      O_new=(w4*100)/nrep;
keep O_new;run;
proc print data=z4;run;
```

2.2.4 การเขียนโปรแกรมสำหรับคำนวณค่าสถิติทดสอบวาลด์ที

```
proc nlmixed data=sim;by nrep;
      parms b0=0 b1=0 k=0;
      eta= b0 + b1*x;
      mu=exp(eta);
      loglike=(y*log(mu/(1+k*mu)))-log(fact(y))+((y-1)*log(1+(k*y)))-
(mu*(1+(k*y)))/(1+(k*mu));
prob=exp(loglike);
model y~general(loglike);
ods output parameterestimates=para_est;
run;
data para_est;set sim;set para_est;
      para=parameter;
      est=estimate;
      se=standarderror;
keep para est se;run;
data wald1;set para_est;
      if para='k' then do;
      k_hat=est;
```

```
        se_k=se;
        output;end;
keep k_hat se_k;run;
proc print data=wald1;run;
data wald;set wald1;
        t=k_hat/se_k;
keep t;run;
proc print data=wald;run;
data order;nrep=&nrep;
        do nrep=1 to nrep;
        output;end;
run;
data w1;set wald;set order;by nrep;
retain w1;
        if (first.nrep) then do;
                w1=0;end;
                if t<-1.645 then w1=w1+1;
                else;
        if (last.nrep) then output;
run;
proc sql;
        create table t1 as
        select sum(w1) as w1
        from w1;
quit;
data z1;set t1;nrep=&nrep;
        wald=(w1*100)/nrep;
keep wald;run;
proc print data=z1;run;
```

2.2.5 การเขียนโปรแกรมสำหรับคำนวณค่าสถิติทดสอบอัตราส่วนภาวะน่าจะเป็น

```
proc nlmixed data=sim;by nrep;
        parms b0=0 b1=0 k=0;
        eta= b0 + b1*x;
        mu=exp(eta);
        loglike=(y*log(mu/(1+k*mu)))-log(fact(y))+((y-1)*log(1+(k*y)))-
(mu*(1+(k*y)))/(1+(k*mu));
prob=exp(loglike);
model y~general(loglike);
ods output parameterestimates=para_est;
ods output fitstatistics=fitstat_GP;
run;
proc nlmixed data=sim;by nrep;
        parms b0=0 b1=0;
        eta1= b0 + b1*x;
        mu1= exp(eta1);
        model y~poisson(mu1);
ods output fitstatistics=fitstat_poi;
run;
data fitstat_GP;set fitstat_GP;
        Descri=Descr;
        if Descri='-2 Log Likelihood' then do;
        neg2logL_GP=value;
        output;end;
keep neg2logL_GP;run;
proc print data=fitstat_GP;run;
data fitstat_poi;set fitstat_poi;
```

```
        Descri=Descr;
        if Descri='-2 Log Likelihood' then do;
        neg2logL_poi=value;
        output;end;
keep neg2logL_poi;run;
proc print data=fitstat_poi;run;
data para_est;set sim;set para_est;
        para=parameter;
        est=estimate;
        se=standarderror;
keep para est se;run;
data k_hat;set para_est;
        if para='k' then do;
        k_hat=est;
        output;end;
keep k_hat;run;
proc print data=k_hat;run;
data sign;set k_hat;
        if k_hat>=0 then sign=1;
        if k_hat<0 then sign=-1;
run;
proc print data=sign;run;
data LRT;set fitstat_poi;set fitstat_GP;set sign;
        if neg2logL_GP>=neg2logL_poi then LRT=0;
        if neg2logL_GP<neg2logL_poi then LRT=sign*(SQRT(neg2logL_poi-
neg2logL_GP));
keep LRT;run;
proc print data=LRT;run;
data order;nrep=&nrep;
        do nrep=1 to nrep;
        output;end;
run;
data w2;set LRT;set order;by nrep;
retain w2;
        if (first.nrep) then do;
                w2=0;end;
                if LRT<-1.645 then w2=w2+1;
                else;
        if (last.nrep) then output;
run;
proc sql;
        create table t2 as
        select sum(w2) as w2
        from w2;
quit;
data z2;set t2;nrep=&nrep;
        LRT=(w2*100)/nrep;
keep LRT;run;
proc print data=z2;run;
```

2.2.6 การเขียนโปรแกรมสำหรับคำนวณค่าสถิติทดสอบสกอร์

```
proc nlmixed data=sim;by nrep;
        parms b0=0 b1=0;
        eta1= b0 + b1*x;
        mu1= exp(eta1);
        model y~poisson(mu1);
predict mu1 out=mu_p;
```

```
ods output fitstatistics=fitstat_poi;
run;
data myObs;set mu_p;set sim;by nrep;
      p=pred;
      a=((y-p)**2)-y;
      b=p**2;
keep nrep y x p a b;run;
proc print data=myObs;run;
ods output simplestats=simpout;
ods output simplestats=simpout1;
proc corr data=myObS;by nrep;
      VAR a b;run;
data simpout;set simpout;
      para1=variable;
      sum1=sum;
      if para1='a' then do;
      sum_a=sum1;
      output;end;
keep sum_a;run;
proc print data=simpout;run;
data simpout1;set simpout1;
      para1=variable;
      sum1=sum;
      if para1='b' then do;
      sum_b=sum1;
      output;end;
keep sum_b;run;
proc print data=simpout1;run;
data score;set simpout;set simpout1;
      score=sum_a/sqrt(2*sum_b);
run;
proc print data=score;run;
data order;nrep=&nrep;
      do nrep=1 to nrep;
      output;end;
run;
data w3;set score;set order;by nrep;
retain w3;
      if (first.nrep) then do;
            w3=0;end;
            if score<-1.645 then w3=w3+1;
            else;
      if (last.nrep) then output;
run;
proc sql;
      create table t3 as
      select sum(w3) as w3
      from w3;
quit;
data z3;set t3;nrep=&nrep;
      score=(w3*100)/nrep;
keep score;run;
proc print data=z3;run;
```

2.2.7 การเขียนโปรแกรมสำหรับคำนวณค่าสถิติทดสอบ Q^*

2.2.7.1 กรณีขนาดตัวอย่างเท่ากับ 30

```
ods output simplestats=simpout5;
proc corr data=sim;by nrep;
      VAR y;run;
data sumY;set simpout5;
      para1=variable;
      sumY=sum;
      if para1='y' then do;
      sumY=sumY;
      output;end;
keep sumY nrep;run;
proc print data=sumY;run;
data Y2;set sim;
      y2=y**2;
keep y y2 nrep;run;
proc print data=Y2;run;
ods output simplestats=simpout6;
proc corr data=Y2;by nrep;
      VAR y2;run;
data sumY2;set simpout6;
      para1=variable;
      sumY2=sum;
      if para1='y2' then do;
      sumY2=sumY2;
      output;end;
keep sumY2 nrep;run;
proc print data=sumY2;run;
data Q;set sumY2;set sumY;n=&n;
      Q=((n*sumY2)/sumY)-sumY;
keep Q nrep;run;
proc print data=Q;run;
data order;nrep=&nrep;
      do nrep=1 to nrep;
      output;end;
run;
data w7n30;set Q;set order;by nrep;
retain w7n30;
      if (first.nrep) then do;
            w7n30=0;end;
            if Q<18.00 then w7n30=w7n30+1;
            else;
      if (last.nrep) then output;
run;
proc sql;
      create table t7n30 as
      select sum(w7n30) as w7n30
      from w7n30;
quit;
data z7n30;set t7n30;nrep=&nrep;
      Q=(w7n30*100)/nrep;
keep Q;run;
proc print data=z7n30;run;
```

2.2.7.2 กรณีขนาดตัวอย่างเท่ากับ 50

```
ods output simplestats=simpout5;
proc corr data=sim;by nrep;
      VAR y;run;
data sumY;set simpout5;
      para1=variable;
      sumY=sum;
      if para1='y' then do;
      sumY=sumY;
      output;end;
keep sumY nrep;run;
proc print data=sumY;run;
data Y2;set sim;
      y2=y**2;
keep y y2 nrep;run;
proc print data=Y2;run;
ods output simplestats=simpout6;
proc corr data=Y2;by nrep;
      VAR y2;run;
data sumY2;set simpout6;
      para1=variable;
      sumY2=sum;
      if para1='y2' then do;
      sumY2=sumY2;
      output;end;
keep sumY2 nrep;run;
proc print data=sumY2;run;
data Q;set sumY2;set sumY;n=&n;
      Q=((n*sumY2)/sumY)-sumY;
keep Q nrep;run;
proc print data=Q;run;
data order;nrep=&nrep;
      do nrep=1 to nrep;
      output;end;
run;
data w7n50;set Q;set order;by nrep;
retain w7n50;
      if (first.nrep) then do;
            w7n50=0;end;
            if Q<34.231 then w7n50=w7n50+1;
            else;
      if (last.nrep) then output;
run;
proc sql;
      create table t7n50 as
      select sum(w7n50) as w7n50
      from w7n50;
quit;
data z7n50;set t7n50;nrep=&nrep;
      Q=(w7n50*100)/nrep;
keep Q;run;
proc print data=z7n50;run;
```

3. การเรียกใช้แมคโครภายใต้เงื่อนไขต่าง ๆ

3.1 กรณี Overdispersion

3.1.1 กรณีของขนาดตัวอย่างเท่ากับ 30

```
%case(nrep=5000,n=30,k1=0.0);
%case(nrep=5000,n=30,k1=0.015);
%case(nrep=5000,n=30,k1=0.02);
%case(nrep=5000,n=30,k1=0.025);
%case(nrep=5000,n=30,k1=0.03);
%case(nrep=5000,n=30,k1=0.035);
%case(nrep=5000,n=30,k1=0.04);
%case(nrep=5000,n=30,k1=0.05);
%case(nrep=5000,n=30,k1=0.06);
%case(nrep=5000,n=30,k1=0.07);
%case(nrep=5000,n=30,k1=0.08);
%case(nrep=5000,n=30,k1=0.09);
%case(nrep=5000,n=30,k1=0.1);
%case(nrep=5000,n=30,k1=0.2);
```

3.1,2 กรณีของขนาดตัวอย่างเท่ากับ 50

```
%case(nrep=5000,n=50,k1=0.0);
%case(nrep=5000,n=50,k1=0.015);
%case(nrep=5000,n=50,k1=0.02);
%case(nrep=5000,n=50,k1=0.025);
%case(nrep=5000,n=50,k1=0.03);
%case(nrep=5000,n=50,k1=0.035);
%case(nrep=5000,n=50,k1=0.04);
%case(nrep=5000,n=50,k1=0.05);
%case(nrep=5000,n=50,k1=0.06);
%case(nrep=5000,n=50,k1=0.07);
%case(nrep=5000,n=50,k1=0.08);
%case(nrep=5000,n=50,k1=0.09);
%case(nrep=5000,n=50,k1=0.1);
%case(nrep=5000,n=50,k1=0.2);
```

3.1.3 กรณีของขนาดตัวอย่างเท่ากับ 100

```
%case(nrep=5000,n=100,k1=0.0);
%case(nrep=5000,n=100,k1=0.015);
%case(nrep=5000,n=100,k1=0.02);
%case(nrep=5000,n=100,k1=0.025);
%case(nrep=5000,n=100,k1=0.03);
%case(nrep=5000,n=100,k1=0.035);
%case(nrep=5000,n=100,k1=0.04);
%case(nrep=5000,n=100,k1=0.05);
%case(nrep=5000,n=100,k1=0.06);
%case(nrep=5000,n=100,k1=0.07);
%case(nrep=5000,n=100,k1=0.08);
%case(nrep=5000,n=100,k1=0.09);
%case(nrep=5000,n=100,k1=0.1);
%case(nrep=5000,n=100,k1=0.2);
```

3.1.4 กรณีของขนาดตัวอย่างเท่ากับ 200

```
%case(nrep=5000,n=200,k1=0.0);
%case(nrep=5000,n=200,k1=0.015);
%case(nrep=5000,n=200,k1=0.02);
%case(nrep=5000,n=200,k1=0.025);
%case(nrep=5000,n=200,k1=0.03);
%case(nrep=5000,n=200,k1=0.035);
%case(nrep=5000,n=200,k1=0.04);
%case(nrep=5000,n=200,k1=0.05);
%case(nrep=5000,n=200,k1=0.06);
%case(nrep=5000,n=200,k1=0.07);
%case(nrep=5000,n=200,k1=0.08);
%case(nrep=5000,n=200,k1=0.09);
%case(nrep=5000,n=200,k1=0.1);
%case(nrep=5000,n=200,k1=0.2);
```

3.2 กรณี Underdispersion

3.2.1 กรณีของขนาดตัวอย่างเท่ากับ 30

```
%case(nrep=5000,n=30,k1=0.0);
%case(nrep=5000,n=30,k1=-0.015);
%case(nrep=5000,n=30,k1=-0.02);
%case(nrep=5000,n=30,k1=-0.025);
%case(nrep=5000,n=30,k1=-0.03);
%case(nrep=5000,n=30,k1=-0.035);
%case(nrep=5000,n=30,k1=-0.04);
%case(nrep=5000,n=30,k1=-0.05);
%case(nrep=5000,n=30,k1=-0.06);
%case(nrep=5000,n=30,k1=-0.07);
%case(nrep=5000,n=30,k1=-0.08);
%case(nrep=5000,n=30,k1=-0.09);
%case(nrep=5000,n=30,k1=-0.1);
%case(nrep=5000,n=30,k1=-0.2);
```

3.2.2 กรณีของขนาดตัวอย่างเท่ากับ 50

```
%case(nrep=5000,n=50,k1=0.0);
%case(nrep=5000,n=50,k1=-0.015);
%case(nrep=5000,n=50,k1=-0.02);
%case(nrep=5000,n=50,k1=-0.025);
%case(nrep=5000,n=50,k1=-0.03);
%case(nrep=5000,n=50,k1=-0.035);
%case(nrep=5000,n=50,k1=-0.04);
%case(nrep=5000,n=50,k1=-0.05);
%case(nrep=5000,n=50,k1=-0.06);
%case(nrep=5000,n=50,k1=-0.07);
%case(nrep=5000,n=50,k1=-0.08);
%case(nrep=5000,n=50,k1=-0.09);
%case(nrep=5000,n=50,k1=-0.1);
%case(nrep=5000,n=50,k1=-0.2);
```

3.2.3 กรณีของขนาดตัวอย่างเท่ากับ 100

```
%case(nrep=5000,n=100,k1=0.0);
%case(nrep=5000,n=100,k1=-0.015);
%case(nrep=5000,n=100,k1=-0.02);
%case(nrep=5000,n=100,k1=-0.025);
%case(nrep=5000,n=100,k1=-0.03);
%case(nrep=5000,n=100,k1=-0.035);
%case(nrep=5000,n=100,k1=-0.04);
%case(nrep=5000,n=100,k1=-0.05);
%case(nrep=5000,n=100,k1=-0.06);
%case(nrep=5000,n=100,k1=-0.07);
%case(nrep=5000,n=100,k1=-0.08);
%case(nrep=5000,n=100,k1=-0.09);
%case(nrep=5000,n=100,k1=-0.1);
%case(nrep=5000,n=100,k1=-0.2);
```

3.2.4 กรณีของขนาดตัวอย่างเท่ากับ 200

```
%case(nrep=5000,n=200,k1=0.0);
%case(nrep=5000,n=200,k1=-0.015);
%case(nrep=5000,n=200,k1=-0.02);
%case(nrep=5000,n=200,k1=-0.025);
%case(nrep=5000,n=200,k1=-0.03);
%case(nrep=5000,n=200,k1=-0.035);
%case(nrep=5000,n=200,k1=-0.04);
%case(nrep=5000,n=200,k1=-0.05);
%case(nrep=5000,n=200,k1=-0.06);
%case(nrep=5000,n=200,k1=-0.07);
%case(nrep=5000,n=200,k1=-0.08);
%case(nrep=5000,n=200,k1=-0.09);
%case(nrep=5000,n=200,k1=-0.1);
%case(nrep=5000,n=200,k1=-0.2);
```

ภาคผนวก ค

โปรแกรมสำเร็จรูป SAS® version 9.1

สำหรับข้อมูลจำนวนปูเพศผู้จำแนกตามคุณลักษณะของปูเพศเมีย

กำหนดให้ ตัวแปรตอบสนอง (y) คือ จำนวนปูเพศผู้ และตัวแปรอธิบาย (x) แทน น้ำหนักของปูเพศเมีย (หน่วยเป็นกิโลกรัม) k แทน พารามิเตอร์การกระจาย($\hat{\square}$) ซึ่งโปรแกรมที่ใช้ในการคำนวณค่าพารามิเตอร์ต่าง ๆ ของตัวแบบการถดถอยปัวซงและตัวแบบการถดถอยปัวซงนัยทั่วไปแบบที่ 2 มีดังนี้

```
data crab;
      input y x@@;
datalines;
8 3.05 4 2.25 0 1.60 5 2.00 6 1.95 0 1.85 6 2.50 1 2.00 0 2.45
4 2.60 0 2.90 2 1.85 0 2.75 7 3.05 6 2.80 0 1.80 4 1.95 0 2.25
0 2.15 3 2.25 3 2.28 3 2.45 6 2.25 5 3.30 6 2.50 3 2.00 1 2.87
0 1.85 0 1.70 0 2.20 10 3.20 3 2.92 4 2.10 2 1.65 0 2.60 1 2.00
1 3.00 0 3.20 4 3.28 7 2.80 4 3.73 5 2.90 4 1.47 0 2.00 2 1.90
3 2.30 1 1.97 0 2.35 0 1.90 4 2.85 15 3.00 0 1.80 0 2.65 0 2.10
0 1.30 1 1.60 0 1.55 0 1.20 0 1.90 0 2.25 0 2.20 3 3.10 12 3.23
0 2.10 1 2.90 0 2.10 0 1.65 0 1.80 5 2.15 6 2.63 9 3.25 6 1.80
8 2.00 4 2.30 0 2.15 0 3.05 8 3.05 0 2.40 0 2.00 3 3.00 3 2.90
6 3.15 2 2.10 14 2.30 5 3.85 0 1.80 1 1.65 4 3.02 6 2.70 2 2.02
5 2.80 0 1.40 0 2.20 0 1.55 0 2.62 0 1.60 0 2.30 3 2.70 4 2.90
4 2.80 2 3.28 1 1.60 1 2.20 9 2.30 5 2.10 4 1.95 0 2.55 5 3.10
3 3.60 0 2.30 3 3.15 1 2.55 0 1.90 4 2.55 4 3.50 1 2.80 7 5.20
4 1.60 6 2.30 4 3.20 1 2.40 0 2.65 0 2.75 0 2.15 0 1.30 0 2.40
3 2.30 10 2.25 5 2.70 3 3.25 8 2.95 0 2.55 2 2.17 0 1.80 10 1.90
5 2.05 5 2.40 0 1.90 2 3.33 5 2.70 1 2.80 0 2.63 3 2.20 0 2.00
8 3.05 3 3.32 6 2.50 5 2.40 2 2.60 1 3.00 0 2.10 0 2.25 0 3.20
3 2.40 8 2.10 6 2.60 0 2.22 5 2.70 4 2.55 0 1.95 0 2.30 4 2.35
6 2.25 9 3.00 5 2.10 3 3.20 0 2.60 1 3.10 11 3.05 0 1.90 7 2.75
3 2.75 0 2.00

;
proc nlmixed data=crab;
      parms b0=0 b1=0;
      eta1= b0 + b1*x;
      mu1= exp(eta1);
      model y~poisson(mu1);
run;
proc nlmixed data=crab;
      parms b0=0 b1=0 k=0;
      eta=b0+(b1*x);
      mu=exp(eta);
      loglike=(y*log(mu/(1+k*mu)))-log(fact(y))+((y-1)*log(1+(k*y)))-
(mu*(1+(k*y)))/(1+(k*mu));
prob=exp(loglike);
model y~general(loglike);
run;
```

ประวัติผู้วิจัย

ชื่อ-สกุล นางสาวเมษิยา แย้มเจริญกิจ
ที่อยู่ 343/9 หมู่ 2 ต. เกาะหลัก อ. เมืองฯ จ. ประจวบคีรีขันธ์ 77000

ประวัติการศึกษา

พ.ศ. 2551 สำเร็จการศึกษาปริญญาวิทยาศาสตรบัณฑิต สาขาสถิติ มหาวิทยาลัยศิลปากร เกียรตินิยมอันดับสอง

พ.ศ. 2552 ศึกษาต่อระดับปริญญาวิทยาศาสตรมหาบัณฑิต สาขาสถิติประยุกต์ มหาวิทยาลัยศิลปากร