ช่วงความเชื่อมั่นค่าสัมประสิทธิ์การแปรผันควอไทล์ โดยวิธีโบเน็ต และวิธีบูตสแตรป

พรรณนา เอี่ยมสุวรรณ์*
อาจารย์ ภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยารามคำแหง

* ผู้นิพนธ์ประสานงาน โทรศัพท์ 08-7672-3249 อีเมล: phannana18@gmail.com
รับเมื่อ 25 กันยายน 2557 ตอบรับเมื่อ 12 มกราคม 2558
DOI: 10.14416/j.kmutnb.2015.01.002

บทคัดย่อ

วัตถุประสงค์ของบทความทางวิชาการนี้ เป็นการศึกษาวิธีการประมาณช่วงความเชื่อมั่นค่าสัมประสิทธิ์การแปรผันควอไทล์ทั้งหมด 3 วิธี ได้แก่ 1) วิธีโบเน็ต 2) วิธีบูตสแตรป และ 3) วิธีโบเน็ตบูตสแตรป ใช้การจำลองข้อมูลจากประชากรที่มีการแจกแจง *Normal*(4,1), *Lognormal*(0,1) และ *Gamma*(1.5,1) ขนาดตัวอย่าง 10, 25, 50 และ 100 ด้วยการทำซ้ำโดยเทคนิคมอนติคาร์โล ผลการศึกษาพบว่า วิธีบูตสแตรปให้ค่าประมาณค่าความน่าจะเป็นครอบคลุมเข้าใกล้ระดับความเชื่อมั่นที่กำหนดและให้ค่าความยาวเฉลี่ยของช่วงสั้นที่สุด ยกเว้นเมื่อข้อมูลมีการแจกแจงแบบ *Lognormal*(0,1) และ *Gamma*(1.5,1) และมีขนาดตัวอย่างเท่ากับ 100 วิธีโบเน็ตมีประสิทธิภาพ

คำสำคัญ: บูตสแตรป เทคนิคมอนติคาร์โล

การอ้างอิงบทความ: พรรณนา เอี่ยมสุวรรณ์, "ช่วงความเชื่อมั่นค่าสัมประสิทธิ์การแปรผันควอไทล์ โดยวิธีโบเน็ต และวิธีบูตสแตรป," *วารสารวิชาการพระจอมเกล้าพระนครเหนือ*, ปีที่ 25, ฉบับที่ 1, หน้า 161-168, ม.ค. - เม.ย. 2558. http://dx.doi.org/10.14416/j.kmutnb.2015.01.002

Confidence Intervals for a Coefficient of Quartile Variation with Bonett Method and Bootstrap Method

Phannana Aiemsuwan*
Lecturer, Department of Statistics, Faculty of Science, Ramkhamhaeng University, Bangkok, Thailand

* Corresponding Author, Tel. 08-7672-3249, E-mail: phannana18@gmail.com
Received 25 September 2014; Accepted 12 January 2015
DOI: 10.14416/j.kmutnb.2015.01.002

Abstract

The purpose of this paper was to study the confidence intervals for a coefficient of quartile variation. The three methods were: 1) the Bonett method 2) the Bootstrap method and 3) the Bonett Bootstrap method. The study was carried out by generating random samples of 10, 25, 50 and 100 with *Normal*(4,1), *Lognormal*(0,1) and *Gamma*(0.5,1) distributions. The sample size and distribution were then checked for computing confidence intervals for a coefficient of quartile variation using Monte Carlo techniques. The result indicated that the bootstrap method provided the coverage probability close to the nominal level with the smallest average interval length for all sample sizes and all distributions. The Bonett method performed better when the distributions of the samples were *Lognormal*(0,1) and *Gamma*(0.5,1) with the size of 100.

Keywords: Bootstrap, Monte Carlo Techniques

Please cite this article as: P. Aiemsuwan, "Confidence Intervals for a Coefficient of Quartile Variation with Bonett Method and Bootstrap Method," *J. KMUTNB*, Vol. 25, No. 1, pp. 161-168, Jan. - Apr. 2015 (in Thai). http://dx.doi.org/10.14416/j.kmutnb.2015.01.002

1. บทนำ

การวัดการกระจาย (Measure of Dispersion) เป็นสถิติประเภทหนึ่งที่ใช้อธิบายลักษณะของข้อมูลที่ศึกษา ซึ่งบอกถึงการกระจายของกลุ่มข้อมูลว่าค่าต่างๆ นั้นมีค่าแตกต่างกันมากน้อยเพียงใด ถ้าข้อมูลมีค่าต่างกันมาก เรียกว่า ข้อมูลมีการกระจายมาก และในทางตรงกันข้าม ถ้าข้อมูลมีค่าต่างกันน้อยเรียกว่า ข้อมูลมีการกระจายน้อย ซึ่งการวัดการกระจายทำให้เห็นภาพรวมของข้อมูลชุดนั้นๆ ได้ชัดเจนยิ่งขึ้น โดยจะแบ่งเป็น 2 ลักษณะ คือ การวัดการกระจายแบบสัมบูรณ์ (Absolute Dispersion) เป็นการพิจารณาการกระจายของข้อมูลภายในกลุ่มเดียวกัน เพื่อดูว่าข้อมูลกลุ่มนั้นแต่ละค่ามีค่าแตกต่างกันมากน้อยเพียงใด และการวัดการกระจายแบบสัมพัทธ์ (Relative Dispersion) เป็นการเปรียบเทียบการกระจายของข้อมูลต่างกลุ่ม โดยใช้อัตราส่วนของค่าที่ได้จากการวัดการกระจายแบบสัมบูรณ์ กับค่ากลางของข้อมูลนั้นๆ ซึ่งในบทความนี้ ผู้เขียนสนใจค่าวัดการกระจายแบบสัมพัทธ์ที่เรียกว่า “ค่าสัมประสิทธิ์การแปรผันควอไทล์” (Coefficient of Quartile Variation: CQV)

$$CQV = \frac{Q_3 - Q_1}{Q_3 + Q_1}$$

เมื่อ Q_1 คือค่าควอไทล์ที่ 1 หรือค่าเปอร์เซ็นไทล์ที่ 25 ของประชากร Q_3 คือค่าควอไทล์ที่ 3 หรือค่าเปอร์เซ็นไทล์ที่ 75 ของประชากร ในปี ค.ศ.1959 Wilks [1] ได้นำเสนอการประมาณค่าช่วงความเชื่อมั่น 95% ของค่าควอไทล์ที่ 1 เป็น $(Y_{(a)}, Y_{(b)})$ และค่าควอไทล์ที่ 3 เป็น $(Y_{(c)}, Y_{(d)})$ โดยที่ $Y_{(j)}$ เป็นค่าสถิติลำดับที่ j, $a = n/4 - 1.96(3n/16)^{1/2}$, $b = n/4 + 1.96(3n/16)^{1/2}$, $c = n + 1 - b$ และ $d = n + 1 - a$ เมื่อ $a \geq 1$ ต่อมาในปี ค.ศ.1994 Stuard and Ord [2] ได้นำเสนอค่าความแปรปรวนของตัวประมาณค่าควอไทล์ที่ 1 ตัวประมาณค่าควอไทล์ที่ 3 และความแปรปรวนร่วมของตัวประมาณค่าควอไทล์ที่ 1 กับตัวประมาณค่าควอไทล์ที่ 3 ดังนี้

$$v_1(\hat{Q}_1) = 3/(n16 f_1^2)$$

$$v_1(\hat{Q}_3) = 3/(n16 f_3^2)$$

และ $$cov(\hat{Q}_1, \hat{Q}_3) = 1/(n16 f_1 f_3)$$

โดยมี f_1 และ f_3 เป็นฟังก์ชันความหนาแน่นของความน่าจะเป็นของตัวประมาณค่าควอไทล์ที่ 1 และตัวประมาณค่าควอไทล์ที่ 3 ตามลำดับ ต่อมาได้มีการนำเสนอค่าความแปรปรวนของตัวประมาณค่าควอไทล์ที่ 1 กับตัวประมาณค่าควอไทล์ที่ 3 ซึ่งปราศจาก f_1 และ f_3 แต่จะใช้ช่วงความเชื่อมั่นของค่าควอไทล์ที่ 1 และค่าควอไทล์ที่ 3 ซึ่งนำมาเป็นสัดส่วนกับค่า $z_{1-\alpha^*/2}$ แสดงค่าความแปรปรวนได้ดังนี้

$$v_2(\hat{Q}_1) = (Y_{(b)} - Y_{(a)})^2 / (2z_{1-\alpha^*/2})^2$$

และ $$v_2(\hat{Q}_3) = (Y_{(d)} - Y_{(c)})^2 / (2z_{1-\alpha^*/2})^2$$

ซึ่งตัวประมาณค่านี้ถูกนำเสนอในปี ค.ศ. 1998 โดย Hettmansperger and Mckean [3] ต่อมาในปี ค.ศ. 2001 วิธีการนี้ Price and Bonett [4] ได้นำไปประมาณค่าความแปรปรวนของตัวประมาณค่ามัธยฐานเมื่อข้อมูลมีการแจกแจงต่างๆ ซึ่งกล่าวได้ว่าเป็นตัวประมาณค่าที่ง่ายต่อการคำนวณและให้ค่าความเอนเอียงและค่าเฉลี่ยความคลาดเคลื่อนต่ำ จากการประมาณค่าความแปรปรวนของตัวประมาณค่าควอไทล์ทั้งสองวิธี ในปีค.ศ. 2006 Bonett [5] ได้ทำการประมาณค่า f_1 และ f_3 แสดงได้ดังนี้

$$\hat{f}_1^2 = 3(z_{1-\alpha^*/2})^2 / \left\{4n((Y_{(b)} - Y_{(a)})^2)\right\}$$

และ $$\hat{f}_3^2 = 3(z_{1-\alpha^*/2})^2 / \left\{4n((Y_{(d)} - Y_{(c)})^2)\right\}$$

เพื่อใช้ในการประมาณค่าช่วงความเชื่อมั่นค่าสัมประสิทธิ์การแปรผันควอไทล์ ซึ่งช่วงความเชื่อมั่นของค่าสัมประสิทธิ์การแปรผันควอไทล์ที่ถูกพัฒนาขึ้นนั้นจะมี

ประสิทธิภาพในกรณีที่นำไปใช้กับการแจกแจงที่ไม่ปกติ ต่อมาในปี ค.ศ.2014 Tongkaw and Pongsakchat [7] ได้ทำการสร้างช่วงความเชื่อมั่นสำหรับสัมประสิทธิ์การแปรผันควอไทล์ โดยทำการแก้ไขดัดแปลงจากวิธีของ Bonett [5] และวิธีบูตสแตรป (Bootstrap Method) ที่นำเสนอโดย Efron and Tibshirani [6] โดยทำการประมาณค่าแบบช่วงทั้งหมดสองวิธีคือ วิธีบูตสแตรป (Bootstrap Method) และวิธีโบเน็ตบูตสแตรป (Bonett Bootstrap Method) ซึ่งใช้การจำลองข้อมูลโดยเทคนิคมอนติคาร์โล จำลองข้อมูลทั้งที่มีการแจกแจงแบบปกติและไม่ปกติกล่าวคือ *Normal*(0,1), *Lognormal*(0,1) และ *Gamma*(1.5,1) ซึ่งค่าสัมประสิทธิ์การแปรผันควอไทล์ เป็นค่าวัดการกระจายที่มักพบในการศึกษาไม่ว่าจะเป็นด้านวิทยาศาสตร์ ด้านสังคมศาสตร์ ผู้เขียนจึงสนใจศึกษาวิธีการประมาณค่าช่วงความเชื่อมั่นสำหรับสัมประสิทธิ์การแปรผันควอไทล์ ของวิธีที่กล่าวมาข้างต้น โดยเกณฑ์การวัดประสิทธิภาพจะพิจารณาจากค่าประมาณความน่าจะเป็นครอบคลุม (Coverage Probability) และค่าความยาวเฉลี่ย (Average Length)

2. วิธีบูตสแตรป (Bootstrap Method)

ปี ค.ศ.1993 Efron and Tibshirani [6] ทำการเสนอวิธีการประมาณค่าพารามิเตอร์โดยใช้เทคนิคการสุ่มตัวอย่างซ้ำ (Resampling Technique) จะใช้การสุ่มตัวอย่างชุดใหม่จากตัวอย่างสุ่มที่มีเพียงชุดเดียวโดยการสุ่มตัวอย่างแบบใส่คืน ทำการสุ่มตัวอย่างขนาด n จากชุดตัวอย่าง $y_1, y_2, \ldots, y_n$ โดยค่าที่ได้จะใส่คืนกลับไปในชุดตัวอย่างก่อนที่จะมีการสุ่มตัวอย่างครั้งต่อไป ให้ $y^*_1, y^*_2, \ldots, y^*_n$ เป็นชุดของตัวอย่างขนาด n ที่สุ่มได้ ซึ่งจะเรียกชุดของตัวอย่างดังกล่าวนี้ว่า ตัวอย่างบูตสแตรป (Bootstrap Sample) แล้วคำนวณหาตัวประมาณที่ได้จากการสุ่มตัวอย่าง บูตสแตรปแต่ละรอบ $\hat{\theta}_i$, $i = 1, 2, \ldots, B$ โดยที่ B คือจำนวนครั้งในการสุ่มซ้ำ ซึ่ง $\hat{\theta}^{(B)}$ จะเป็นตัวประมาณค่าของพารามิเตอร์ นอกจากนี้ Efron and Tibshiran [6] ได้นำเสนอวิธีการประมาณค่าช่วงความเชื่อมั่นเรียกว่า วิธีบูตสแตรปที ซึ่งจะทำการประมาณค่า z^*_i จากตัวอย่างบูตสแตรปที่ได้จากการสุ่มซ้ำ B ครั้ง อีกทั้งยังนำเสนอวิธีเปอร์เซ็นไทล์บูตสแตรป ซึ่งเป็นวิธีการประมาณค่าช่วงความเชื่อมั่นที่ใช้ค่าเปอร์เซ็นไทล์ของตัวประมาณค่าที่สนใจในจากการสุ่มซ้ำ B ครั้ง

3. การประมาณช่วงความเชื่อมั่นค่าสัมประสิทธิ์การแปรผันควอไทล์

3.1 ช่วงความเชื่อมั่นค่าสัมประสิทธิ์การแปรผันควอไทล์โดยวิธีโบเน็ต (The Bonett Method)

ในปี ค.ศ. 2006 Bonett [5] ให้ Y เป็นตัวแปรสุ่มที่มีการแจกแจงแบบต่อเนื่อง $y_1, y_2, \ldots, y_n$ เป็นตัวอย่างสุ่มขนาด n ให้ $\hat{Q}_1$ และ $\hat{Q}_3$ เป็นค่าเปอร์เซ็นไทล์ที่ 25 และ 75 ของประชากร ตามลำดับ ช่วงความเชื่อมั่น (1 - α)100% สำหรับสัมประสิทธิ์การแปรผันควอไทล์

$$\exp\left\{\log\left(\frac{D}{S}\right)c \pm z_{1-\alpha/2} v^{\frac{1}{2}}\right\}$$

เมื่อ $c = \dfrac{n}{n-1}$, $D = \hat{Q}_3 - \hat{Q}_1$, $S = \hat{Q}_3 + \hat{Q}_1$

จากนั้นทำการประมาณค่าความแปรปรวนของ $\log\left(\frac{D}{S}\right)$ เป็นดังนี้

$$v = \left(\frac{1}{16n}\right)\left\{\frac{\left(\frac{3}{\hat{f}_1^2} + \frac{3}{\hat{f}_3^2} - \frac{2}{\hat{f}_1^2\hat{f}_3^2}\right)}{D^2} + \frac{\left(\frac{3}{\hat{f}_1^2} + \frac{3}{\hat{f}_3^2} + \frac{2}{\hat{f}_1^2\hat{f}_3^2}\right)}{S^2} - \frac{\left(\frac{3}{\hat{f}_3^2} - \frac{3}{\hat{f}_1^2}\right)}{D^2}\right\}$$

จากนั้นประมาณค่าฟังก์ชันความหนาแน่นของความน่าจะเป็นของตัวประมาณค่าควอไทล์ที่ 1 และตัวประมาณค่าควอไทล์ที่ 3 โดย

$$\hat{f}_1 = \sqrt{\frac{3\left(z_{1-\alpha^*/2}\right)^2}{\left\{4n\left(Y_{(b)} - Y_{(a)}\right)^2\right\}}}$$

และ

$$\hat{f}_3 = \sqrt{\frac{3\left(z_{1-\alpha^*/2}\right)^2}{\left\{4n\left(Y_{(d)} - Y_{(c)}\right)^2\right\}}}$$

$\hat{f}_1$ แทนตัวประมาณค่าฟังก์ชันความหนาแน่นของความน่าจะเป็นของตัวประมาณค่าควอไทล์ที่ 1

$\hat{f}_3$ แทนตัวประมาณค่าฟังก์ชันความหนาแน่นของความน่าจะเป็นของตัวประมาณค่าควอไทล์ที่ 3

โดยที่

$$(y_{(a)}, y_{(b)}) = \left(y_{\left(\frac{n}{4} - 1.96\sqrt{\frac{3n}{16}}\right)}, y_{\left(\frac{n}{4} + 1.96\sqrt{\frac{3n}{16}}\right)} \right)$$

$$(y_{(c)}, y_{(d)}) = \left(y_{n+1-b}, y_{n+1-a} \right)$$

เมื่อ $a \geq 1$

$z_{1-\alpha^*/2}$ แทนค่าควอนไทล์ $1-\alpha^*/2$ ของการแจกแจงปกติมาตรฐาน โดยที่ α^* ขึ้นอยู่กับขนาดของตัวอย่าง ดังนี้

- ถ้าขนาดตัวอย่างมีขนาดเล็ก ($n < 30$)

$$\alpha^* = 1 - \sum_{i=a}^{b-1} \binom{n}{i} \left(\frac{1}{4}\right)^i \left(\frac{3}{4}\right)^{n-i}$$

- ถ้าขนาดตัวอย่างมีขนาดใหญ่ ($n \geq 30$) ค่า α^* เท่ากับ 0.05

3.2 ช่วงความเชื่อมั่นค่าสัมประสิทธิ์การแปรผันควอไทล์โดยวิธีบูตสแตรป (The Bootstrap Method)

วิธีนี้ Tongkaw and Pongsakchat [7] ได้นำเสนอช่วงความเชื่อมั่นสำหรับสัมประสิทธิ์การแปรผันควอไทล์โดยใช้เทคนิคบูตสแตรป โดยทำการประมาณค่าช่วงความเชื่อมั่นโดยเทคนิคที่เรียกว่า เปอร์เซ็นไทล์บูตสแตรป (Percentile Bootstrap) ดังนั้นช่วงความเชื่อมั่น $(1-\alpha)100\%$ สำหรับสัมประสิทธิ์การแปรผันควอไทล์ คือ

$$[CQV_{BL}, CQV_{BU}]$$

โดย CQV_{BL} เป็นเปอร์เซ็นไทล์ที่ $(\alpha/2)100\%$ ของสัมประสิทธิ์การแปรผันควอไทล์ของตัวอย่างบูตสแตรป CQV_{BU} เป็นเปอร์เซ็นไทล์ที่ $(1-\alpha/2)100\%$ ของสัมประสิทธิ์การแปรผันควอไทล์ของตัวอย่างบูตสแตรป ซึ่งวิธีนี้จะทำการประมาณค่าสัมประสิทธิ์การแปรผันควอไทล์จากวิธีการสุ่มแบบบูตสแตรป จากนั้นทำการเรียงค่าประมาณที่ได้จากน้อยไปหามาก แล้วหาค่าเปอร์เซ็นไทล์ ซึ่งเปอร์เซ็นไทล์ที่ $(\alpha/2)100\%$ เป็นขีดจำกัดความเชื่อมั่นด้านล่าง และเปอร์เซ็นไทล์ที่ $(1-\alpha/2)100\%$ เป็นขีดจำกัดความเชื่อมั่นด้านบน ของช่วงความเชื่อมั่นค่าสัมประสิทธิ์การแปรผันควอไทล์

3.3 ช่วงความเชื่อมั่นค่าสัมประสิทธิ์การแปรผันควอไทล์โดยวิธีโบเน็ตบูตสแตรป (The Bonett Bootstrap Method)

วิธีโบเน็ตบูตสแตรปถูกสร้างขึ้นโดย Tongkaw and Pongsakchat [7] ในปี ค.ศ.2014 การผสมผสานระหว่างวิธีโบเน็ตเข้ากับวิธีบูตสแตรป

ตัวประมาณค่าแบบจุดของค่าความแปรปรวนของ $\log\left(\frac{D}{S}\right)$ หรือ $\log\left(\frac{\hat{Q}_3^* - \hat{Q}_1^*}{\hat{Q}_3^* + \hat{Q}_1^*}\right)$ ที่ได้นั้นจะทำการประมาณมาจากวิธีบูตสแตรปเพราะฉะนั้นช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของสัมประสิทธิ์การแปรผันควอไทล์โดยวิธีโบเน็ตบูตสแตรป คือ

$$\exp\left\{\log\left(\frac{D}{S}\right) c \pm t^*_{1-\alpha/2} v_B^{1/2}\right\}$$

เมื่อ

$$c = \frac{n}{n-1}$$

ให้ $\hat{Q}_1^*$ และ $\hat{Q}_3^*$ เป็นค่าเปอร์เซ็นไทล์ที่ 25 และ 75 ของตัวอย่างที่ได้จากวิธีบูตสแตรป

ค่าความแปรปรวนของ $\log\left(\frac{D}{S}\right)$ คือ

$$v = \left(\frac{1}{16n}\right)\left\{\frac{\left(\frac{3}{\hat{f}_{1B}^2} + \frac{3}{\hat{f}_{3B}^2} - \frac{2}{\hat{f}_{1B}^2 \hat{f}_{3B}^2}\right)}{D_B^2} + \frac{\left(\frac{3}{\hat{f}_{1B}^2} + \frac{3}{\hat{f}_{3B}^2} + \frac{2}{\hat{f}_{1B}^2 \hat{f}_{3B}^2}\right)}{S_B^2} - \frac{\left(\frac{3}{\hat{f}_{3B}^2} - \frac{3}{\hat{f}_{1B}^2}\right)}{D_B S_B}\right\}$$

ทำการประมาณค่าฟังก์ชันความหนาแน่นของความน่าจะเป็นของตัวประมาณค่าควอไทล์ที่ 1 และตัวประมาณค่าควอไทล์ที่ 3 โดยวิธีบูตสแตรป

$$\hat{f}_{1B} = \sqrt{\frac{3\left(z_{1-\alpha^*/2}\right)^2}{\left\{4n\left(\hat{Y}^*_{(b)} - \hat{Y}^*_{(a)}\right)^2\right\}}}$$

และ

$$\hat{f}_{3B} = \sqrt{\frac{3\left(z_{1-\alpha^*/2}\right)^2}{\left\{4n\left(\hat{Y}^*_{(d)} - \hat{Y}^*_{(c)}\right)^2\right\}}}$$

$\hat{f}_{1B}$ แทนตัวประมาณค่าฟังก์ชันความหนาแน่นของความน่าจะเป็นของตัวประมาณค่าควอไทล์ที่ 1 โดยวิธีบูตสแตรป

$\hat{f}_{3B}$ แทนตัวประมาณค่าฟังก์ชันความหนาแน่นของความน่าจะเป็นของตัวประมาณค่าควอไทล์ที่ 3 โดยวิธีบูตสแตรป

โดยที่

$$(y^*_{(a)}, y^*_{(b)}) = \left(y^*_{\left(\frac{n}{4}-1.96\sqrt{\frac{3n}{16}}\right)}, y^*_{\left(\frac{n}{4}+1.96\sqrt{\frac{3n}{16}}\right)} \right)$$

$$(y^*_{(c)}, y^*_{(d)}) = \left(y^*_{n+1-b}, y^*_{n+1-a} \right)$$

$\hat{Y}^*_j$ เป็นสถิติลำดับที่ j จากวิธีบูตสแตรป โดยจะใช้ $t^*_{1-\alpha/2}$ จากวิธีบูตสแตรป เมื่อ $a \geq 1$

$t^*_{1-\alpha/2}$ แทนค่าควอนไทล์ (1 - α/2)100% จากวิธีบูตสแตรป

3.4 เกณฑ์ในการพิจารนา

3.4.1 ค่าประมาณความน่าจะเป็นครอบคลุม (Coverage Probability)

ค่าประมาณความน่าจะเป็นครอบคลุมในการประมาณช่วงความเชื่อมั่นค่าสำหรับสัมประสิทธิ์การแปรผันควอไทล์คำนวณจากค่าความน่าจะเป็นที่ช่วงความเชื่อมั่นค่าสัมประสิทธิ์การแปรผันควอไทล์ที่คำนวณได้จากวิธีการประมาณต่าง ๆ จะครอบคลุมค่าสัมประสิทธิ์การแปรผันควอไทล์ที่แท้จริง

$$Coverage\ Probability = P(CQV \in CI)$$

โดย CI คือช่วงความเชื่อมั่นค่าสำหรับสัมประสิทธิ์การแปรผันควอไทล์

3.4.2 ค่าความยาวเฉลี่ย (Average Length)

ค่าความยาวเฉลี่ยในการประมาณค่าช่วงความเชื่อมั่นค่าสัมประสิทธิ์การแปรผันควอไทล์จะพิจารณาค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นที่ประมาณได้จากวิธีต่าง ๆ วิธีคำนวณโดยนำค่าขีดจำกัดความเชื่อมั่นด้านบนลบด้วยค่าขีดจำกัดความเชื่อมั่นด้านล่างของช่วงจะได้ค่าความยาวช่วง จากนั้นนำค่าความยาวช่วงมาหาค่าเฉลี่ย

4. ผลการทดลอง

4.1 ผลการทดลองวิธีโบเน็ต

Bonett [5] ได้ทำการสร้างช่วงความเชื่อมั่นค่าสัมประสิทธิ์การแปรผันควอไทล์ ที่ระดับความเชื่อมั่น 95% โดยทำการจำลองข้อมูลที่มีการแจกแจง *Normal*(4,1), *Lognormal*(0,1) และ *Gamma*(1.5,1) ให้มีขนาดตัวอย่าง 10, 25, 50 และ 100 ได้ค่าประมาณความน่าจะเป็นครอบคลุมของช่วงความเชื่อมั่นดังตารางที่ 1 พบว่าช่วงความเชื่อมั่นสำหรับสัมประสิทธิ์การแปรผันควอไทล์โดยวิธีโบเน็ตให้ค่าประมาณความน่าจะเป็นครอบคลุม (Coverage Probability) เข้าใกล้ระดับความเชื่อมั่นที่กำหนดเมื่อตัวอย่างสุ่มมีการแจกแจงแบบไม่ปกติคือ *Lognormal*(0,1) และ *Gamma*(1.5,1) และยังพบว่าช่วงความเชื่อมั่นนี้จะมีประสิทธิภาพดีเมื่อตัวอย่างมีขนาดใหญ่

4.2 ผลการทดลองวิธีบูตสแตรปและวิธีโบเน็ตบูตสแตรป

Tongkaw and Pongsakchat [7] สร้างช่วงความเชื่อมั่นค่าสัมประสิทธิ์การแปรผันควอไทล์ ที่ระดับความเชื่อมั่น

95% โดยทำการจำลองข้อมูลที่มีการแจกแจง *Normal*(4,1), *Lognormal*(0,1) และ *Gamma*(1.5,1) ให้มีขนาดตัวอย่าง 10, 25, 50 และ 100 ได้ค่าประมาณความน่าจะเป็นครอบคลุมของช่วงความเชื่อมั่นดังตารางที่ 2 พบว่าเมื่อตัวอย่างสุ่มมีการแจกแจง *Normal*(4,1) วิธีบูตสแตรปให้ค่าประมาณความน่าจะเป็นครอบคลุมเข้าใกล้ระดับความเชื่อมั่น 95% มากกว่าวิธีโบเน็ตบูตสแตรปที่ทุกขนาดตัวอย่าง สำหรับรูปแบบการแจกแจงที่ไม่ปกติ คือ *Lognormal*(0,1) และ *Gamma*(1.5,1) วิธีบูตสแตรปให้ค่าประมาณความน่าจะเป็นครอบคลุมเข้าใกล้ระดับความเชื่อมั่น 95% เมื่อตัวอย่างสุ่มมีขนาด 10, 25 และ 50 ยกเว้นเมื่อตัวอย่างสุ่มมีขนาด 100 วิธีโบเน็ตบูตสแตรปให้ค่าประมาณความน่าจะเป็นครอบคลุมเข้าใกล้ระดับความเชื่อมั่น 95% หากพิจารณาค่าความยาวเฉลี่ยของช่วง วิธีบูตสแตรปให้ค่าความยาวเฉลี่ยของช่วงสั้นกว่าวิธีโบเน็ตบูตสแตรปทั้งกรณีที่ตัวอย่างสุ่มมีการแจกแจงแบบปกติและไม่ปกติ คือ *Normal*(4,1), *Lognormal*(0,1) และ *Gamma*(1.5,1) ที่ทุกขนาดตัวอย่างยกเว้นที่ตัวอย่างมีขนาด 100

ตารางที่ 1 ค่าประมาณความน่าจะเป็นครอบคลุมของช่วงความเชื่อมั่นค่าสัมประสิทธิ์การแปรผันควอไทล์ โดยวิธีโบเน็ต

การแจกแจง	cqv	n	ค่าประมาณความน่าจะเป็นครอบคลุม
Normal(4,1)	0.169	10	0.965
		25	0.961
		50	0.956
		100	0.952
Lognormal(0,1)	0.587	10	0.970
		25	0.956
		50	0.955
		100	0.954
Gamma(1.5,1)	0.544	10	0.974
		25	0.958
		50	0.954
		100	0.952

แหล่งข้อมูล: Bonett [5]

ตารางที่ 2 ค่าประมาณความน่าจะเป็นครอบคลุมและค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นค่าสัมประสิทธิ์การแปรผันควอไทล์โดยวิธีโบเน็ต วิธีบูตสแตรป และวิธีโบเน็ตบูตสแตรป

การแจกแจง		***Normal*(4,1)**				
วิธี	***n***	**10**	**15**	**25**	**50**	**100**
bootstrap	CP	0.9442	0.9464	0.9570	0.9564	0.9560
	(AL)	0.6222	0.4571	0.3696	0.2506	0.1756
Bonett-bootstrap	CP	0.9260	0.9618	0.9754	0.9752	0.9350
	(AL)	0.7118	0.5708	0.4996	0.2912	0.1673
การแจกแจง		***Lognormal*(0,1)**				
วิธี	***n***	**10**	**15**	**25**	**50**	**100**
bootstrap	CP	0.9535	0.9589	0.9686	0.9704	0.9689
	(AL)	0.6413	0.5254	0.4348	0.2979	0.2095
Bonett-bootstrap	CP	0.9266	0.9676	0.9757	0.9842	0.9497
	(AL)	0.7556	1.1337	0.7912	0.3801	0.2042
การแจกแจง		***Gamma*(1.5,1)**				
วิธี	***n***	**10**	**15**	**25**	**50**	**100**
bootstrap	CP	0.9488	0.9512	0.9674	0.9666	0.9652
	(AL)	0.6218	0.5154	0.4332	0.3002	0.2131
Bonett-bootstrap	CP	0.9254	0.9576	0.9642	0.9802	0.9460
	(AL)	0.6221	0.5800	0.4900	0.3403	0.2007

CP ค่าประมาณความน่าจะเป็นครอบคลุม AL ค่าความยาวเฉลี่ย
แหล่งข้อมูล: Tongkaw and Pongsakchat [7]

5. อภิปรายผลและสรุป

วิธีบูตสแตรปมีประสิทธิภาพ กล่าวคือ ให้ค่าประมาณความน่าจะเป็นครอบคลุมเข้าใกล้ระดับความเชื่อมั่นที่กำหนดและให้ความยาวเฉลี่ยของช่วงที่สั้นทั้งในกรณีที่ตัวอย่างมีการแจกแจงแบบปกติและไม่ปกติ คือ *Normal*(4,1), *Lognormal*(0,1) และ *Gamma*(1.5,1) ยกเว้นเมื่อตัวอย่างสุ่มมีขนาดใหญ่คือ 100 และมีการแจกแจงแบบไม่ปกติคือ *Lognormal*(0,1) และ *Gamma*(1.5,1) วิธีโบเน็ตมีประสิทธิภาพดี

เนื่องจากคุณสมบัติของบูตสแตรปนั้นเป็นตัวประมาณที่คงเส้นคงวา (Consistency) อยู่แล้ว Efron and Tibshirani [6] ดังนั้นเมื่อขนาดตัวอย่างลู่เข้าสู่อนันต์แล้ว ส่วนเบี่ยงเบนมาตรฐานลู่เข้าสู่ศูนย์ ซึ่งจะส่งผลให้ค่าความยาวเฉลี่ยของช่วงนั้นสั้นเสมอ แต่มีข้อเสียคือ การคำนวณที่ค่อนข้างยุ่งยากและไม่มีรูปแบบที่ชัดเจน (Implicit Formula) ในขณะที่วิธีโบเน็ตไม่ยุ่งยากนัก อีกทั้งมีรูปแบบการคำนวณที่ชัดเจน

เอกสารอ้างอิง

[1] S.S. Wilks, *Nonparametric Statistical Inference*, New York: Wiley, 1959.

[2] A. Stuart and D.G. Ord, *Kendalls' Advanced Theory of Statistics*, Volume I: Distribution Theory. Sixth ed., London: Edward Arnold, 1994.

[3] T.P. Hettmansperger and J.W. McKean, *Robust Nonparametric Statistical Methods*, London: Wiley, 1998.

[4] R.M. Price and D.G. Bonett, "Estimating the variance of the sample median," *J. Statist. Comput*. Simulation, vol. 68, pp. 295-305, 2001.

[5] D.G. Bonett, "Confidence interval for a coefficient of quartile variation," *Comput. Statist. Data Anal*, pp. 2953-2957, 2006.

[6] B. Efron and R.J. Tibshirani, *An Introduction to the Bootstrap*, U.K.: Chapman & Hall, 1993.

[7] A. Tongkaw and V. Pongsakchat, "Confidence intervals for a coefficient of quartile variation with bootstrap method," *ICAS 2014*, Khon Kaen, 2014.