การประมาณค่าสัมประสิทธิ์ความเบ้ของการแจกแจงอินเวอร์สเกาส์เซียน เมื่อข้อมูลมีค่าสูญหาย: การศึกษาเชิงจำลอง

The Estimation for the Coefficients of Skewness of Inverse Gaussian Distribution with Missing Values: A Simulation Study

วราฤทธิ์ พานิชกิจโกศลกุล[1*]

Wararit Panichkitkosolkul[1*]

บทคัดย่อ

การศึกษาครั้งนี้มีวัตถุประสงค์เพื่อเสนอวิธีการประมาณค่าสัมประสิทธิ์ความเบ้ของการแจกแจงอินเวอร์สเกาส์เซียน เมื่อข้อมูลมีค่าสูญหาย และเปรียบเทียบวิธีการประมาณค่าสัมประสิทธิ์ความเบ้ 3 วิธี คือ 1) วิธีอย่างง่าย 2) วิธีปรับค่าสัมประสิทธิ์ความเบ้ด้วยเทอมค่าคงที่ $(n-1)/n$ และ 3) วิธีปรับค่าสัมประสิทธิ์ความเบ้ด้วยเทอมค่าคงที่ $(n-2)/n$ การวิจัยครั้งนี้ใช้การจำลองแบบมอนติคาร์โล โดยการเปรียบเทียบค่าความเอนเอียงสัมบูรณ์ ($|Bias|$) และค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error: MSE) ของค่าประมาณ กำหนดขนาดตัวอย่าง (n) เท่ากับ 20, 40, 60, 80, 100 และ 200 กำหนดค่าพารามิเตอร์ μ เท่ากับ 1 พารามิเตอร์ λ เท่ากับ 1, 3, 5, 10, 15 และ 20 และร้อยละของค่าสูญหาย (p) เท่ากับ ร้อยละ 5 และ ร้อยละ 10 ผลการวิจัยสรุปได้ดังนี้

วิธีปรับค่าสัมประสิทธิ์ความเบ้ด้วยเทอมค่าคงที่ $(n-2)/n$ ให้ค่าความเอนเอียงสัมบูรณ์ ($|Bias|$) และค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ต่ำที่สุด ทุกกรณีที่ศึกษา รองลงมาคือ วิธีปรับค่าสัมประสิทธิ์ความเบ้ด้วยเทอมค่าคงที่ $(n-1)/n$ และวิธีอย่างง่าย ตามลำดับ

คำสำคัญ : การจำลองแบบมอนติคาร์โล ค่าสัมประสิทธิ์ความเบ้ การแจกแจงอินเวอร์สเกาส์เซียน ค่าสูญหาย

---

[1] ภาควิชาคณิตศาสตร์และสถิติ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์ อำเภอคลองหลวง จังหวัดปทุมธานี 12121

[1] Department of Mathematics and Statistics, Faculty of Science and Technology, Thammasat University, Khlongluang, Pathumthani 12121, Thailand.

* ผู้นิพนธ์ประสานงาน ไปรษณีย์อิเล็กทรอนิกส์ (Corresponding author, e-mail): wararit@mathstat.sci.tu.ac.th

ABSTRACT

The objectives of this study are to propose the estimation method for coefficients of skewness of Inverse Gaussian distribution with missing values and to compare three estimation methods for coefficients of skewness of Inverse Gaussian distribution. The following estimation methods are considered: the simple method, adjusted coefficients of skewness method with term $(n-1)/n$, and adjusted coefficients of skewness method with term $(n-2)/n$. Using Monte Carlo simulations, we compare the absolute bias ($|Bias|$) and the mean square errors (MSE) of these estimation methods. The comparisons were done by using sample sizes (n) equal to 20, 40, 60, 80, 100 and 200 whereas parameter μ is 1, parameters λ are 1, 3, 5, 10, 15, and 20, and the percentage of missing values are equal to 5% and 10%. Results of the study are as follows:

For all sample sizes and all parameter values, the $|Bias|$ and the MSE of adjusted coefficients of skewness method with term $(n-2)/n$ are the lowest. We can arrange the methods as follows: the adjusted coefficients of skewness method with term $(n-1)/n$ and simple method, respectively.

Key words: Monte Carlo simulation, coefficients of skewness, inverse Gaussian distribution, missing value

บทนำ

การแจกแจงอินเวอร์สเกาส์เซียน (Inverse Gaussian Distribution) เสนอขึ้นครั้งแรกโดย Schrodinger (1915) ซึ่งเป็นการแจกแจงหนึ่งที่นำมาประยุกต์ใช้กับปัญหาต่างๆ เช่น การวิเคราะห์ราคาหุ้น ปัญหาด้านชีววิทยา อุทกภูมิศาสตร์ อุตุนิยมวิทยา และความเชื่อถือได้ของผลิตภัณฑ์ เป็นต้น (Cohen and Whitten, 1988) โดยมีนักวิจัยหลายท่านที่สนใจศึกษาเกี่ยวกับการแจกแจงดังกล่าว อาทิ Tweedie (1957), Shuster (1968), Chhikara and Folks (1989), Seshadri (1994; 1999), Balakrisnan and Chen (1997) เป็นต้น ฟังก์ชันความหนาแน่นความน่าจะเป็น (Probability Density Function: pdf) ของการแจกแจงอินเวอร์สเกาส์เซียน เขียนได้ดังนี้ (Krishnamoorthy, 2006)

$$f(x\,|\,\mu,\lambda) = \sqrt{\frac{\lambda}{2\pi x^3}}\exp\left(\frac{-\lambda(x-\mu)^2}{2\mu^2 x}\right), \quad x>0 \ , \ \mu>0, \lambda>0$$

กราฟของฟังก์ชันความหนาแน่นความน่าจะเป็นแสดงดังรูป 1 สำหรับสมบัติที่สำคัญ

ของการแจกแจงอินเวอร์สเกาส์เซียน (Balakrishnan and Nevzorov, 2003) มีดังนี้

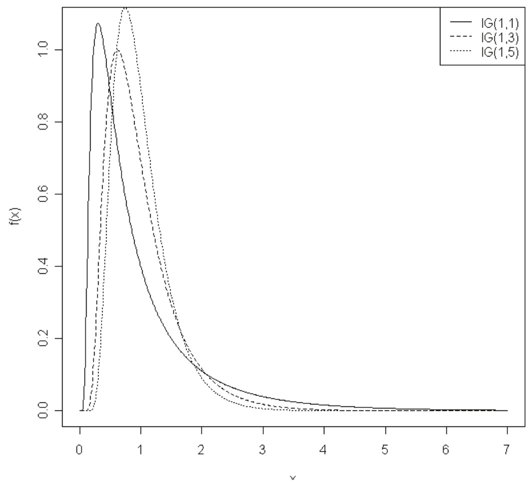
ค่าเฉลี่ย μ

ความแปรปรวน μ^3/λ

สัมประสิทธิ์ความเบ้ (γ_1) $3\sqrt{\dfrac{\mu}{\lambda}}$

สัมประสิทธิ์ความโด่ง (γ_2) $3+\dfrac{15\mu}{\lambda}$

ในบางกรณีผู้วิเคราะห์ต้องการศึกษาเกี่ยวกับความเบ้ (Skewness) ของประชากรซึ่งมีการแจกแจงอินเวอร์สเกาส์เซียน ซึ่งสามารถประมาณค่าโดยใช้ค่าสัมประสิทธิ์ความเบ้ของตัวอย่าง (Sample Coefficient of Skewness) ดังนั้นผู้วิเคราะห์จะต้องเก็บรวบรวมข้อมูลจากตัวอย่างเพื่อศึกษาถึงลักษณะความเบ้ของประชากร โดยปัญหาหนึ่งที่ผู้วิเคราะห์อาจประสบในการเก็บรวบรวมข้อมูล คือ ปัญหาค่าสูญหาย (Missing Values) ซึ่งส่งผลกระทบทำให้ค่าประมาณของพารามิเตอร์มีความคลาดเคลื่อน โดยทั่วไปมักแก้ไขปัญหาการสูญหายของข้อมูลด้วยการตัดค่าสูญหายทิ้ง และนำข้อมูลที่สมบูรณ์มาทำการวิเคราะห์ จึงมีผลทำให้ข้อมูลมีจำนวนลดน้อยลง แต่ในบางครั้งข้อมูลอาจมีจำนวนไม่มากนัก จึงไม่สามารถตัดค่าสูญหายออกได้ ทำให้ต้องมีการประมาณค่าสูญหาย ดังนั้นเมื่อผู้วิเคราะห์ทำการประมาณค่าสูญหายแล้วก็จะนำข้อมูลทั้งหมดไปประมาณค่าสัมประสิทธิ์ความเบ้ ซึ่งในการประมาณค่าสัมประสิทธิ์ความเบ้ของการแจกแจงอินเวอร์สเกาส์เซียน จะต้องประมาณค่าพารามิเตอร์ μ และ λ ของการแจกแจงดังกล่าว ซึ่งสามารถประมาณค่าได้หลายวิธี เช่น วิธีความควรจะเป็นสูงสุด (Maximum Likelihood Method: ML) และวิธีการประมาณ



รูป 1 ฟังก์ชันความหนาแน่นความน่าจะเป็นของการแจกแจงอินเวอร์สเกาส์เซียน เมื่อ $\mu = 1$ และ $\lambda = 1, 3$ และ 5

แบบไม่เอนเอียงซึ่งมีความแปรปรวนต่ำสุด (Minimum Variance Unbiased Estimation Method: MVUE) เป็นต้น (Evans *et al.*, 2000) เมื่อได้ค่าประมาณพารามิเตอร์ $\hat{\mu}$ และ $\hat{\lambda}$ แล้ว ขั้นตอนต่อไปจะนำค่าประมาณพารามิเตอร์ $\hat{\mu}$ และ $\hat{\lambda}$ ไปประมาณค่าสัมประสิทธิ์ความเบ้ วิธีดังกล่าวเรียกว่า วิธีอย่างง่าย (Simple Method) อย่างไรก็ดี วิธีอย่างง่ายให้ค่าประมาณพารามิเตอร์มากเกินกว่าค่าพารามิเตอร์ที่แท้จริง (True Parameter) ดังนั้นผู้วิจัยจึงมีแนวคิดที่จะนำเทอมค่าคงที่ที่มีค่าน้อยกว่า 1 มาปรับให้ค่าประมาณของสัมประสิทธิ์ความเบ้มีความคลาดเคลื่อนน้อยลง โดยทดลองกับเทอมค่าคงที่ในรูปแบบต่างๆ จากการค้นคว้ายังไม่พบว่ามีนักวิจัยคนใดทำการศึกษาในลักษณะดังกล่าวกับการประมาณค่าประสิทธิ์ความเบ้ของการแจกแจงอินเวอร์สเกาส์เซียน เมื่อข้อมูลมีค่าสูญหาย

ดังนั้นในการศึกษาครั้งนี้จึงมีวัตถุประสงค์เพื่อเสนอวิธีการประมาณค่าสัมประสิทธิ์ความเบ้ของการแจกแจงอินเวอร์สเกาส์เซียน เมื่อข้อมูลมีค่าสูญหาย และเปรียบเทียบวิธีการประมาณค่าสัมประสิทธิ์ความเบ้ 3 วิธี คือ 1) วิธีอย่างง่าย 2) วิธีปรับค่าสัมประสิทธิ์ความเบ้ด้วยเทอมค่าคงที่ $(n-1)/n$ และ 3) วิธีปรับค่าสัมประสิทธิ์ความเบ้ด้วยเทอมค่าคงที่ $(n-2)/n$ โดยการเปรียบเทียบค่าความเอนเอียงสัมบูรณ์ ($|Bias|$) และค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error: MSE) ของค่าประมาณ สำหรับขอบเขตของการศึกษามีดังต่อไปนี้

1. กำหนดขนาดตัวอย่างที่ใช้ (n) เท่ากับ 20, 40, 60, 80, 100 และ 200
2. กำหนดค่าพารามิเตอร์ μ เท่ากับ 1
3. กำหนดค่าพารามิเตอร์ λ เท่ากับ 1, 3, 5, 10, 15 และ 20
4. กำหนดร้อยละของค่าสูญหาย (p) เท่ากับ ร้อยละ 5 และ ร้อยละ 10
5. ประมาณค่าพารามิเตอร์ และ λ ด้วยวิธีความควรจะเป็นสูงสุด

การศึกษาครั้งนี้ใช้การจำลองแบบมอนติคาร์โลด้วยเครื่องไมโครคอมพิวเตอร์ โดยใช้โปรแกรม R เวอร์ชัน 2.8.0 และกำหนดให้เครื่องคอมพิวเตอร์ทำงานซ้ำๆ กัน 10,000 ครั้งในแต่ละกรณีที่ศึกษา

งานวิจัยที่เกี่ยวข้อง

วราฤทธิ์ (2550) เสนอวิธีการประมาณค่าฐานนิยมของข้อมูลที่มีการแจกแจงอินเวอร์สเกาส์เซียนโดยนำเทอมค่าคงที่ซึ่งขึ้นกับขนาดตัวอย่างมาปรับให้ค่าประมาณฐานนิยมมีความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ลดลง

วราฤทธิ์ (2552) ศึกษาวิธีการประมาณค่าโมเมนต์อันดับที่สองของการแจกแจงอินเวอร์สเกาส์เซียน พบว่าวิธีปรับค่าโมเมนต์อันดับที่สองด้วยเทอมค่าคงที่ $(n-2)/n$ ให้ค่าความเอนเอียงสัมบูรณ์ และความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุด

Gupta (1973) เสนอวิธีความควรจะเป็นสูงสุด (Maximum Likelihood Method: ML)

สำหรับประมาณค่าพารามิเตอร์ของการแจกแจงอินเวอร์สเกาส์เซียนแบบตัดปลายหาง (Truncated Inverse Gaussian Distribution)

Sanchez and He (2003) ได้ศึกษาระยะเวลาที่ผู้ใช้เข้าชมเว็บไซต์หนึ่งซึ่งข้อมูลมีการแจกแจงอินเวอร์สเกาส์เซียน โดยประมาณค่าพารามิเตอร์ด้วยวิธีความควรจะเป็นสูงสุด

Lee and Tang (2007) เสนอวิธีการประมาณค่าพารามิเตอร์ของการแจกแจงอินเวอร์สเกาส์เซียนโดยใช้วิธีอีเอ็มแบบดัดแปลง (Modified EM-Algorithm) ได้แก่ ตัวประมาณอีเอ็มแบบดัดแปลง (Modified EM Estimator: MEME) และตัวประมาณความควรจะเป็นแบบดัดแปลง (Modified Maximum Likelihood Estimators: MMLE)

Arefi *et al.* (2008) ศึกษาการประมาณค่าแบบช่วง (Interval Estimation) ของค่าเฉลี่ยของการแจกแจงอินเวอร์สเกาส์เซียน พบว่า ช่วงความเชื่อมั่นโดยใช้วิธีอัตราส่วนความควรจะเป็น (Likelihood Ratio Interval) มีประสิทธิภาพสูงที่สุด ส่วนวิธีของวาร์ล (Wald Interval) มีประสิทธิภาพต่ำที่สุด

วิธีการศึกษา

ในการศึกษาครั้งนี้มีวิธีการศึกษาดังนี้

3.1 การจำลองข้อมูลที่ใช้ในการศึกษา

จำลองข้อมูลที่ใช้ในการศึกษาโดยการสร้างตัวอย่างสุ่ม $X_1, X_2, ..., X_n$ จากการแจกแจงอินเวอร์สเกาส์เซียน ซึ่งมีพารามิเตอร์ μ และ λ ตามที่กำหนดไว้ในขอบเขตของการศึกษา

3.2 การประมาณค่าสูญหาย

สมมติว่า $X_k;\ 1 \le k \le n$ เป็นค่าสูญหาย ประมาณค่า X_k โดยเลือกข้อมูลมาอย่างสุ่มจากข้อมูลที่เก็บรวบรวมได้

3.3 การประมาณค่าพารามิเตอร์ μ และ λ ด้วยวิธีความควรจะเป็นสูงสุด

เมื่อสร้างตัวอย่างสุ่ม $X_1, X_2, ..., X_n$ จากการแจกแจงอินเวอร์สเกาส์เซียน และประมาณค่าสูญหายแล้ว ต่อไปจะประมาณค่าพารามิเตอร์ μ และ λ ด้วยวิธีความควรจะเป็นสูงสุด (Evans *et al*., 2000)

ตัวประมาณความควรจะเป็นสูงสุดของ μ

คือ $$\hat{\mu} = \bar{X} = n^{-1}\sum_{i=1}^{n} X_i$$

ตัวประมาณความควรจะเป็นสูงสุดของ λ

คือ $$\hat{\lambda} = \frac{n}{\left(\sum_{i=1}^{n} X_i^{-1} - (\bar{X})^{-1}\right)}$$

3.4 การประมาณค่าสัมประสิทธิ์ความเบ้ของข้อมูลด้วยวิธีการที่ศึกษา

เมื่อประมาณค่าพารามิเตอร์ μ และ λ นำค่าประมาณ $\hat{\mu}$ และ $\hat{\lambda}$ ที่ได้ ประมาณค่าสัมประสิทธิ์ความเบ้ด้วยวิธีการที่ศึกษา

1) วิธีอย่างง่าย ตัวประมาณค่าสัมประสิทธิ์ความเบ้ คือ

$$\hat{\gamma}_{1\_S} = 3\sqrt{\frac{\hat{\mu}}{\hat{\lambda}}}$$

2) วิธีปรับค่าสัมประสิทธิ์ความเบ้ด้วยเทอมค่าคงที่ $(n-1)/n$ ตัวประมาณค่าสัมประสิทธิ์ความเบ้ คือ

$$\hat{\gamma}_{1\_C_1} = 3\left(\frac{n-1}{n}\right)\sqrt{\frac{\hat{\mu}}{\hat{\lambda}}}$$

3) วิธีปรับค่าสัมประสิทธิ์ความเบ้ด้วยเทอมค่าคงที่ $(n-2)/n$ ตัวประมาณค่าสัมประสิทธิ์ความเบ้ คือ

$$\hat{\gamma}_{1\_C_2} = 3\left(\frac{n-2}{n}\right)\sqrt{\frac{\hat{\mu}}{\hat{\lambda}}}$$

3.5 คำนวณค่าความเอนเอียงสัมบูรณ์และค่าคลาดเคลื่อนกำลังสองเฉลี่ยของค่าประมาณของวิธีการประมาณค่าแต่ละวิธี แล้วทำการเปรียบเทียบ

ค่าความเอนเอียงสัมบูรณ์ ($|Bias|$) ของค่าประมาณ คำนวณจาก

$$|Bias| = \left|\frac{\sum_{i=1}^{10,000}\hat{\gamma}_{1\_ki}}{10,000} - \gamma_1\right|$$

ค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ของค่าประมาณ คำนวณจาก

$$MSE = \frac{\sum_{i=1}^{10,000}(\gamma_1 - \hat{\gamma}_{1\_ki})^2}{10,000}$$

เมื่อ γ_1 แทน ค่าสัมประสิทธิ์ความเบ้ของประชากร

$\hat{\gamma}_{1\_ki}$ แทน ค่าประมาณของสัมประสิทธิ์ความเบ้ ซึ่งคำนวณจากวิธีที่ k ทำซ้ำครั้งที่ i

k แทน วิธีการประมาณค่าสัมประสิทธิ์ความเบ้

ผลการศึกษา

ในการนำเสนอผลการศึกษาเพื่อความสะดวกจะใช้สัญลักษณ์ต่างๆ แทนความหมายดังนี้

S หมายถึง วิธีอย่างง่าย

C_1 หมายถึง วิธีปรับค่าสัมประสิทธิ์ความเบ้ด้วยเทอมค่าคงที่ $(n-1)/n$

C_2 หมายถึง วิธีปรับค่าสัมประสิทธิ์ความเบ้ด้วยเทอมค่าคงที่ $(n-2)/n$

$|Bias|$ หมายถึง ค่าความเอนเอียงสัมบูรณ์

MSE หมายถึง ค่าคลาดเคลื่อนกำลังสองเฉลี่ย

ผลการเปรียบเทียบค่าความเอนเอียงสัมบูรณ์ และค่าคลาดเคลื่อนกำลังสองเฉลี่ยของค่าประมาณ นำเสนอดังตาราง 1 และ 2 ซึ่งมีรายละเอียดดังนี้

4.1 การเปรียบเทียบค่าความเอนเอียงสัมบูรณ์

ในทุกระดับของขนาดตัวอย่าง (n = 20, 40, 60, 80, 100 และ 200) ทุกระดับของค่าพารามิเตอร์ (λ = 1, 3, 5, 10, 15 และ 20) และทุกระดับของร้อยละของค่าสูญหาย (p = 0.05 และ 0.10) วิธีปรับค่าสัมประสิทธิ์ความเบ้ด้วยเทอมค่าคงที่ $(n-2)/n$ ให้ค่าความเอนเอียงสัมบูรณ์ ($|Bias|$) ต่ำที่สุด รองลงมาคือวิธีปรับค่าสัมประสิทธิ์ความเบ้ด้วยเทอมค่าคงที่ $(n-1)/n$ และวิธีอย่างง่าย ตามลำดับ

4.2 การเปรียบเทียบค่าคลาดเคลื่อนกำลังสองเฉลี่ย

ในทุกระดับของขนาดตัวอย่าง (n = 20, 40, 60, 80, 100 และ 200) ทุกระดับของค่าพารามิเตอร์ (λ = 1, 3, 5, 10, 15 และ 20) และทุกระดับของร้อยละของค่าสูญหาย ($p =$ 0.05 และ 0.10) วิธีปรับค่าสัมประสิทธิ์ความเบ้ด้วยเทอมค่าคงที่ $(n-2)/n$ ให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ต่ำที่สุด รองลงมาคือ วิธีปรับค่าสัมประสิทธิ์ความเบ้ด้วยเทอมค่าคงที่ $(n-1)/n$ และวิธีอย่างง่าย ตามลำดับ

วิจารณ์ผลการศึกษา

ในการประมาณค่าสัมประสิทธิ์ความเบ้ของการแจกแจงอินเวอร์สเกาส์เซียน เมื่อข้อมูลมีค่าสูญหาย ควรเลือกใช้วิธีปรับค่าสัมประสิทธิ์ความเบ้ด้วยเทอมค่าคงที่ $(n-2)/n$ เนื่องจากวิธีนี้เป็นวิธีที่ให้ค่าความเอนเอียงสัมบูรณ์ ($|Bias|$) และค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ต่ำที่สุดในทุกกรณีที่ศึกษา รองลงมาคือ วิธีปรับค่าสัมประสิทธิ์ความเบ้ด้วยเทอมค่าคงที่ $(n-1)/n$ และวิธีอย่างง่าย ตามลำดับ ซึ่งมีสูตรในประมาณค่าดังนี้

$$\hat{\gamma}_1 = 3\left(\frac{n-2}{n}\right)\sqrt{\frac{\hat{\mu}}{\hat{\lambda}}}$$

โดยที่ $\hat{\mu} = \bar{X}$ และ

$$\hat{\lambda} = \frac{n}{\left(\sum_{i=1}^{n} X_i^{-1} - (\bar{X})^{-1}\right)}$$

ตาราง 1 ค่าความเอนเอียงสัมบูรณ์ ($|Bias|$)และค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ของค่าประมาณสัมประสิทธิ์ความเบ้ของการแจกแจงอินเวอร์สเกาส์เซียน เมื่อ μ เท่ากับ 1 และ $p = 0.05$

n	λ	$\lvert Bias \rvert$			MSE		
		วิธี S	วิธี C_1	วิธี C_2	วิธี S	วิธี C_1	วิธี C_2
20	1	1.1086	0.9031	0.6977*	1.4023	0.9721	0.6272*
	3	1.6370	1.4686	1.3001*	2.7034	2.1779	1.7094*
	5	1.8580	1.6980	1.5380*	3.4608	2.8911	2.3726*
	10	2.1171	1.9638	1.8105*	4.4843	3.8585	3.2797*
	15	2.2448	2.0938	1.9428*	5.0400	4.3850	3.7755*
	20	2.3249	2.1751	2.0253*	5.4056	4.7315	4.1023*

ตาราง 1 ค่าความเอนเอียงสัมบูรณ์ ($|Bias|$)และค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ของค่าประมาณสัมประสิทธิ์ความเบ้ของการแจกแจงอินเวอร์สเกาส์เซียน เมื่อ μ เท่ากับ 1 และ $p = 0.05$ (ต่อ)

n	λ	$\lvert Bias \rvert$	MSE	0.9645*	1.4651	1.2268	1.0104*
		วิธี S	วิธี C_1	วิธี C_2	วิธี S	วิธี C_1	วิธี C_2
	5	1.9015	1.8204	1.7393*	3.6202	3.3183	3.0294*
	10	2.1569	2.0793	2.0017*	4.6536	4.3246	4.0077*
	15	2.2841	2.2077	2.1312*	5.2178	4.8743	4.5425*
	20	2.3644	2.2885	2.2127*	5.5908	5.2377	4.8962*
60	1	1.2004	1.1304	1.0604*	1.5011	1.3359	1.1806*
	3	1.7009	1.6437	1.5864*	2.9009	2.7093	2.5243*
	5	1.9159	1.8616	1.8073*	3.6735	3.4683	3.2691*
	10	2.1709	2.1189	2.0669*	4.7134	4.4903	4.2727*
	15	2.2974	2.2462	2.1950*	5.2786	5.0459	4.8185*
	20	2.3776	2.3268	2.2759*	5.6530	5.4140	5.1801*
80	1	1.2110	1.1584	1.1057*	1.5123	1.3864	1.2661*
	3	1.7095	1.6665	1.6235*	2.9283	2.7829	2.6412*
	5	1.9232	1.8823	1.8415*	3.7008	3.5454	3.3934*
	10	2.1782	2.1391	2.1000*	4.7449	4.5762	4.4105*
	15	2.3043	2.2658	2.2273*	5.3101	5.1343	4.9613*
	20	2.3839	2.3457	2.3075*	5.6830	5.5024	5.3247*
100	1	1.2166	1.1745	1.1323*	1.5161	1.4146	1.3166*
	3	1.7143	1.6798	1.6454*	2.9436	2.8265	2.7118*
	5	1.9280	1.8953	1.8626*	3.7189	3.5939	3.4710*
	10	2.1817	2.1504	2.1190*	4.7601	4.6245	4.4908*
	15	2.3082	2.2774	2.2465*	5.3280	5.1866	5.0472*
	20	2.3879	2.3574	2.3268*	5.7024	5.5572	5.4140*
200	1	1.2284	1.2072	1.1861*	1.5273	1.4756	1.4248*
	3	1.7223	1.7050	1.6877*	2.9686	2.9094	2.8508*
	5	1.9364	1.9200	1.9036*	3.7505	3.6873	3.6246*
	10	2.1895	2.1738	2.1581*	4.7942	4.7258	4.6578*
	15	2.3157	2.3003	2.2848*	5.3627	5.2914	5.2205*
	20	2.3955	2.3802	2.3649*	5.7386	5.6654	5.5927*

** หมายถึง วิธีการประมาณค่าที่ให้ค่า* $|Bias|$ *ต่ำที่สุด หรือค่า* MSE *ต่ำที่สุด*

ตาราง 2 ค่าความเอนเอียงสัมบูรณ์ ($|Bias|$) และค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ของค่าประมาณสัมประสิทธิ์ความเบ้ของการแจกแจงอินเวอร์สเกาส์เซียน เมื่อ μ เท่ากับ 1 และ $p = 0.10$

n	λ	$\lvert Bias \rvert$			MSE		
		วิธี S	วิธี C_1	วิธี C_2	วิธี S	วิธี C_1	วิธี C_2
20	1	1.0976	0.8928	0.6879*	1.3913	0.9653	0.6242*
	3	1.6359	1.4675	1.2991*	2.7013	2.1762	1.7080*
	5	1.8577	1.6977	1.5378*	3.4605	2.8909	2.3724*
	10	2.1174	1.9641	1.8108*	4.4860	3.8600	3.2811*
	15	2.2442	2.0932	1.9423*	5.0374	4.3826	3.7734*
	20	2.3246	2.1749	2.0251*	5.4045	4.7306	4.1015*
40	1	1.1702	1.0659	0.9617*	1.4639	1.2261	1.0102*
	3	1.6843	1.5989	1.5135*	2.8499	2.5689	2.3024*
	5	1.9006	1.8195	1.7385*	3.6172	3.3154	3.0268*
	10	2.1571	2.0795	2.0018*	4.6545	4.3255	4.0085*
	15	2.2839	2.2075	2.1310*	5.2170	4.8735	4.5418*
	20	2.3642	2.2883	2.2125*	5.5899	5.2368	4.8953*
60	1	1.1967	1.1267	1.0568*	1.5001	1.3353	1.1804*
	3	1.7016	1.6443	1.5871*	2.9042	2.7124	2.5272*
	5	1.9159	1.8616	1.8073*	3.6739	3.4687	3.2694*
	10	2.1705	2.1185	2.0665*	4.7118	4.4888	4.2712*
	15	2.2980	2.2468	2.1956*	5.2812	5.0484	4.8209*
	20	2.3771	2.3263	2.2755*	5.6507	5.4117	5.1780*
80	1	1.2107	1.1581	1.1055*	1.5164	1.3904	1.2701*
	3	1.7078	1.6648	1.6218*	2.9232	2.7780	2.6365*
	5	1.9232	1.8824	1.8416*	3.7013	3.5459	3.3939*
	10	2.1774	2.1383	2.0992*	4.7415	4.5729	4.4073*
	15	2.3041	2.2656	2.2271*	5.3092	5.1334	4.9604*
	20	2.3840	2.3458	2.3076*	5.6835	5.5029	5.3252*

* หมายถึง วิธีการประมาณค่าที่ให้ค่า $|Bias|$ ต่ำที่สุด หรือค่า MSE ต่ำที่สุด

ตาราง 2 ค่าความเอนเอียงสัมบูรณ์ ($|Bias|$) และค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ของค่าประมาณสัมประสิทธิ์ความเบ้ของการแจกแจงอินเวอร์สเกาส์เซียน เมื่อ μ เท่ากับ 1 และ $p = 0.10$ (ต่อ)

n	λ	$\lvert Bias \rvert$			MSE		
		วิธี S	วิธี C_1	วิธี C_2	วิธี S	วิธี C_1	วิธี C_2
100	1	1.2126	1.1705	1.1284*	1.5105	1.4093	1.3117*
	3	1.7130	1.6785	1.6441*	2.9395	2.8225	2.7080*
	5	1.9274	1.8947	1.8621*	3.7170	3.5920	3.4692*
	10	2.1813	2.1500	2.1187*	4.7585	4.6229	4.4893*
	15	2.3078	2.2770	2.2462*	5.3264	5.1850	5.0456*
	20	2.3876	2.3571	2.3265*	5.7010	5.5559	5.4126*
200	1	1.2300	1.2088	1.1877*	1.5336	1.4818	1.4309*
	3	1.7226	1.7053	1.6880*	2.9699	2.9107	2.8520*
	5	1.9364	1.9200	1.9036*	3.7505	3.6873	3.6246*
	10	2.1897	2.1740	2.1583*	4.7951	4.7266	4.6586*
	15	2.3160	2.3006	2.2851*	5.3641	5.2928	5.2219*
	20	2.3953	2.3800	2.3647*	5.7376	5.6644	5.5917*

* หมายถึง วิธีการประมาณค่าที่ให้ค่า $|Bias|$ ต่ำที่สุด หรือค่า MSE ต่ำที่สุด

ตัวอย่างการประยุกต์ใช้

การประมาณค่าสัมประสิทธิ์ความเบ้ของการแจกแจงอินเวอร์สเกาส์เซียน เมื่อข้อมูลมีค่าสูญหายสามารถนำไปประยุกต์ใช้ได้ดังนี้

ในทดสอบช่วงชีวิต (Life test) ของลูกบอล โดยเก็บรวบรวมข้อมูลความคงทนของลูกบอล จำนวน 23 ลูก (Pavur *et al.*, 1992) โดยสมมติว่าข้อมูลตัวสุดท้ายเป็นค่าสูญหาย (Not Available: NA) ข้อมูลมีดังนี้

17.88	51.96	93.12	28.92
54.12	98.64	33.00	55.56
105.12	41.52	67.80	105.84
42.12	68.64	127.92	45.60
68.64	127.92	48.48	68.88
173.40	51.84	NA	

จากข้อมูลข้างต้นทำการประมาณค่าข้อมูลสูญหาย โดยการเลือกข้อมูลมาจำนวน 1 ค่า อย่างสุ่มจากข้อมูลที่เก็บรวบรวมได้โดยให้เครื่อง

คอมพิวเตอร์เลือกข้อมูลมา 1 ค่าอย่างสุ่ม ค่าประมาณของข้อมูลสูญหาย คือ $\hat{X}_{23} = 42.12$

จากนั้นประมาณค่าพารามิเตอร์ของการแจกแจงอินเวอร์สเกาส์เซียน ดังนี้

$$\hat{\mu} = n^{-1}\sum_{i=1}^{n} X_i = 70.39304$$

$$\hat{\lambda} = \frac{n}{\left(\sum_{i=1}^{n} X_i^{-1} - (\bar{X})^{-1}\right)} = 55.37035$$

ขั้นต่อไปทำการประมาณค่าสัมประสิทธิ์ความเบ้โดยใช้วิธีปรับค่าสัมประสิทธิ์ความเบ้ด้วยเทอมค่าคงที่ $(n-2)/n$ ดังนี้

$$\hat{\gamma}_{1\_C_2} = 3\left(\frac{n-2}{n}\right)\sqrt{\frac{\hat{\mu}}{\hat{\lambda}}} = 3\left(\frac{21}{23}\right)\sqrt{\frac{70.39304}{55.37035}} = 3.08844$$

ข้อเสนอแนะ

ในการศึกษาครั้งนี้ประมาณค่าข้อมูลสูญหายโดยการเลือกข้อมูลอย่างสุ่มจากข้อมูลที่เก็บรวบรวมได้ ดังนั้นควรศึกษาเพิ่มเติมในกรณีที่ใช้การประมาณค่าข้อมูลสูญหายวิธีอื่น เช่น การประมาณค่าข้อมูลสูญหายด้วยค่าเฉลี่ยหรือมัธยฐาน เป็นต้น

กิตติกรรมประกาศ

ผู้วิจัยขอขอบพระคุณผู้ทรงคุณวุฒิที่ได้ให้คำแนะนำและข้อเสนอแนะ ซึ่งเป็นประโยชน์อย่างยิ่งในการปรับปรุงแก้ไขต้นฉบับให้สมบูรณ์มากยิ่งขึ้น ผู้วิจัยขอขอบพระคุณมา ณ โอกาสนี้

เอกสารอ้างอิง

วราฤทธิ์ พานิชกิจโกศลกุล. 2550. วิธีประมาณค่าฐานนิยมของข้อมูลที่มีการแจกแจงอินเวอร์สเกาส์เซียน. **วารสารวิชาการและวิจัย มทร.พระนคร** 1(2): 179-185.

วราฤทธิ์ พานิชกิจโกศลกุล. 2552. การจำลองแบบมอนติคาร์โลสำหรับประมาณค่าโมเมนต์อันดับที่สองของการแจกแจงอินเวอร์สเกาส์เซียน. **วารสารนเรศวร พะเยา** 2(1): 19-23.

Arefi, M., Mohtashami, G.R. and Vaghei, Y. 2008. A note on interval estimation for the mean of inverse Gaussian distribution. **Statistics and Operations Research Transactions** 32(1): 49-56.

Balakrisnan, N. and Chen, W.S. 1997. **Tables for Order Statistics from Inverse Gaussian Distributions with Applications.** CRC Press, Florida.

Balakrisnan, N. and Nevzorov, V.B. 2003. **A Primer on Statistical Distribution.** John Wiley & Sons, New Jersey.

Chhikara, R.S. and Folks J.L. 1989. **The Inverse Gaussian Distribution: Theory, Methodology and Applications.** Marcel Dekker, New York.

Cohen, A.C. and Whitten, B.J. 1988. **Parameter Estimation in Reliability and Life Span Models.** Marcel Dekker, New York.

Evans, M., Hastings, N., and Peacock, B. 2000**. Statistical Distributions.** John Wiley & Sons, New York.

Gupta, R.P. 1973. Maximum likelihood estimate of the parameters of a truncated inverse Gaussian distribution. **Metrika** 20(1): 51-53.

Krishnamoorthy, K. 2006. **Handbook of Statistical Distributions with Applications.** Chapman & Hall, Florida.

Lee, Y.M. and Tang, J. 2007. A modified EM-algorithm for estimating the parameters of inverse Gaussian distribution based on time-censored wiener degradation data. **Statistica Sinica** 17: 873-893.

Pavur, R.J., Edgeman, R.L. and Scott, R.C. 1992. Quadratic statistics for the goodness-of-fit test of the inverse Gaussian distribution. **IEEE Transactions on Reliability** 41:118-123.

Sanchez, J. and He, Y. 2003. **Internet data analysis for the undergraduate statistics curriculum.** Available Source: http://preprints.stat.ucla.edu/370/ams2004.pdf, May 11, 2009.

Schrodinger, E. 1915. Zur Theorie der Fall-und Steigversuche an Teilchen mit Brownscher Bewegung, **Physikalische Zeitschrift** 16:289-295.

Seshadri, V. 1994. **Inverse Gaussian Distributions.** Oxford University Press, Oxford.

Seshadri, V. 1999. **The Inverse Gaussian Distribution: Statistical Theory and Applications.** Springer-Verlag, New York.

Shuster, J. 1968. On the inverse Gaussian distribution. **Journal of the American Statistical Association** 63: 1514-1516.

Tweedie, M.C.K. 1957. Statistical properties of inverse Gaussian distributions. **Annals of Mathematical Statistics** 28:362-377.