Development of Negative Binomial –Lindley distribution and Its Application.

การพัฒนาการแจกแจงความน่าจะเป็นของตัวแปรสุ่มทวินามลบ-ลินเลย์และการประยุกต์

ทวีป พรหมอยู่

ABSTRACT

This study proposes development of negative binomial –lindley distribution and its application distribution based on negative binomial – generalized Lindley distribution. The probability mass function of random variable for negative binomial – generalized Lindley which have four parameters (r, α ,θ ,γ). The probability mass function of negative binomial – generalized Lindley becomes probability mass function of negative binomial – Lindley when substituting $\alpha=1$ and $\gamma=1$. The probability mass function of negative binomial – generalized Lindley has the mathematical properties of negative binomial – generalized Lindley distribution such as factorial moments, first - fourth moments, mean, variance, coefficient of skewness, coefficient of kurtosis, moment generation function and characteristic function. Next, comparison of estimating parameters for negative binomial – generalized Lindley distribution with maximum likelihood method and moment method by using simulation study. The sample sizes considered are 50, 100, 200 and 500 with repeated 500 times for each simulation. The simulation study found that estimating parameters of negative binomial – generalized Lindley distribution by using moment method is given mean square error maximum likelihood method for almost situation. Finally, application study are used a real data set which number of injured from the accident on major road in Nakorn-Phathom province of Thailand in 2012. The real data are fitted by the Poisson distribution, negative binomial distribution and negative binomial – generalized Lindley distribution by using the chi-square goodness of fit test. The maximum likelihood method provides very poor fit for the Poisson distribution and negative binomial distribution and acceptable fits for the negative binomial – generalized Lindley distribution.

Keywords : Negative Binomial –Lindley, Distribution, Application

บทคัดย่อ

การวิจัยครั้งนี้เป็นการพัฒนาการแจกแจงความน่าจะเป็นของตัวแปรสุ่มทวินามลบ-ลินเลย์และการประยุกต์ โดยสร้างฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มทวินามลบ – ลินเลย์วางนัยทั่วไปที่มีพารามิเตอร์ทั้งหมด 4 ตัว (r, α ,θ , γ) แต่เมื่อพารามิเตอร์ $\alpha=1$ และ $\gamma=1$ จะลดรูปได้การแจก ทวินามลบ-ลินเลย์ และศึกษาคุณสมบัติเชิงคณิตศาสตร์ของตัวแปรสุ่มที่มีการแจกแจงทวินามลบ-ลินเลย์วางนัยทั่วไป ได้แก่ แฟกทอเรียลโมเมนต์ (Factorial Moments) โมเมนต์อันดับที่ 1 - 4 (First - Fourth Moments) ค่าเฉลี่ย (Mean) ความแปรปรวน (Variance) สัมประสิทธิ์ความเบ้ (Coefficient of Skewness) สัมประสิทธิ์ความโด่ง (Coefficient of Kurtosis) ฟังก์ชันโมเมนต์เวียนบังเกิด (Moment Generation Function) และ ฟังก์ชันลักษณะเฉพาะ (Characteristic Function) พร้อมทั้งเปรียบเทียบการประมาณค่าพารามิเตอร์ของการแจกแจงทวินามลบ-ลินเลย์วางนัยทั่วไปโดยใช้วิธีภาวะน่าจะเป็นสูงสุดและวิธีโมเมนต์ ได้จำลองข้อมูลตามสถานการณ์ที่กำหนดโดยใช้ขนาดตัวอย่างเท่ากับ 50,100,200 และ 500 และ ทำการจำลองซ้ำ 500 ครั้ง พบว่าการประมาณค่าพารามิเตอร์การแจกแจงทวินามลบ-ลินเลย์วางนัยทั่วไปด้วยวิธีโมเมนต์ จะให้ค่าเฉลี่ย กำลังสองน้อยที่สุด (Mean Square Error) ต่ำกว่า วิธีภาวะน่าจะเป็นสูงสุด ทุกขนาด

ทวีป พรหมอยู่ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏสวนสุนันทา, กรุงเทพมหานคร โทร 0 2160 1 111 โทรสาร 0 2160 1 010, E-mail : promyoo_t@hotmail.com

ตัวอย่างเกือบทุกกรณี และได้นำไป ประยุกต์ กับข้อมูลจริง เป็นข้อมูลอุบัติเหตุบนถนนทางหลวงของจังหวัดนครปฐม ปี 2555 โดยใช้การ ทดสอบภาวะสารรูปสนิทดี ที่ใช้ตัวสถิติทดสอบไคสแควร์ ใช้การแจกแจงปัวซอง การแจกแจงทวินามลบ และ การแจกแจงทวินามลบ -ลินเลย์ วางใน ทั่วไป พบว่าข้อมูลชุดนี้มีการแจกแจงทวินามลบ -ลินเลย์ วางใน ทั่วไป เหมาะสมที่สุด

คำสำคัญ : ทวินามลบ-ลินเลย์, การแจกแจงความน่าจะเป็น, การประยุกต์

บทนำ

การแจกแจงความน่าจะเป็นของแต่ละตัวแปรสุ่มมีหลากหลายรูปแบบขึ้นอยู่กับลักษณะของการทดลองสุ่มที่ทำให้เกิดตัวแปรสุ่มนั้น ๆ เช่น การทดลองที่ตัวแปรสุ่มมีโอกาสเกิดขึ้นได้เพียง 2 ค่า คือค่าที่ต้องการกับค่าที่ไม่ต้องการด้วยความน่าจะเป็นเท่ากับ p และ 1-p ตามลำดับ เรียกว่าการทดลองสุ่มลักษณะนี้ว่าการทดลองเบอร์นูลี ตัวแปรสุ่มที่เกิดจากการทดลองนี้เรียกว่าตัวแปรสุ่มเบอร์นูลี และเรียกการแจกแจงความน่าจะเป็นของตัวแปรสุ่มนี้ว่าการแจกแจงเบอร์นูลีซึ่งเป็นการแจกแจงไม่ต่อเนื่อง หรือ การทดลองที่ตัวแปรสุ่มมีได้ n ค่าที่แตกต่างกันแต่ละค่ามีความน่าจะเป็นเท่ากันและเท่ากับ 1/ n แล้วเรียกว่าการทดลองสุ่มลักษณะนี้ว่าการทดลองสม่ำเสมอ(uniform distribution) ตัวแปรสุ่มที่เกิดจากการทดลองนี้เรียกว่าตัวแปรสุ่มสม่ำเสมอ และเรียกการแจกแจงความน่าจะเป็นของตัวแปรสุ่มนี้ว่าการแจกแจงสม่ำเสมอ (uniform distribution) ซึ่งเป็นการแจกแจงไม่ต่อเนื่อง เป็นต้น การได้มาซึ่งการแจกแจงความน่าจะเป็นของตัวแปรสุ่มลักษณะนี้ถือเป็นการได้มาโดยวิธีสังเกตจากการทดลองสุ่ม นอกจากการได้มาด้วยวิธีการสังเกตจากการทดลองสุ่มโดยตรงแล้วอาจจะได้จากการพัฒนาจากการแจกแจงความน่าจะเป็นที่มีอยู่แล้วก็ได้ เช่น การแจกแจงเรขาคณิตได้จากการพัฒนาแจกแจงทวินามลบ (Negative Binomial Random Variable)การแจกแจงทวินามลบเกิดจากการผสมระหว่างการแจกแจงปัวซงกับการแจกแจงแกมมา การแจกแจงลินเลย์วางนัยทั่วไป (Generalized Lindley Distribution)ได้จากการพัฒนาการแจกแจงลินเลย์ (Lindley Distribution) การแจกแจงลินเลย์เกิดจากการผสมระหว่างการแจกแจงเอกโปเนนเชียลกับการแจกแจงแกมมา เป็นต้น

การสร้างหรือพัฒนาการแจกแจงความน่าจะเป็นของตัวแปรสุ่มเป็นสาระสำคัญในทฤษฎีความน่าจะเป็นที่นักคณิตศาสตร์ได้ดำเนินการมาเป็นเวลาช้านาน ซึ่งจะเป็นประโยชน์ให้การคำนวณความน่าจะเป็นของการเกิดเหตุการณ์ในหลากหลายรูปแบบทำได้ง่ายและรวดเร็วขึ้น และการได้มาซึ่งการแจกแจงความน่าจะเป็นของตัวแปรสุ่มแบบนี้ยังแสดงถึงความก้าวหน้าในการพัฒนาเนื้อหาสาระทางคณิตศาสตร์ที่เกี่ยวกับความน่าจะเป็นอีกด้วย และการสร้างหรือพัฒนาการแจกแจงความน่าจะเป็นของตัวแปรสุ่มมักจะดำเนินการไปพร้อมๆ กับการพัฒนาคุณสมบัติเชิงความน่าจะเป็นของตัวแปรสุ่มอีกด้วย ได้แก่ ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่ม (probability function) ฟังก์ชันความน่าจะเป็นสะสม (cumulative probability function) ค่าเฉลี่ย (mean) ค่าความแปรปรวน (variance) ค่าสัมประสิทธิ์ความเบ้ (skewness coefficient) ค่าสัมประสิทธิ์ความโด่ง (kurtosis coefficient) ฟังก์ชันโมเมนต์เวียนบังเกิด (moment generation function) และฟังก์ชันลักษณะเฉพาะ (characteristic function)

สำหรับการศึกษาในครั้งนี้เป็นการสร้างการแจกแจงความน่าจะเป็นของตัวแปรสุ่มขึ้นมาใหม่โดยการผสมกันระหว่างการแจกแจงทวินามลบ (Negative Binomial Distribution) กับการแจกแจงลินเลย์วางนัยทั่วไป (Generalized Lindley Distribution) โดยจะเรียกการแจกแจงแบบใหม่นี้ว่า การแจกแจงทวินามลบ - ลินเลย์วางนัยทั่วไป (Negative Binomial – Generalized Lindley Distribution) นับว่าเป็นการสร้างการแจกแจงแบบไม่ต่อเนื่องที่สำคัญให้เกิดขึ้นในวงวิชาการคณิตศาสตร์ได้อีกการแจกแจงหนึ่ง

วัตถุประสงค์ของการวิจัย

1. เพื่อสร้างฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มทวินามลบ - ลินเลย์วางนัยทั่วไป
2. เพื่อศึกษาคุณสมบัติเชิงความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงทวินามลบ-ลินเลย์วางนัยทั่วไป
3. เพื่อศึกษาและเปรียบเทียบการประมาณค่าพารามิเตอร์ของการแจกแจงแบบทวินามลบ-ลินเลย์วางนัย

ทั่วไปโดยใช้วิธีภาวะน่าจะเป็นสูงสุด (maximum likelihoodmethod) และวิธีโมเมนต์(moment method)

4. เพื่อศึกษาการแจกแจงแบบทวินามลบ-ลินเลย์วางนัยทั่วไปในเชิงประยุกต์กับข้อมูลไม่ต่อเนื่อง เช่นจำนวนผู้เสียชีวิตจากการเกิดอุบัติเหตุ จำนวนผู้ป่วย เป็นต้น

ระเบียบวิธีวิจัย

ประชากร คือ การแจกแจงความน่าจะเป็นแบบไม่ต่อเนื่อง กลุ่มตัวอย่าง คือ การแจกแจงความน่าจะเป็นของตัวแปรสุ่มปัวซง การแจกแจงความน่าจะเป็นของตัวแปรสุ่มทวินามลบ การแจกแจงความน่าจะเป็นของตัวแปรสุ่มทวินามลบ-ลินเลย์ และการแจกแจงความน่าจะเป็นของตัวแปรสุ่มทวินามลบ-ลินเลย์วางนัยทั่วไป เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูลครั้งนี้ใช้ คือโปรแกรมสำเร็จรูป R ในการจำลองสถานการณ์และวิเคราะห์ข้อมูล

ขั้นตอนในการดำเนินงานวิจัย

1. ศึกษาสาระสนเทศต่าง ๆ ที่เกี่ยวข้องกับการแจกแจงความน่าจะเป็นของตัวแปรสุ่ม

2. ทำการสร้างฟังก์ชันการแจกแจงทวินามลบ -ลินเลย์วางนัยทั่วไป พร้อมทั้งประมาณค่าพารามิเตอร์ และประยุกต์กับข้อมูลจริง โดยใช้โปรแกรม R และสร้างโปรแกรมขึ้นมาเองมีดังนี้

ตอนที่ 1. วิเคราะห์คุณสมบัติเชิงความน่าจะเป็นที่เกี่ยวข้องกับตัวแปรสุ่มที่มีการแจกแจงแบบทวินามลบ-ลินเลย์วางนัยทั่วไป ดังต่อไปนี้

- ฟังก์ชันความน่าจะเป็น
- ฟังก์ชันความน่าจะเป็นสะสม
- คุณสมบัติสถิติเชิงคณิตศาสตร์ซึ่งประกอบไปด้วยค่าเฉลี่ยความแปรปรวน ค่าสัมประสิทธิ์ความเบ้และค่าสัมประสิทธิ์ความโด่ง เป็นต้น
- ฟังก์ชันโมเมนต์เวียนบังเกิด (moment generation function) และฟังก์ชันลักษณะเฉพาะ
- การสร้างค่าตัวแปรสุ่มทวินามลบ-ลินเลย์วางนัยทั่วไป
- การทดสอบภาวะรูปสนิทดี โดยใช้สถิติทดสอบไคสแควร์

ตอนที่ 2. ประมาณค่าพารามิเตอร์โดยใช้วิธีการประมาณค่าพารามิเตอร์ 2 วิธีคือวิธีภาวะน่าจะเป็นสูงสุดและวิธีโมเมนต์

- จำลองข้อมูลตามสถานการณ์ที่กำหนดโดยใช้โปรแกรม R ของแต่ละขนาดตัวอย่างที่ใช้ในการศึกษาเท่ากับ 50,100,200,100 และ 500 โดยทำการจำลองซ้ำ 500 ครั้งในแต่ละสถานการณ์ที่กำหนดและประมาณค่าค่าพารามิเตอร์โดยวิธีภาวะน่าจะเป็นสูงสุด และวิธีโมเมนต์ของแต่ละสถานการณ์
- เปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าพารามิเตอร์โดยพิจารณาจากค่าเฉลี่ยของความคลาดเคลื่อนกำลังสองระหว่างค่าประมาณกับค่าจริง (Mean Square Error:MSE)

ตอนที่ 3. วิเคราะห์ข้อมูลจริงกับการแจกแจงทวินามลบ-ลินเลย์วางนัยทั่วไป โดยข้อมูลจริงที่นำมาใช้ในประยุกต์ จะทำการวิเคราะห์ดังนี้

- การทดสอบภาวะสารูปสนิทดี
- การประมาณค่าพารามิเตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุดและวิธีโมเมนต์

ผลการวิจัย

ผลการวิเคราะห์การประมาณค่าพารามิเตอร์ของการแจกแจงทวินามลบ -ลินเลย์ทั่วไปโดยวิธีการประมาณค่าพารามิเตอร์ 2 วิธีคือวิธีภาวะน่าจะเป็นสูงสุด และวิธีโมเมนต์ ในการศึกษาครั้งนี้ได้จำลองข้อมูลตามสถานการณ์ที่กำหนดโดยใช้โปรแกรม R ของแต่ละขนาดตัวอย่างที่ใช้ในการศึกษาเท่ากับ 50,100,200 และ 500 โดยทำการจำลองซ้ำ 500 ครั้ง และกำหนดค่าพารามิเตอร์ดังนี้

1) $r=1, \theta=5, \alpha=2, \gamma=4$
2) $r=3, \theta=6, \alpha=2, \gamma=5$
3) $r=5, \theta=8, \alpha=1, \gamma=9$

ผลการประมาณค่าพารามิเตอร์ในแต่ละสถานการณ์ที่กำหนด ได้ค่าประมาณพารามิเตอร์โดยวิธีภาวะน่าจะเป็นสูงสุด และวิธีโมเมนต์ และได้เปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าพารามิเตอร์โดยพิจารณาจากค่าเฉลี่ยของความคลาดเคลื่อนกำลังสองระหว่างค่าประมาณกับค่าจริง (Mean Square Error:MSE) ได้ผลดังนี้

1) ค่าเฉลี่ยของค่าประมาณ และค่า MSE ของวิธี MLE และ MM ในการประมาณค่าพารามิเตอร์ $r, \theta, \alpha, \gamma$ ของการแจกแจงทวินามลบ -ลินเลย์ทั่วไป เมื่อกำหนด

$r=1,\theta=5,\alpha=2,\gamma=4$ พบว่า การประมาณค่าพารามิเตอร์ r ด้วยวิธี MLE จะให้ค่าเฉลี่ยของค่าประมาณพารามิเตอร์ที่ใกล้เคียงกับพารามิเตอร์ที่กำหนดมากกว่าวิธี MM เมื่อมีขนาดตัวอย่างที่ 50 และ 100 แต่เมื่อมีขนาดตัวอย่างเท่ากับ 200 และ 500 พบว่า วิธี MM จะให้ค่าเฉลี่ยของค่าประมาณพารามิเตอร์ที่ใกล้เคียงกับพารามิเตอร์ที่กำหนดมากกว่าวิธี MLE เมื่อพิจารณาถึงค่า MSE จะพบว่าวิธี MLE ให้ค่า MSE ต่ำกว่าวิธี MM เมื่อมีขนาดตัวอย่างที่ 50 และ 100 และ วิธี MM ให้ค่า MSE ต่ำกว่าวิธี MLE เมื่อมีขนาดตัวอย่างที่ 200 และ 500

- การประมาณค่าพารามิเตอร์ θ ด้วยวิธี MLE จะให้ค่าเฉลี่ยของค่าประมาณพารามิเตอร์ที่ใกล้เคียงกับพารามิเตอร์ที่กำหนดมากกว่าวิธี MM ทุกขนาดตัวอย่าง เมื่อพิจารณาถึงค่า MSE จะพบว่าวิธี MLE ให้ค่า MSE ต่ำกว่าวิธี MM ทุกขนาดตัวอย่าง

- การประมาณค่าพารามิเตอร์ α ด้วยวิธี MM จะให้ค่าเฉลี่ยของค่าประมาณพารามิเตอร์ที่ใกล้เคียงกับพารามิเตอร์ที่กำหนดมากกว่าวิธี MLE ทุกขนาดตัวอย่าง เมื่อพิจารณาถึงค่า MSE จะพบว่าวิธี MM ให้ค่า MSE ต่ำกว่าวิธี MLE ทุกขนาดตัวอย่าง

- การประมาณค่าพารามิเตอร์ γ ด้วยวิธี MLE จะให้ค่าเฉลี่ยของค่าประมาณพารามิเตอร์ที่ใกล้เคียงกับพารามิเตอร์ที่กำหนดมากกว่าวิธี MM ทุกขนาดตัวอย่าง เมื่อพิจารณาถึงค่า MSE จะพบว่าวิธี MLE ให้ค่า MSE ต่ำกว่าวิธี MM ทุกขนาดตัวอย่าง

2) ค่าเฉลี่ยของค่าประมาณ และค่า MSE ของวิธี MLE และ MM ในการประมาณค่าพารามิเตอร์ r,θ,α,γ ของการแจกแจงทวินามลบ -ลินเลย์ทั่วไป เมื่อกำหนด $r=3,\theta=6,\alpha=2,\gamma=5$ พบว่า

การประมาณค่าพารามิเตอร์ r ด้วยวิธี MM จะให้ค่าเฉลี่ยของค่าประมาณพารามิเตอร์ที่ใกล้เคียงกับพารามิเตอร์ที่กำหนดมากกว่าวิธี MLE เมื่อมีขนาดตัวอย่างที่ 50 และ 100 แต่เมื่อมีขนาดตัวอย่างเท่ากับ 200 และ 500 พบว่า วิธี MLE จะให้ค่าเฉลี่ยของค่าประมาณพารามิเตอร์ที่ใกล้เคียงกับพารามิเตอร์ที่กำหนดมากกว่าวิธี MM เมื่อพิจารณาถึงค่า MSE จะพบว่าวิธี MM ให้ค่า MSE ต่ำกว่าวิธี MLE เมื่อมีขนาดตัวอย่างที่ 50 และ 100 และ วิธี MLE ให้ค่า MSE ต่ำกว่าวิธี MM เมื่อมีขนาดตัวอย่างที่ 200 และ 500

การประมาณค่าพารามิเตอร์ θ ด้วยวิธี MM จะให้ค่าเฉลี่ยของค่าประมาณพารามิเตอร์ที่ใกล้เคียงกับพารามิเตอร์ที่กำหนดมากกว่าวิธี MLE ทุกขนาดตัวอย่าง เมื่อพิจารณาถึงค่า MSE จะพบว่าวิธี MM ให้ค่า MSE ต่ำกว่าวิธี MLE ทุกขนาดตัวอย่าง ส่วนวิธี MLE จะให้ค่า MSE ที่มีแนวโน้มลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้น

การประมาณค่าพารามิเตอร์ α ด้วยวิธี MM จะให้ค่าเฉลี่ยของค่าประมาณพารามิเตอร์ที่ใกล้เคียงกับพารามิเตอร์ที่กำหนดมากกว่าวิธี MLE ทุกขนาดตัวอย่าง เมื่อพิจารณาถึงค่า MSE จะพบว่าวิธี MM ให้ค่า MSE ต่ำกว่าวิธี MLE ทุกขนาดตัวอย่าง ส่วนวิธี MLE จะให้ค่า MSE ที่มีแนวโน้มลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้น

การประมาณค่าพารามิเตอร์ γ ด้วยวิธี MM จะให้ค่าเฉลี่ยของค่าประมาณพารามิเตอร์ที่ใกล้เคียงกับพารามิเตอร์ที่กำหนดมากกว่าวิธี MLE ทุกขนาดตัวอย่าง เมื่อพิจารณาถึงค่า MSE จะพบว่าวิธี MM ให้ค่า MSE ต่ำกว่าวิธี MLE ทุกขนาดตัวอย่าง ส่วนวิธี MLE จะให้ค่า MSE ที่มีแนวโน้มลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้น

3) ค่าเฉลี่ยของค่าประมาณ และค่า MSE ของวิธี MLE และ MM ในการประมาณค่าพารามิเตอร์ r,θ,α,γ ของการแจกแจงทวินามลบ -ลินเลย์ทั่วไป เมื่อกำหนด $r=5,\theta=8,\alpha=1,\gamma=9$ พบว่า

การประมาณค่าพารามิเตอร์ r ด้วยวิธี MM จะให้ค่าเฉลี่ยของค่าประมาณพารามิเตอร์ที่ใกล้เคียงกับพารามิเตอร์ที่กำหนดมากกว่าวิธี MLE ทุกขนาดตัวอย่าง เมื่อพิจารณาถึงค่า MSE จะพบว่าวิธี MM ให้ค่า MSE ต่ำกว่าวิธี MLE ทุกขนาดตัวอย่าง ส่วนวิธี MLE จะให้ค่า MSE ที่มีแนวโน้มลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้น

การประมาณค่าพารามิเตอร์ θ ด้วยวิธี MM จะให้ค่าเฉลี่ยของค่าประมาณพารามิเตอร์ที่ใกล้เคียงกับพารามิเตอร์ที่กำหนดมากกว่าวิธี MLE ทุกขนาดตัวอย่าง เมื่อพิจารณาถึงค่า MSE จะพบว่าวิธี MM ให้ค่า MSE ต่ำกว่าวิธี MLE ทุกขนาดตัวอย่าง ส่วนวิธี MLE จะให้ค่า MSE ที่มีแนวโน้มลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้น

การประมาณค่าพารามิเตอร์ α ด้วยวิธี MM จะให้ค่าเฉลี่ยของค่าประมาณพารามิเตอร์ที่ใกล้เคียงกับพารามิเตอร์ที่กำหนดมากกว่าวิธี MLE ทุกขนาดตัวอย่าง เมื่อพิจารณาถึงค่า MSE จะพบว่าวิธี MM ให้ค่า MSE ต่ำกว่าวิธี MLE ทุกขนาดตัวอย่าง ส่วนวิธี MLE จะให้ค่า MSE ที่มีแนวโน้มลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้น

การประมาณค่าพารามิเตอร์ γ ด้วยวิธี MM จะให้ค่าเฉลี่ยของค่าประมาณพารามิเตอร์ที่ใกล้เคียงกับพารามิเตอร์ที่กำหนดมากกว่าวิธี MLE ทุกขนาดตัวอย่าง เมื่อพิจารณาถึงค่า MSE จะพบว่าวิธี MM ให้ค่า MSE ต่ำกว่าวิธี MLE ทุกขนาดตัวอย่าง ส่วนวิธี MLE จะให้ค่า MSE ที่มีแนวโน้มลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้น

การประยุกต์ข้อมูลจริงกับการแจกแจงทวินามลบ-ลินเลย์ทั่วไป ได้นำเสนอการประยุกต์ใช้ข้อมูลอุบัติเหตุบนถนนทางหลวงของจังหวัดนครปฐม ปี 2555 (กรมขนส่งทางบก, 2556) ซึ่งได้นำมาจากการเก็บรวบรวมข้อมูลจำนวนผู้เสียชีวิตจากการเกิดอุบัติเหตุบนถนนทางหลวงของจังหวัดนครปฐม พบว่า ข้อมูลอุบัติเหตุบนถนนทางหลวงของจังหวัดนครปฐม ปี 2555 โดยทดสอบภาวะสารรูปสนิทดี โดยใช้ตัวสถิติทดสอบไคสแควร์ พบว่า ข้อมูลชุดนี้มีการแจกแจงทวินามลบ-ลินเลย์ทั่วไป ที่ระดับนัยสำคัญ 0.05

สรุปและอภิปรายผล

ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบทวินามลบ-ลินเลย์วางนัยทั่วไปเป็นดังนี้

$$f(x;r,\alpha,\theta,\gamma)=\binom{r+x-1}{x}\sum_{j=0}^{x}\binom{x}{j}(-1)^{j}\left(\frac{\theta}{\theta+r+j}\right)^{\alpha+1}\frac{\theta+r+j+\gamma}{\theta+\gamma},$$

x=0,1,2,..., r, α , θ , γ > 0

คุณสมบัติสถิติเชิงคณิตศาสตร์

1) แฟกทอเรียลโมเมนต์ (Factorial Moments) อันดับที่ k

$$\mu_{[k]}(\mathrm{X})=\frac{\Gamma(r+k)}{\Gamma(r)}\sum_{j=0}^{k}\binom{k}{j}(-1)^{j}\left(\frac{\theta}{\theta-(k-j)}\right)^{\alpha+1}\frac{\theta-(k-j)+\gamma}{\theta+\gamma},k=1,2,\ldots,\mathrm{r},\theta,\alpha,\gamma>0$$

2) โมเมนต์อันดับที่ 1 ถึง 4

โมเมนต์อันดับที่ 1 (First Moment)

$$E(\mathrm{X})=r\frac{\theta^{\alpha+1}}{\theta+\gamma}\left[\frac{\theta-1+\gamma}{(\theta-1)^{\alpha+1}}-\frac{\theta+\gamma}{\theta^{\alpha+1}}\right]$$

โดยที่ $\theta>1$

โมเมนต์อันดับที่ 2 (Second Moment)

$$E(\mathrm{X}^2)=r(r+1)\frac{\theta^{\alpha+1}}{\theta+\gamma}\left[\frac{\theta-2+\gamma}{(\theta-2)^{\alpha+1}}-\frac{2(\theta-1+\gamma)}{(\theta-1)^{\alpha+1}}+\frac{\theta+\gamma}{\theta^{\alpha+1}}\right]$$

โดยที่ $\theta>2$

โมเมนต์อันดับที่ 3 (Third Moment)

$$E(\mathrm{X}^3)=r(r+1)(\mathrm{r}+2)\frac{\theta^{\alpha+1}}{\theta+\gamma}\left[\frac{\theta-3+\gamma}{(\theta-3)^{\alpha+1}}-\frac{3(\theta-2+\gamma)}{(\theta-2)^{\alpha+1}}+\frac{3(\theta-1+\gamma)}{(\theta-1)^{\alpha+1}}-\frac{\theta+\gamma}{\theta^{\alpha+1}}\right]$$

โดยที่ $\theta>3$

โมเมนต์อันดับที่ 4 (Fourth Moment)

$$E(\mathrm{X}^4)=r(r+1)(\mathrm{r}+2)(\mathrm{r}+3)\frac{\theta^{\alpha+1}}{\theta+\gamma}\left[\frac{\theta-4+\gamma}{(\theta-4)^{\alpha+1}}-\frac{4(\theta-3+\gamma)}{(\theta-3)^{\alpha+1}}+\frac{6(\theta-2+\gamma)}{(\theta-2)^{\alpha+1}}-\frac{4(\theta-1+\gamma)}{(\theta-1)^{\alpha+1}}-\frac{\theta+\gamma}{\theta^{\alpha+1}}\right]$$

โดยที่ $\theta>4$

3) ค่าเฉลี่ย (Mean)

$$Mean=E(\mathrm{X})=r\frac{\theta^{\alpha+1}}{\theta+\gamma}\left[\frac{\theta-1+\gamma}{(\theta-1)^{\alpha+1}}-\frac{\theta+\gamma}{\theta^{\alpha+1}}\right]$$

โดยที่ $\theta>1$

4) ความแปรปรวน (Variance)

$$Var(X)=\frac{\delta_3}{\delta_1}\left(r^2+r\right)-\frac{\delta_2}{\delta_1}\left(2r-\frac{r^2\delta_2}{\delta_1}\right)+r$$

โดยที่ $\delta_1=\frac{\theta+\gamma}{\theta^{\alpha+1}}$, $\delta_2=\frac{(\theta-1+\gamma)}{(\theta-1)^{\alpha+1}}$,

$\delta_3=\frac{\theta-2+\gamma}{(\theta-2)^{\alpha+1}}$ และ $\theta>2$

5) ค่าสัมประสิทธิ์ความเบ้ (Coefficient of Skewness)

$$Skewness(X)=\frac{E(X^3)-3E(X)Var(X)-[E(X)]^3}{[Var(X)]^{3/2}}$$

โดยที่ $\theta>3$

6) ค่าสัมประสิทธิ์ความโด่ง (Coefficient of Kurtosis)

$$Kurtosis(X)=\frac{E(X^4)-4E(X^3)E(X)+6E(X^2)[E(X)]^2-3[E(X)]^4}{[Var(X)]^2}$$

โดยที่ $\theta>4$

7) ฟังก์ชันโมเมนต์เวียนบังเกิด (Moment Generation Function)

$$M_X(t)=\sum_{x=0}^{\infty}e^{tx}\binom{r+x-1}{x}\sum_{j=0}^{x}\left[\binom{x}{j}(-1)^{j}\left(\frac{\theta}{\theta+r+j}\right)^{\alpha+1}\frac{\theta+r+j+\gamma}{\theta+\gamma}\right]$$

, r, α , θ , γ > 0

8) ฟังก์ชันลักษณะเฉพาะ (Characteristic Function)

$$\varphi_X(t)=\sum_{x=0}^{\infty}e^{itx}\binom{r+x-1}{x}\sum_{j=0}^{x}\left[\binom{x}{j}(-1)^j\left(\frac{\theta}{\theta+r+j}\right)^{\alpha+1}\frac{\theta+r+j+\gamma}{\theta+\gamma}\right],$$

$r,\alpha,\theta,\gamma>0$

จากการจำลองสถานการณ์เพื่อทำการประมาณค่าพารามิเตอร์ของการแจกแจงทวินามลบ -ลินเลย์ทั่วไปโดยวิธิการประมาณค่าพารามิเตอร์ 2 วิธีคือวิธีภาวะน่าจะเป็นสูงสุด และวิธีโมเมนต์ ในการศึกษาครั้งนี้ได้จำลองข้อมูลตามสถานการณ์ที่กำหนดโดยใช้ขนาดตัวอย่างเท่ากับ 50,100,200 และ 500 โดยทำการจำลองซ้ำ 500 ครั้ง พบว่าการประมาณค่าพารามิเตอร์ r,θ,α,γ ด้วยวิธีโมเมนต์ จะให้ค่าเฉลี่ยของค่าประมาณพารามิเตอร์ที่ใกล้เคียงกับพารามิเตอร์ที่กำหนดมากกว่าวิธี MLE ทุกขนาดตัวอย่างเกือบทุกกรณี และเมื่อพิจารณาถึงค่า MSE จะพบว่าวิธี MM ให้ค่า MSE ต่ำกว่าวิธี MLE ทุกขนาดตัวอย่างเกือบทุกกรณี

การประยุกต์ข้อมูลจริง โดยเป็นข้อมูลอุบัติเหตุบนถนนทางหลวงของจังหวัดนครปฐม ปี 2555 โดยทดสอบภาวะสารรูปสนิทดี โดยใช้ตัวสถิติทดสอบไคสแควร์ พบว่าข้อมูลชุดนี้มี การแจกแจงทวินามลบ -ลินเลย์ทั่วไป ที่ระดับนัยสำคัญ 0.05

ข้อเสนอแนะ

จากการศึกษาพบว่าเมื่อข้อมูลที่นำมาศึกษามีการแจกแจงทวินามลบ-ลินเลย์ทั่วไป วิธีการประมาณค่าพารามิเตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุด และวิธีโมเมนต์ โดยพบว่าวิธีโมเมนต์ให้ประสิทธิภาพ ดีกว่าวิธีภาวะน่าจะเป็นสูงสุด เกือบทุกกรณี แต่วิธีโมเมนต์มีข้อเสียคือ ค่าประมาณพารามิเตอร์ของ θ จะหาค่าได้เมื่อ $\theta>4$ ดังนั้นบางกรณีควรใช้วิธีภาวะน่าจะเป็นสูงสุดใน การประมาณค่าพารามิเตอร์ ของการแจกแจงทวินามลบ-ลินเลย์ทั่วไป

เอกสารอ้างอิง

จิรัชย์ สุขะเกตุ. (2548). **ความน่าจะเป็น และทฤษฎีสถิติเบื้องต้น.** กรุงเทพฯ, สำนักพิมพ์มหาวิทยาลัยเกษตรศาสตร์.

ประชุม สุวัตถี. (2545). **ทฤษฎีการอนุมานเชิงสถิติ.** (พิมพ์ครั้งที่2). กรงเทพฯ. โครงการส่งเสริมเอกสารวิชาการ สถาบันบัณฑิตพัฒนบริหารศาสตร์.

วิชัย สุรเชิดเกียรติ. (2538). การพัฒนาโปรแกรมย่อยฟังก์ชันสำหรับสร้างค่าตัวแปรสุ่มกรุงเทพฯ: สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ.

อารียา สุดสุข. (2555). **วิทยานิพนธ์ ตัวแปรสุ่มเบตาเอกซ์โปเนนเชียลวางในทั่วไป และการประยุกต์ทางด้านความเชื่อถือได้**. บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์.

Felle, W.(1968). **An Introduction to Probability Theory and Its Applications.** (vol.1) New York : Wiley.

Ghitany, M.E., B. Atieh and S. Nadarajah, 2008b. **Lindley distribution and its application.** Math. Comput. Simulat., 78: 493-506.

Pudprommarat, C., W. Bodhisuwan and P. Zeephongsekul, (2012). **A New Mix Negative Binomial Distribution**. journal of Applied Science. 12 (17):1853-1858.

Rainer, W. (2000). **Econometric Analysis of Count Data**. 3rdEdn., Springer-Verlag, Berlin, Germany.

Simon, L.J., (1961), **Fitting negative binomial distributions by the method of maximum likelihood.** Proc. Casual. Actuar. Soc., 48: 45-53.

Wang, Z., (2011). **One mixed negative binomial distribution with application.** Journal of Statistical Planning and Inference, 141(3) : 153-1160.

Zhaoliang Wang. **(2011) One Mixed Negative Binomial Distribution with application.** Journal of statistical Planning and Inference,vol.141(3):1153-1160

Zakerzadeh, H and A. Dolati (2009), **Generralized Lindley Distribution.** Journal of mathematical Extension, vol.3 No.2 :13-25.