KKU Res. J. 2012; 17(2):287-292
http : resjournal.kku.ac.th

การประมาณค่าสัมประสิทธิ์การแปรผันแบบใหม่ของข้อมูลที่อยู่ในรูปอัตราส่วนที่สัมพันธ์กัน
The New Estimation of Coefficient of Variation of Ratio Correlated Data

จินดา สวัสดิ์ทวี[1], สุชาดา กรเพชรปาณี[2]*

Jinda Sawattawee[1], Suchada Kornpetpanee[2]*

[1]นิสิตปริญญาเอก สาขาการวิจัยและสถิติทางวิทยาการปัญญา วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา
[2]วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา
**Correspondent author: Sjinny033@hotmail.com*

Received December 7,2011
Accepted April 24,2012

บทคัดย่อ

การศึกษาครั้งนี้มีวัตถุประสงค์เพื่อเสนอวิธีการประมาณค่าสัมประสิทธิ์การแปรผันแบบใหม่ของข้อมูลที่อยู่ในรูปอัตราส่วนที่สัมพันธ์กัน และได้เปรียบเทียบวิธีการใหม่นี้กับวิธีการคำนวณแบบดั้งเดิม โดยการเปรียบเทียบค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสองของค่าสัมประสิทธิ์การแปรผัน (Root Mean Square Error: RMSE) ของค่าประมาณ ซึ่งกำหนดขนาดตัวอย่าง เท่ากับ 5, 10, 15, 30, 50 และ 100 และกำหนดค่าสัมประสิทธิ์สหสัมพันธ์ (coefficient correlation values : ρ) เท่ากับ ±0.3, ±0.5, ±0.7 และ ±0.9 การศึกษาครั้งนี้ใช้วิธีจำลองข้อมูลแบบมอนติคาร์โล และทำซ้ำ ๆ กัน 1000 ครั้ง ในแต่ละสถานการณ์ ผลการศึกษาสรุปได้ดังนี้ วิธีการประมาณค่าสัมประสิทธิ์การแปรผันแบบใหม่ของข้อมูลที่อยู่ในรูปอัตราส่วน จะให้ค่า RMSE ต่ำกว่าวิธีการคำนวณแบบดั้งเดิมทุกเงื่อนไข

Abstract

The objectives of this study are to propose the new estimation coefficient of variation method of ratio correlated data and to compare with original coefficient of variation method. The research was considered by Root Mean Square Error (RMSE). The comparisons were done by using sample sizes (n) equal to 5, 10, 15, 30, 50, and 100 whereas coefficient correlation values (ρ) are ±0.3, ±0.5, ±0.7 and ±0.9. This study used the Monte Carlo Simulation method and the experiment was repeated 1000 times for each condition. The results showed that the new methods to estimate the coefficient variation generates the RMSE value lower than the original methods in every condition.

คำสำคัญ: การประมาณค่าสัมประสิทธิ์การแปรผัน, ข้อมูลในรูปอัตราส่วนที่สัมพันธ์

Keywords: the new estimation coefficient of variation, ratio correlated data

1. บทนำ

ข้อมูลที่อยู่ในรูปอัตราส่วน (Ratio Data) ที่พบได้บ่อยในปัจจุบัน เช่น อัตราดอกเบี้ย อัตราผลกำไรตอบแทน อัตราการซื้อขาย เป็นต้น ซึ่งอัตราเหล่าเหล่านี้จะประกอบด้วยข้อมูล 2 ชุด หรือ ตัวแปร 2 ตัว ที่อยู่ในรูปอัตราส่วน มีงานวิจัยมากมายที่กล่าวถึงอัตราส่วนของสองตัวแปร โดยเฉพาะอย่างยิ่งงานวิจัยทางด้านการวิเคราะห์ทางการเกษตรและทางชีววิทยาที่เกี่ยวข้องกับอัตราส่วนของตัวแปรสุ่มปกติ 2 ตัว เช่น การแยกแยะอัตราส่วนของน้ำหนักขององค์ประกอบของพืชกับส่วนทั้งหมดของพืช หรืออัตราส่วนของพารามิเตอร์สองตัวในตัวแบบ เป็นต้น (1) นอกจากนี้ ในเรื่องของการแจกแจงอัตราส่วน x/y เป็นปัญหาหนึ่งที่ได้รับความสนใจในด้านชีววิทยาและวิทยาศาสตร์กายภาพ ได้นำเทคนิคทางคณิตศาสตร์ไปใช้ในการทดสอบทฤษฎีทางเศรษฐศาสตร์ เช่น อัตราส่วนของการสืบทอดทางพันธุกรรมที่เป็นไปตามกฎของเมนเดลในยีน มวลที่นำไปสู่อัตราส่วนของพลังงานนิวเคลียสในฟิสิกส์ เป้าหมายในการควบคุมการตกตะกอนในอุตุนิยมวิทยา หรืออัตราส่วนของพัสดุคงคลังในด้านเศรษฐศาสตร์ เป็นต้น (2) และจากการศึกษางานวิจัยที่ผ่านมาก่อนหน้านี้ พบว่า มีนักวิจัยหลายท่านได้ให้ความสนใจเกี่ยวกับอัตราส่วนของ x/y เช่น Shoukri, M. M.และคณะ (3), Hayya, J. และคณะ (4), Fieller (5) และ Hinkley (6) เป็นต้น ซึ่งโดยส่วนใหญ่จะให้ความสนใจเฉพาะในกรณีที่ X, Y เป็นตัวแปรสุ่มที่อิสระต่อกันและมาจากกลุ่มประชากรที่คล้ายกัน แต่ในกรณีของตัวแปรสุ่ม X, Y ที่มีความสัมพันธ์กัน จะพบว่ามีงานวิจัยค่อนข้างน้อย (2)

การหาค่าสัมประสิทธิ์การแปรผัน พารามิเตอร์ที่สำคัญที่ช่วยในการคำนวณหาค่า คือ ค่าเฉลี่ย เท่ากับ $\mu_x = E(x)$ และค่าความแปรปรวน เท่ากับ $\sigma^2 = Var(x)$ ซึ่งโดยทั่วไปสามารถคำนวณหาได้จาก

$$CV(X) = \frac{\sigma_X}{\mu_X} = \frac{\sqrt{Var(x)}}{E(x)} \quad [1]$$

ต่อมามีนักวิจัยหลายท่านได้นำสูตรสัมประสิทธิ์การแปรผันมาพัฒนาเพื่อให้เหมาะสมกับข้อมูลในงานวิจัย ในปี 2005 Kim, Y.Y. และคณะ (7) ได้พัฒนาสูตรสัมประสิทธิ์การแปรผันเพื่อใช้ในการสำรวจสาเหตุของวัฏจักรการแปรผันในการฉีดเชื้อเพลิงไฮโดรเจนเข้าไปในเครื่องจักร ซึ่งสูตรสัมประสิทธิ์การแปรผันแบบใหม่ของเขาที่ได้ คือ

$$CV_{cp} = \frac{\left[\frac{1}{N}\sum_{i=1}^{N}\{cp(i)-\overline{cp}\}^2\right]^{\frac{1}{2}}}{\frac{1}{N}\sum_{i=1}^{N}cp(i)} \quad [2]$$

ต่อมา ในปี 2007 Holmes, D. T. & Buhr, K. A (8) ได้พัฒนาสูตรสัมประสิทธิ์การแปรผันขึ้นมาเพื่อใช้ในการคำนวณเชิงปริมาณของผลลัพธ์ที่ได้จากห้องปฏิบัติ

การซึ่งอยู่ในรูปของอัตราส่วน $R = \frac{X}{Y}$ โดยที่ X, Y เป็นตัวแปรสุ่มแบบ Gaussian ที่อิสระต่อกัน ซึ่งสูตรที่ได้จะอยู่ในรูป

$$CV_R \cong \sqrt{CV_X^2 + CV_Y^2} \quad [3]$$

ต่อมาในปี 2008 Pang, W. K. และคณะ (9) ได้พัฒนาสูตรสัมประสิทธิ์การแปรผันโดยใช้วิธีการจำลองขั้นพื้นฐานเพื่อหาอัตราตอบแทนเงินปันผล ซึ่งรูปแบบที่เหมาะสมในการหาอัตราตอบแทนเงินปันผลคือการแจกแจงแบบเบตา ค่าสัมประสิทธิ์การแปรผัน (CV) ที่พัฒนาขึ้นมาใหม่ถูกใช้ในการวัดความเสถียรภาพของการให้อัตราผลตอบแทนเงินปันผลของดัชนีฮั่งเส็งและดัชนีย่อยของตลาดหุ้นฮ่องกง (Hong Kong) และเปรียบเทียบผลลัพธ์ที่ได้กับวิธีดั้งเดิม สูตรสัมประสิทธิ์การแปรผันที่ได้ คือ

$$CV = \frac{\sqrt{Var(x)}}{E(x)} = \frac{\sqrt{\frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}}}{\frac{\alpha}{\alpha+\beta}} = \sqrt{\frac{\beta}{\alpha}} \times \frac{1}{\sqrt{\alpha+\beta+1}} \quad [4]$$

ในปีเดียวกัน Euser, A. M และคณะ (10) ได้พัฒนาสูตรขึ้นมาเพื่อเป็นตัวชี้วัดความสามารถในการทำซ้ำของตัวแปรที่เปลี่ยนรูปเป็นค่า log ค่าสัมประสิทธิ์การแปรผันแบบใหม่ที่ได้สามารถอธิบายความหมายได้โดยตรงจากความผิดพลาดมาตรฐานของการวัดของตัวแปรที่เปลี่ยนรูปเป็นค่า log สูตรที่ได้ คือ

$$CV_{inter} = 100\% \cdot \ln(10)\sqrt{(\hat{\sigma}_o^2 + \hat{\sigma}_E^2)} \quad [5]$$

ในปี 2010 Cox, J. C. & Sadiraj, V. (11) ได้พัฒนาสูตรสัมประสิทธิ์การแปรผันขึ้นมาเพื่อเป็นเกณฑ์สำหรับการตัดสินใจภายใต้ความเสี่ยง ในการวิเคราะห์ข้อมูลแบบวิธีการวิเคราะห์อภิมาน (meta-analysis) ที่เกี่ยวข้องกับทางเลือกของมนุษย์และสัตว์ที่มีความเสี่ยง โดยใช้สัมประสิทธิ์การแปรผันเป็นเกณฑ์ในการอธิบายการตัดสินใจภายใต้ความเสี่ยง ซึ่งสูตรสัมประสิทธิ์การแปรผันที่ได้ใหม่นี้จะใช้ในการทำนายโอกาสที่จะเกิดขึ้น และสูตรที่ได้คือ

$$CV = \sqrt{\frac{\frac{l^2}{p}+\frac{n^2 l^2}{1-p}}{(n+1)^2 l^2} - 1} = \frac{|p-1/(n+1)|}{\sqrt{p(1-p)}} \quad [6]$$

จากที่กล่าวมา ผู้วิจัยจึงมีแนวคิดที่จะพัฒนาสูตรสัมประสิทธิ์การแปรผันแบบใหม่เพื่อนำไปใช้กับ

ข้อมูลในรูปอัตราส่วน $Z = \frac{X}{Y}$ โดยที่ X, Y เป็นตัวแปรสุ่มที่สัมพันธ์กัน ดังนั้นในการศึกษาครั้งนี้ จึงได้เสนอวิธีการประมาณค่าสัมประสิทธิ์การแปรผันของ

ข้อมูลที่อยู่ในรูปอัตราส่วน $Z = \frac{X}{Y}$ โดยเปรียบเทียบค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง (RMSE) กับวิธีการคำนวณค่าสัมประสิทธิ์การแปรผันแบบดั้งเดิม

2. วิธีวิจัย

การศึกษาครั้งนี้ใช้วิธีการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โลด้วยเครื่องคอมพิวเตอร์ ด้วยโปรแกรมภาษาซี โดยใช้ Dev C++ เพื่อจำลองข้อมูลในการทดสอบการประมาณค่าสัมประสิทธิ์การแปรผันในสถานการณ์ต่าง ๆ และกำหนดให้เครื่องคอมพิวเตอร์ทำงานซ้ำ ๆ กัน 1000 ครั้ง ในแต่ละสถานการณ์ โดยการกำหนดขนาดตัวอย่าง (n) เท่ากับ 5, 10, 15, 30, 50 และ 100 และกำหนดค่าสัมประสิทธิ์สหสัมพันธ์ (coefficient correlation values: ρ) เท่ากับ ±0.3, ±0.5, ±0.7 และ ±0.9

ในการจำลองข้อมูล จะจำลองตัวแปรสุ่ม X, Y ที่กำหนดด้วยค่าสัมประสิทธิ์สหสัมพันธ์ (ρ) ในแต่ละสถานการณ์ เพื่อนำมาประมาณค่าพารามิเตอร์ $\mu(Z) = E(\frac{X}{Y})$ และ $\sigma^2(Z) = Var(\frac{X}{Y})$ เมื่อ $Z = \frac{X}{Y}$ ด้วยการใช้ทฤษฎีบททางคณิตศาสตร์ เช่น ทฤษฎีบทเทย์เลอร์ (Taylor's Theorem) ค่าคาดหวังทางคณิตศาสตร์ (Expected values) ค่าความแปรปรวน (Variance Values) และค่าความแปรปรวนร่วม (Covariance Values) เป็นต้น

ดังนั้น ค่า $E(\frac{X}{Y})$ และค่า $Var(\frac{X}{Y})$ สามารถประมาณได้ดังนี้

$$\mu(Z) = \mu(\frac{X}{Y}) = E(\frac{X}{Y}) = E(X)\cdot E(\frac{1}{Y}) + Cov(X, \frac{1}{Y}) \quad (12) \quad [7]$$

$$E(\frac{X}{Y}) = E(X)\cdot E(\frac{1}{Y}) + \rho_{X,\frac{1}{Y}} \sqrt{E(x^2) - E^2(X)} \cdot \sqrt{E(\frac{1}{Y^2}) - E^2(\frac{1}{Y})} \quad (13) \quad [8]$$

การประมาณ $\frac{1}{Y}$ และ $\frac{1}{Y^2}$ ด้วยการกระจายอนุกรมเทย์เลอร์ รอบจุด $Y = \mu_Y$ ด้วยอนุกรมดีกรี 5 (14) เพื่อหาค่า $E(\frac{1}{Y})$, $(E(\frac{1}{Y}))^2$ และ $E(\frac{1}{Y^2})$ ได้ดังนี้

$$\frac{1}{Y} = \frac{1}{\mu_Y} - \frac{1}{\mu_Y^2}(Y-\mu_Y) + \frac{1}{\mu_Y^3}\frac{(Y-\mu_Y)^2}{2!} - \ldots - \frac{1}{\mu_Y^6}\frac{(Y-\mu_Y)^5}{5!} + r_1$$

$$\frac{1}{Y^2} = \frac{1}{\mu_Y^2} - \frac{2}{\mu_Y^3}(Y-\mu_Y) + \frac{3}{\mu_Y^4}(Y-\mu_Y)^2 - \ldots - \frac{6}{\mu_Y^7}(Y-\mu_Y)^5 + r_2$$

โดย r_1 และ r_2 คือเศษเหลือของอนุกรมเทย์เลอร์ นำค่า $\frac{1}{Y}$ และ $\frac{1}{Y^2}$ ไปหาค่า $E(\frac{1}{Y})$, $(E(\frac{1}{Y}))^2$ และ $E(\frac{1}{Y^2})$ จะได้ว่า

$$E(\frac{1}{Y}) = \frac{1}{Y}(1 + CV_Y^2 + CV_Y^4) \quad [9]$$

$$E^2(\frac{1}{Y}) = \frac{1}{Y^2}(1 + 2CV_Y^2 + 3CV_Y^4 + 2CV_Y^6 + CV_Y^8) \quad [10]$$

$$E(\frac{1}{Y^2}) = \frac{1}{Y^2}(1 + 3CV_Y^2 + 5CV_Y^4) \quad [11]$$

นำค่า $E(\frac{1}{Y})$, $(E(\frac{1}{Y}))^2$ และ $E(\frac{1}{Y^2})$ แทนในสมการ [8] ดังนั้น

$$E(\frac{X}{Y}) = \mu_X\left(\frac{1}{\mu_Y}(1 + CV_Y^2 + CV_Y^4)\right) + \rho_{X,\frac{1}{Y}}\frac{\sigma_X}{\mu_Y}\sqrt{CV_Y^2 + 2CV_Y^4 - 2CV_Y^6 - CV_Y^8} \quad [12]$$

เมื่อ X, Y เป็นตัวแปรสุ่มที่สัมพันธ์กัน และในทำนองเดียวกัน

$$\sigma^2(Z) = \mathrm{Var}(Z) = \mathrm{Var}(\frac{X}{Y})$$

$$\mathrm{Var}(Z) = E(Z^2) - E^2(Z)$$

$$= E(X^2)\cdot E(\frac{1}{Y^2}) - \left(E(X)\cdot E(\frac{1}{Y}) + \rho_{X,\frac{1}{Y}} \sqrt{E(x^2) - E^2(X)} \cdot \sqrt{E(\frac{1}{Y^2}) - E^2(\frac{1}{Y})} \right)^2 \qquad [13]$$

นำค่า $E(\frac{1}{Y})$, $(E(\frac{1}{Y}))^2$ และ $E(\frac{1}{Y^2})$ แทนในสมการ [13] ดังนั้น

$$\begin{aligned}\mathrm{Var}(Z) = {} & \frac{\sigma_X^2}{\mu_Y^2}(1 + 2CV_Y^2 + 3CV_Y^4 + 2CV_Y^6 + CV_Y^8) + \\ & \frac{\mu_X^2}{\mu_Y^2}(CV_Y^2 + 2CV_Y^4 - 2CV_Y^6 - CV_Y^8) - \\ & 2\rho_{X,\frac{1}{Y}} \frac{\mu_X}{\mu_Y}\frac{\sigma_X}{\mu_Y}(1 + CV_Y^2 + CV_Y^4)\sqrt{CV_Y^2 + 2CV_Y^4 - 2CV_Y^6 - CV_Y^8}\end{aligned} \qquad [14]$$

เมื่อ X, Y เป็นตัวแปรสุ่มที่สัมพันธ์กัน

เมื่อได้ค่าประมาณ $E(\frac{X}{Y})$ และ $\mathrm{Var}(\frac{X}{Y})$ แล้ว นำค่าทั้งสองที่ได้มาประมาณค่าสัมประสิทธิ์การแปรผัน

ดังนั้น ค่าสัมประสิทธิ์การแปรผันแบบใหม่ของข้อมูลที่อยู่ในรูปอัตราส่วนที่สัมพันธ์กัน คือ

$$CV(Z) = \frac{\sqrt{\mathrm{Var}(Z)}}{E(Z)} = \frac{\sqrt{\mathrm{Var}(\frac{X}{Y})}}{E(\frac{X}{Y})}$$

$$= \frac{\frac{\mu_X}{\mu_Y}\left(\frac{\sigma_X}{\mu_X}(1 + CV_Y^2 + CV_Y^4) - \rho_{X,\frac{1}{Y}}\sqrt{CV_Y^2 + 2CV_Y^4 - 2CV_Y^6 - CV_Y^8}\right)}{\frac{\mu_X}{\mu_Y}(1 + CV_Y^2 + CV_Y^4) + \rho_{X,\frac{1}{Y}}\frac{\sigma_X}{\mu_X}\sqrt{CV_Y^2 + 2CV_Y^4 - 2CV_Y^6 - CV_Y^8}} \qquad [15]$$

$$= \frac{CV_X^2(1 + CV_Y^2 + CV_Y^4)^2 - CV_Y^2 - 2CV_Y^4 + 2CV_Y^6 + CV_Y^8}{(CV_X + 3CV_X CV_Y^2 + 5CV_X CV_Y^4) + (1 + CV_Y^2 + CV_Y^4 + CV_X^2 + CV_X^2 CV_Y^2 + CV_X^2 CV_Y^4)\rho_{X,\frac{1}{Y}}\sqrt{CV_Y^2 + 2CV_Y^4 - 2CV_Y^6 - CV_Y^8}}$$

จะพบว่าค่าสัมประสิทธิ์การแปรผันแบบใหม่ที่ได้จะมีความยุ่งยาก ดังนั้นในการประมาณค่าสัมประสิทธิ์การแปรผันแบบใหม่ จะพิจารณาเพื่อลดความยุ่งยากลงโดยการตัดเทอมที่มีค่าสัมประสิทธิ์การแปรผันที่มีดีกรีมากกว่าสองออกจากสมการการประมาณ เพื่อทำให้การจัดรูปแบบวิธีการประมาณมีความเหมาะสมและสะดวกต่อการนำไปใช้งาน ดังนั้น สูตรการประมาณค่าสัมประสิทธิ์การแปรผันแบบใหม่ คือ

$$CV(Z) = CV_X - \rho_{X,\frac{1}{Y}} CV_Y \qquad [16]$$

นำวิธีการประมาณค่าสัมประสิทธิ์การแปรผันแบบใหม่ที่ประมาณได้มาเปรียบเทียบกับวิธีการคำนวณค่าสัมประสิทธิ์การแปรผันแบบดั้งเดิม โดยการคำนวณค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง ค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง (RMSE) คำนวณจาก

$$RMSE = \sqrt{\frac{\sum_{i=}^{1000}(CV_i - CV_{mean})^2}{1000}}$$

เมื่อ cv_i แทน ค่าสัมประสิทธิ์การแปรผันแต่ละตัวที่คำนวณได้ในการทำซ้ำครั้งที่ i

cv_{mean} แทน ค่าเฉลี่ยของค่าสัมประสิทธิ์การแปรผันที่คำนวณได้ในการทำซ้ำ 1000 ครั้ง

i แทน จำนวนครั้งในการทำซ้ำ

3. ผลการวิจัยและอภิปราย

ในการนำเสนอผลการศึกษา เพื่อความสะดวกขอใช้สัญลักษณ์ต่อไปนี้แทนความหมายต่าง ๆ ดังนี้

CV1 หมายถึง ค่าสัมประสิทธิ์การแปรผันที่คำนวณด้วยวิธีการแบบดั้งเดิม ซึ่งคำนวณจาก

$$CV(Z) = \frac{\sqrt{\mathrm{Var}(Z)}}{E(Z)} = \frac{\sqrt{\mathrm{Var}(\frac{X}{Y})}}{E(\frac{X}{Y})}$$

CV2 หมายถึง ค่าสัมประสิทธิ์การแปรผันที่คำนวณด้วยสูตรการประมาณค่าแบบใหม่ที่ประมาณได้ ซึ่งคำนวณจาก

$$CV(Z) = CV_X - \rho_{X,\frac{1}{Y}} CV_Y$$

RMSE หมายถึง ค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง

* หมายถึง ค่า RMSE ที่ต่ำกว่า

3.1 ผลการวิเคราะห์ข้อมูล

ผลการเปรียบเทียบค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสองของค่าสัมประสิทธิ์การแปรผันที่ประมาณได้ นำเสนอดังตารางที่ 1 ซึ่งสรุปได้ว่าในการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โลด้วยเครื่องคอมพิวเตอร์ จะพบว่า ทุกสถานการณ์และทุกเงื่อนไขของขนาดตัวอย่างที่ต่างกันและค่าสัมประสิทธิ์สหสัมพันธ์ที่ต่างกัน วิธีการประมาณค่าสัมประสิทธิ์การแปรผันแบบใหม่ของข้อมูลที่อยู่ในรูปอัตราส่วนที่สัมพันธ์กันจะให้ค่า RMSE ต่ำกว่าวิธีการประมาณค่าสัมประสิทธิ์การแปรผันแบบดั้งเดิม

ตารางที่ 1 ค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสองของการประมาณค่าสัมประสิทธิ์การแปรผันของข้อมูลที่อยู่ในรูปอัตราส่วนที่สัมพันธ์กัน

ρ	n	RMSE	
		CV1	CV2
0.3	5	0.426058	0.299353*
	10	0.615285	0.20508*
	15	0.707545	0.161029*
	30	0.955147	0.112749*
	50	1.260413	0.089678*
	100	1.519585	0.069181*
-0.3	5	0.438596	0.301698*
	10	0.607228	0.200713*
	15	0.728165	0.152511*
	30	1.002618	0.111031*
	50	1.190441	0.090445*
	100	1.604756	0.068306*
0.5	5	0.414555	0.284236*
	10	0.58436	0.180194*
	15	0.722186	0.147032*
	30	0.975295	0.109829*
	50	1.190677	0.086874*
	100	1.408441	0.068004*
-0.5	5	0.434121	0.274757*
	10	0.607211	0.188978*
	15	0.727863	0.150032*
	30	0.985726	0.105975*
	50	1.188168	0.088237*
	100	1.47226	0.068507*
0.7	5	0.437103	0.260825*
	10	0.605413	0.170579*
	15	0.695625	0.134565*
	30	1.034193	0.111335*
	50	1.172676	0.090796*
	100	1.568373	0.070725*
-0.7	5	0.43813	0.257941*
	10	0.576154	0.172891*
	15	0.714169	0.140322*
	30	0.981242	0.109997*
	50	1.207258	0.088625*
	100	1.455079	0.07277*
0.9	5	0.402377	0.255571*
	10	0.545083	0.166809*
	15	0.733823	0.146142*
	30	0.971861	0.111759*
	50	1.183837	0.105794*
	100	1.006258	0.082086*
-0.9	5	0.430796	0.261143*
	10	0.577553	0.171237*
	15	0.684767	0.144657*
	30	0.996965	0.120038*
	50	1.113732	0.106658*
	100	1.277382	0.081634*

3.2 การอภิปรายผล

จากการทดสอบโดยการจำลองข้อมูลและทำซ้ำ 1000 ครั้ง สูตรการประมาณค่าสัมประสิทธิ์การแปรผันแบบใหม่ที่พัฒนาขึ้นสามารถใช้ได้ดีในกรณีที่ข้อมูลสองชุดหรือตัวแปรสองตัวมีความสัมพันธ์กัน โดยเฉพาะกรณีที่มีความสัมพันธ์กันสูง เนื่องจากวิธีการประมาณที่ได้มานี้อยู่ภายใต้ข้อตกลงว่าเป็นการหาค่าสัมประสิทธิ์การแปรผันแบบใหม่ของข้อมูลที่อยู่ในรูปอัตราส่วนที่สัมพันธ์กัน แต่วิธีการประมาณค่าสัมประสิทธิ์การแปรผันที่ได้มาใหม่ก็ยังมีความคลาดเคลื่อนในการประมาณเนื่องจากการได้มาของสูตรการประมาณได้มีการตัดเทอมของสัมประสิทธิ์การแปรผันที่มีดีกรีมากกว่าสองออกจากสูตรการประมาณ ทั้งนี้เพื่อให้มีรูปแบบที่เหมาะสมและเหมาะแก่การนำไปใช้

4. สรุป

วิธีการประมาณค่าของสัมประสิทธิ์การแปรผันแบบใหม่ของข้อมูลที่อยู่ในรูปอัตราส่วนที่สัมพันธ์กันนี้ จะให้ค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสองต่ำกว่าทุกกรณี ดังนั้น ในการประมาณค่าสัมประสิทธิ์การแปรผันของข้อมูลที่อยู่ในรูปอัตราส่วนที่สัมพันธ์กัน ควรใช้สูตรสัมประสิทธิ์การแปรผันแบบใหม่ ซึ่งมีสูตรดังนี้

$$CV(z) = CV_x - \rho_{x,\frac{1}{y}} CV_y$$

โดยที่ $CV(Z)$ คือ สัมประสิทธิ์การแปรผันของสองตัวแปรที่สัมพันธ์กันในรูปอัตราส่วนโดยที่ $Z = \frac{X}{Y}$

CV_X คือ สัมประสิทธิ์การแปรผันของตัวแปรสุ่ม X หรือตัวแปรเศษ

CV_Y คือ สัมประสิทธิ์การแปรผันของตัวแปรสุ่ม Y หรือตัวแปรส่วน

$\rho_{X,\frac{1}{Y}}$ คือ สัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรสุ่ม X และ $\frac{1}{Y}$

5. กิตติกรรมประกาศ

ผู้วิจัยขอขอบคุณคณะเทคโนโลยีและสิ่งแวดล้อม มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตภูเก็ต ที่สนับสนุนทุนการศึกษาและการวิจัย และขอขอบคุณอาจารย์วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพาที่ให้คำแนะนำในการวิจัยครั้งนี้

6. เอกสารอ้างอิง

(1) Shanmugalingam S. On the Analysis of the Ratio of Two Correlated Normal Variables. J Royal Statistical Society. Series D (The Statistician). September 1982; 31(3): 251-258.

(2) Nadarajah S. On the ratio X/Y for some elliptically symmetric distributions. J Multivariate Analysis. 2006; 97: 342 – 358.

(3) Shoukri M M, Colak D, Kaya N, Donner A. Comparison of two dependent within subject coefficients of variation to evaluate the reproducibility of measurement devices. J BMC Medical Research Methodology. 2008; 8(24): 1- 11.

(4) Hayya J, Armstrong D, Gressis N. A note on ratio of two normally distributed variables. J Institute of Management Science. July 1975; 21(11): 1338-1341.

(5) Fieller EC. The distribution of the index in a normal bivariate population. J Biometrika.1932; 24: 428-40.

(6) Hinkley DV. On the ratio of two correlated normal random variables. J Biometrika. 1969; 56 (3): 635 – 639.

(7) Kim YY, Lee JT, Choi GH. An investigation on the causes of cycle variation in direct injection hydrogen fueled engines. Inter J Hydrogen Energy. 2005; 30: 69 –76.

(8) Holmes DT, Buhr KA. Error propagation in calculated ratios. J Clinical Biochemistry. 2007; 40: 728–734.

(9) Pang WK, Yu BW, Troutt MD, Hou SH. A simulation-based approach to the study of coefficient of variation of dividend yields. J, European J Operational Research. 2008; 189: 559–569.

(10) Euser A M, Dekker FW, Cessie Sl. A practical approach to Bland-Altman plots and variation Coefficients for log transformed variables. J, Clinical Epidemiology. 2008; 61: 978 – 982.

(11) Cox JC, Sadiraj V. On the coefficient of variation as a criterion for decision under risk. J Mathematical Psychology. 2010; 54: 387-394.

(12) Blumenfeld D.Operations research calculations handbook. Boca Raton London New York Washington, D.C.: CRC Press LLC. 2001.

(13) Hooks T, Marx D, Kachman S, Pedersen J, Eigenberg R. Analysis of Covariance with Spatially Correlated Secondary Variables. J Revista Colombiana de Estadstica. 2008; 31(1): 95- 109.

(14) Steven, R. Lay. Analysis: an introduction to proof. Department of Mathematics Aurora University, Aurora, Illinois. 1986.