ตัวประมาณเบส์และการสุ่มตัวอย่างแบบกิบส์
Bayes Estimator and Gibbs Sampling

รศ.ร.อ.มานพ วราภักดิ์ *

บทคัดย่อ

การประมาณค่าของพารามิเตอร์ในฟังก์ชั่นการแจกแจงความน่าจะเป็น เป็นเรื่องสำคัญเรื่องหนึ่งของการอนุมานเชิงสถิติ ซึ่งอาจจะทำการประมาณค่าแบบจุดหรือประมาณค่าแบบช่วง การประมาณค่าพารามิเตอร์แบบจุดมีหลายวิธีการ วิธีเบส์เป็นวิธีหนึ่งที่ใช้กันมาก แต่หลายกรณีจะมีความยากมากที่จะหาตัวประมาณเบส์หรือค่าประมาณเบส์ด้วยวิธีการเชิงวิเคราะห์ เช่น แคลคูลัส และพีชคณิตเชิงเส้น จำเป็นต้องใช้วิธีการประมาณ เช่น วิธีเชิงตัวเลข และวิธีการจำลอง เป็นต้น ดังนั้นในบทความนี้ขอนำเสนอการหาค่าประมาณเบส์ด้วยวิธีการประมาณที่เป็นการจำลองมอนติคาร์โล (Monte Carlo simulation) โดยใช้วิธีการสุ่มตัวอย่างหรือการจำลองข้อมูลแบบที่เรียกว่าการสุ่มตัวอย่างแบบกิบส์ (Gibbs sampling) ซึ่งเป็นวิธีการหนึ่งที่รู้จักกันดีในวิธีมอนติคาร์โลลูกโซ่มาร์คอฟ (Markov chain Monte Carlo (MCMC) methods)

คำสำคัญ :

การประมาณค่า ตัวประมาณเบส์ ฟังก์ชันความเสียหาย การจำลองมอนติคาร์โล การสุ่มตัวอย่างแบบกิบส์

---

* รองศาสตราจารย์ประจำคณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย

Abstract

Estimation of parameters of a probability distribution function is one of the major tasks in statistical inference. The estimation can be divided into two parts, point estimation and interval estimation. There are many point estimation methods and one of these methods is the Bayesian method. The Bayesian method is one of the widely used methods, but in some cases the method is very difficult to find the Bayes estimators or Bayes estimates through the analytical methods such as calculus and linear algebra. So, for the hard problem, we frequently use approximation methods such as numerical method and simulation method to find out the approximate Bayes estimates. This article presents an approximation method called the Monte Carlo simulation using the Gibbs sampling which is a well-known method in the Markov chain Monte Carlo (MCMC) methods.

Keywords: estimation, Bayes estimator, loss function, Monte Carlo simulation, Gibbs sampling.

1. บทนำ

การอนุมานเชิงสถิติด้วยวิธีการของเบส์มีแนวคิดต่างจากวิธีการสถิติแบบฉบับทั่วไป คือ พิจารณาพารามิเตอร์ที่ทำการอนุมานเป็นตัวแปรสุ่ม เพราะฉะนั้นพารามิเตอร์มีการแจกแจงความน่าจะเป็น ซึ่งอาจกำหนดขึ้นโดยประสบการณ์ ความเชื่อ หรือข้อมูลข้างเคียง การแจกแจงที่กำหนดนี้อาจยังไม่เหมาะสมที่จะใช้ประโยชน์ ดังนั้น จะใช้ข้อมูลหรือค่าสังเกตใหม่ที่ได้มาปรับแก้การแจกแจงแรกเริ่มข้างต้น ได้การแจกแจงภายหลังและใช้ประโยชน์ต่อไปในการอนุมานเชิงสถิติ เช่น การประมาณค่าของพารามิเตอร์ และการทดสอบสมมติฐานเชิงสถิติ สำหรับในบทความนี้ กล่าวถึงการอนุมานเชิงสถิติในเรื่องการประมาณค่าของพารามิเตอร์ที่เป็นการประมาณค่าแบบจุดด้วยวิธีการเบส์ ในกรณีอย่างง่ายและกรณีอย่างยากซึ่งในกรณีนี้จะใช้วิธีการจำลองในการหาค่าประมาณเบส์

2. ตัวประมาณเบส์

ให้ X เป็นตัวแปรสุ่มประเภทต่อเนื่องหรือไม่ต่อเนื่อง โดยมี θ เป็นพารามิเตอร์ในฟังก์ชันการแจกแจงของ X และ θ เป็นค่าเป็นไปได้ของพารามิเตอร์สุ่ม Θ (ประเภทต่อเนื่องหรือไม่ต่อเนื่อง นั่นคือ พารามิเตอร์ในการแจกแจงเป็นตัวแปรสุ่ม) เพราะฉะนั้น การแจกแจงความน่าจะเป็นของ X เป็นการแจกแจงแบบมีเงื่อนไข สมมติ $X|\Theta=\theta \sim f_{X|\Theta}(x|\theta)$ และสมมติ Θ มีฟังก์ชันความหนาแน่น (density function) $f_{\Theta}(\theta)$ ให้ $X_1, X_2, ..., X_n$ เป็นตัวอย่างสุ่มจาก $X|\Theta=\theta \sim f_{X|\Theta}(x|\theta)$ เพราะฉะนั้น ได้ฟังก์ชันความหนาแน่นร่วม

$$f_{\underline{X},\Theta}(\underline{x},\theta) = f_{\underline{X}|\Theta}(\underline{x}|\theta)\, f_{\Theta}(\theta) = f_{X|\Theta}(x_1|\theta)\cdots f_{X|\Theta}(x_n|\theta) f_{\Theta}(\theta)$$

เมื่อทราบข้อมูล $\underline{X} = (X_1, \ldots, X_n) = \underline{x} = (x_1, \ldots, x_n)$ จะนำมาใช้ประโยชน์ปรับ $f_{\Theta}(\theta)$ ได้ฟังก์ชันความหนาแน่นแบบมีเงื่อนไขของ Θ เมื่อกำหนด $\underline{X} = \underline{x}$ ดังนี้ โดยทฤษฎีบทเบส์หรือหลักเกณฑ์ของเบส์ (Bayes' rule)

$$f_{\Theta|\underline{X}}(\theta|\underline{x}) = \frac{f_{\underline{X}|\Theta}(\underline{x}|\theta) f_{\Theta}(\theta)}{f_{\underline{X}}(\underline{x})} \tag{1}$$

โดยที่ $f_{\underline{X}}(\underline{x}) = \int_{-\infty}^{\infty} f_{\underline{X}|\Theta}(\underline{x}|\theta) f_{\Theta}(\theta)\, d\theta$ กรณี Θ เป็นตัวแปรสุ่มต่อเนื่อง

$= \sum_{\theta=-\infty}^{\infty} f_{\underline{X}|\Theta}(\underline{x}|\theta) f_{\Theta}(\theta)$ กรณี Θ เป็นตัวแปรสุ่มไม่ต่อเนื่อง

ฟังก์ชันความหนาแน่น f_{Θ} และ $f_{\Theta|\underline{X}}$ ของ Θ มีชื่อเรียกเฉพาะว่า "ฟังก์ชันความหนาแน่นก่อน" (prior density function) และ "ฟังก์ชันความหนาแน่นภายหลัง" (posterior density function) ตามลำดับ ทั้งนี้เนื่องจาก f_{Θ} เป็นฟังก์ชันก่อนที่จะได้ข้อมูล $\underline{X}$ หรือก่อนทำการทดลองหาข้อมูล แต่เมื่อได้ข้อมูลหรือค่าสังเกต $\underline{x}$ ของ $\underline{X}$ แล้ว จะนำมาปรับ f_{Θ} ได้ฟังก์ชัน $f_{\Theta|\underline{X}}$ จึงเรียก f_{Θ} ว่าฟังก์ชันความหนาแน่นก่อน และเรียก $f_{\Theta|\underline{X}}$ ว่าฟังก์ชันความหนาแน่นภายหลัง และจะนำ $f_{\Theta|\underline{X}}$ ไปใช้ประโยชน์ในการประมาณค่าพารามิเตอร์ θ

การกำหนดรูปแบบฟังก์ชัน f_{Θ} ซึ่งเป็นรูปแบบการแจกแจงก่อนหรือแรกเริ่มของ Θ อาจจะกำหนดจากประสบการณ์ ความเชื่อ ดุลยพินิจ ข้อมูลข้างเคียง หรือข้อมูลแรกเริ่มที่มักจะมีไม่มากนัก เพราะฉะนั้น เมื่อได้ข้อมูลหรือค่าสังเกตของ $\underline{X}$ จึงควรนำมาปรับแก้ f_{Θ}

หลักการของเบส์ในการหาตัวประมาณของ θ แบบจุด คือ พิจารณาความเสียหายที่จะเกิดขึ้นจากการเลือกตัวประมาณ $\delta(\underline{X})$ เป็นตัวประมาณของ θ ให้ $L[\theta, \delta(\underline{X})]$ เป็นฟังก์ชันความเสียหาย (loss function) ที่เกิดจากการเลือก $\delta(\underline{X})$ เป็นตัวประมาณของ θ ด้วยวิธีการเบส์จะเลือก $\delta(\underline{X})$ ที่ให้ได้ค่าคาดหมายของฟังก์ชันความเสียหาย $E[L(\Theta, \delta(\underline{X}))\,|\,\underline{x}] = E[L(\Theta, \delta(\underline{X}))\,|\,\underline{X} = \underline{x}]$ มีค่าต่ำสุด และเรียก $\delta(\underline{X})$ ที่ได้ว่า **ตัวประมาณเบส์** (Bayes estimator) ซึ่งรูปแบบจะขึ้นอยู่กับรูปแบบของฟังก์ชันความเสียหาย $L[\theta, \delta(\underline{X})]$ รูปแบบของ $L[\theta, \delta(\underline{X})]$ ที่นิยมกำหนดกันคือ "ฟังก์ชันความเสียหายแบบค่าคลาดเคลื่อนกำลังสอง" (squared error loss function) $L[\theta, \delta(\underline{X})] = [\delta(\underline{X}) - \theta]^2$ และรองลงไปคือ "ฟังก์ชันความเสียหายแบบค่าคลาดเคลื่อนสัมบูรณ์" (absolute error loss function) $L[\theta, \delta(\underline{X})] = = |\delta(\underline{X}) - \theta|$

ในกรณี $L[\theta, \delta(\underline{X})] = [\delta(\underline{X}) - \theta]^2$ ได้ว่า

$$E[L(\Theta, \delta(\underline{X}))\,|\,\underline{x}] = E[\{\delta(\underline{X}) - \Theta\}^2\,|\,\underline{x}]$$

$$= E[\{\delta(\underline{X}) - E(\Theta|\underline{x}) + E(\Theta|\underline{x}) - \Theta\}^2 | \underline{x}]$$
$$= E[\{\delta(\underline{X}) - E(\Theta|\underline{x})\}^2 | \underline{x}] + E[\{E(\Theta|\underline{x}) - \Theta\}^2 | \underline{x}]$$

[เนื่องจาก $E[\{\delta(\underline{X}) - E(\Theta|\underline{x})\}\{E(\Theta|\underline{x}) - \Theta\}|\underline{x}] = [\delta(\underline{x}) - E(\Theta|\underline{x})][E(\Theta|\underline{x}) - E(\Theta|\underline{x})] = 0$]

$$= [\delta(\underline{x}) - E(\Theta|\underline{x})]^2 + E[\{\Theta - E(\Theta|\underline{x})\}^2 | \underline{x}] \qquad (2)$$

จากสมการ (2) พบว่า ไม่สามารถควบคุมค่าของเทอมที่สองทางขวาของสมการได้ แต่ควบคุมค่าของเทอมแรกได้ ด้วยการเลือกค่าของ $\delta(\underline{x})$ และเนื่องจากค่าของเทอมแรกค่ามากกว่าหรือเท่ากับศูนย์เสมอ เพราะฉะนั้น ค่าของ $E[\{\delta(\underline{X}) - \Theta\}^2 | \underline{x}]$ จะมีค่าต่ำสุดก็ต่อเมื่อเทอมแรก $[\delta(\underline{x}) - E(\Theta|\underline{x})]^2 = 0$ ซึ่งก็ต่อเมื่อ $\delta(\underline{x}) = E(\Theta|\underline{x})$ นั่นคือ ได้ตัวประมาณเบส์เป็น "ค่าคาดหมายแบบมีเงื่อนไข" ของการแจกแจงภายหลังของ Θ เมื่อกำหนด $\underline{X}$ ดังนี้ กรณีฟังก์ชันความเสียหายอยู่ในรูปแบบของค่าคลาดเคลื่อนกำลังสอง

$$\delta(\underline{X}) = E(\Theta| \underline{X}) \qquad (3)$$
$$= \int_{-\infty}^{\infty} \theta f_{\Theta|\underline{X}}(\theta|\underline{X})\, d\theta \quad \text{กรณี } \Theta \text{ เป็นตัวแปรสุ่มต่อเนื่อง}$$
$$= \sum_{\theta=-\infty}^{\infty} \theta p_{\Theta|\underline{X}}(\theta| \underline{X}) \quad \text{กรณี } \Theta \text{ เป็นตัวแปรสุ่มไม่ต่อเนื่อง}$$

ในทำนองเดียวกัน ถ้าประมาณ $h(\theta)$ (ฟังก์ชันของ θ) ด้วยตัวประมาณเบส์ โดยกำหนดฟังก์ชันความเสียหายในรูปแบบค่าคลาดเคลื่อนกำลังสอง จะได้ตัวประมาณเบส์ดังนี้

$$\delta(\underline{X}) = E[h(\Theta)| \underline{X}] \qquad (4)$$
$$= \int_{-\infty}^{\infty} h(\theta) f_{\Theta|\underline{X}}(\theta|\underline{X})\, d\theta \quad \text{กรณี } \Theta \text{ เป็นตัวแปรสุ่มต่อเนื่อง}$$
$$= \sum_{\theta=-\infty}^{\infty} h(\theta) p_{\Theta|\underline{X}}(\theta| \underline{X}) \quad \text{กรณี } \Theta \text{ เป็นตัวแปรสุ่มไม่ต่อเนื่อง}$$

ในกรณีฟังก์ชันความเสียหายอยู่ในรูปแบบค่าคลาดเคลื่อนสัมบูรณ์ $|\delta(\underline{X}) - \theta|$ แสดงได้ว่าตัวประมาณเบส์ $\delta(\underline{X})$ ของ θ คือ "มัธยฐาน" (median) ของการแจกแจงภายหลังของ Θ เมื่อกำหนด $\underline{X}$ นั่นคือ ค่าประมาณเบส์ $\delta(\underline{x})$ เป็นค่าประมาณที่สอดคล้องสมการ (หาได้จากสมการ)

$$P[\Theta \le \delta(\underline{x})| \underline{x}] \ge 0.5 \text{ และ } P[\Theta \ge \delta(\underline{x})| \underline{x}] \ge 0.5 \qquad (5)$$

ซึ่งถ้า Θ เป็นตัวแปรสุ่มต่อเนื่อง ค่าประมาณเบส์ $\delta(\underline{x})$ สอดคล้องสมการ

$$\int_{-\infty}^{\delta(\underline{x})} f_{\Theta|\underline{X}}(\theta|\underline{x})\,d\theta = \frac{1}{2} \tag{6}$$

ตัวอย่างที่ 1 ให้ $X_1, X_2, \ldots, X_n$ เป็นตัวอย่างสุ่มจาก $X|\Theta=\theta \sim Po(\theta)$ และ $\Theta \sim G(\alpha,\lambda)$ ซึ่งทราบค่า α และ λ จะหาตัวประมาณเบส์ของ θ เมื่อกำหนดฟังก์ชันความเสียหายเป็นแบบค่าคลาดเคลื่อนกำลังสอง

จากข้อกำหนด ได้ฟังก์ชันความหนาแน่นร่วมผสม

$$\begin{aligned}
f_{\underline{X},\Theta}(\underline{x},\theta) &= p_{\underline{X}|\Theta}(\underline{x}|\theta)f_{\Theta}(\theta)\\
&= \left[\frac{e^{-\theta}\theta^{x_1}}{x_1!}\cdots\frac{e^{-\theta}\theta^{x_n}}{x_n!}\right]\left[\frac{\lambda^{\alpha}}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\lambda\theta}\right]\\
&= \left[\frac{e^{-n\theta}\theta^{\sum_{i=1}^{n}x_i}}{\prod_{i=1}^{n}(x_i!)}\right]\left[\frac{\lambda^{\alpha}}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\lambda\theta}\right]\\
&= \frac{\lambda^{\alpha}}{\prod_{i=1}^{n}(x_i!)\Gamma(\alpha)}\theta^{\sum_{i=1}^{n}x_i+\alpha-1}e^{-(n+\lambda)\theta},\quad x_i = 0,1,\ldots,n;\ \theta \ge 0
\end{aligned}$$

เพราะฉะนั้น ได้ฟังก์ชันความน่าจะเป็นขอบ $p_{\underline{X}}$ ของ $\underline{X}$ เป็น

$$\begin{aligned}
p_{\underline{X}}(\underline{x}) = \int_0^{\infty} f_{\underline{X},\Theta}(\underline{x},\theta)\,d\theta &= \frac{\lambda^{\alpha}}{\prod_{i=1}^{n}(x_i!)\Gamma(\alpha)}\int_0^{\infty}\theta^{\sum_{i=1}^{n}x_i+\alpha-1}e^{-(n+\lambda)\theta}d\theta\\
&= \frac{\lambda^{\alpha}}{\prod_{i=1}^{n}(x_i!)\Gamma(\alpha)}\times\frac{\Gamma\left(\sum_{i=1}^{n}x_i+\alpha\right)}{(n+\lambda)^{\sum_{i=1}^{n}x_i+\alpha}},\quad x_i = 0,1,\ldots,n
\end{aligned}$$

และดังนั้น ได้ฟังก์ชันความหนาแน่นภายหลัง $f_{\Theta|\underline{X}}$ ของ Θ เมื่อกำหนด $\underline{X}=\underline{x}$ ดังนี้

$$f_{\Theta|\underline{X}}(\theta|\underline{x}) = \frac{p_{\underline{X}|\Theta}(\underline{x}|\theta)f_{\Theta}(\theta)}{p_{\underline{X}}(\underline{x})}$$

$$= \frac{(n+\lambda)^{\sum_{i=1}^{n}x_i+\alpha}}{\Gamma\left(\sum_{i=1}^{n}x_i+\alpha\right)}\theta^{\sum_{i=1}^{n}x_i+\alpha-1} e^{-(n+\lambda)\theta}, \quad \theta \geq 0$$

ซึ่งเป็นฟังก์ชันความหนาแน่นของการแจกแจงแกมมา โดยมี $\sum_{i=1}^{n} x_i + \alpha$ และ $n + \lambda$ เป็น พารามิเตอร์ นั่นคือ ได้ $\Theta|\underline{X} = \underline{x} \sim G\left(\sum_{i=1}^{n} x_i + \alpha, n + \lambda\right)$

เมื่อกำหนดฟังก์ชันความเสียหายเป็นฟังก์ชันของค่าเคลื่อนกำลังสอง ได้ตัวประมาณเบส์เป็นค่าเฉลี่ยของ $\Theta|\underline{X}$ คือค่าเฉลี่ยของการแจกแจงแกมมา $G\left(\sum_{i=1}^{n} x_i + \alpha, n + \lambda\right)$:

$$\delta(\underline{X}) = \hat{\theta} = \frac{\sum_{i=1}^{n} X_i + \alpha}{n + \lambda}$$

◄

เนื่องจากฟังก์ชันความหนาแน่นขอบหรือฟังก์ชันความน่าจะเป็นขอบ $f_{\underline{X}}(\underline{x})$ ของ $\underline{X}$ ไม่ขึ้นอยู่กับพารามิเตอร์ θ (ไม่เป็นฟังก์ชันของ θ) ดังนั้น สามารถเขียน $f_{\Theta|\underline{X}}(\theta|\underline{x})$ ได้เป็น

$$f_{\Theta|\underline{X}}(\theta|\underline{x}) = c(\underline{x})f_{\underline{X}|\Theta}(\underline{x}|\theta)f_{\Theta}(\theta)$$

โดยพิจารณา $f_{\underline{X}}(\underline{x})$ เป็นค่าคงที่ แทนด้วย $c(\underline{x})$ และเมื่อกำหนดเทอมประกอบต่าง ๆ ที่ไม่ขึ้นอยู่กับ θ เป็นค่าคงที่ด้วย แทนทั้งหมดด้วย $c(\underline{x})$ เป็นตัวประกอบที่ไม่ขึ้นอยู่กับ θ [$c(\underline{x})$ เป็นตัวประกอบหรือค่าคงที่ที่ทำให้ $f_{\Theta|\underline{X}}(\theta|\underline{x})$ เป็นฟังก์ชันความหนาแน่นหรือฟังก์ชันความน่าจะเป็น] เพราะฉะนั้น บ่อยครั้งจะเขียน $f_{\Theta|\underline{X}}(\theta|\underline{x})$ เป็นสัดส่วนกับ $f_{\underline{X}|\Theta}(\underline{x}|\theta)f_{\Theta}(\theta)$ นั่นคือ

$$f_{\Theta|\underline{X}}(\theta|\underline{x}) \propto f_{\underline{X}|\Theta}(\underline{x}|\theta)f_{\Theta}(\theta) \tag{7}$$

เมื่อมีความเพียงพอในการอธิบายหรือสามารถระบุการแจกแจงของ $\Theta|\underline{X}$ ได้ (หรือเมื่อมีความเพียงพอที่จะใช้ประโยชน์ต่อไป) ตัวอย่างเช่น จากตัวอย่างที่ 1 สามารถเขียน $f_{\Theta|\underline{X}}(\theta|\underline{x})$ เป็น

$$f_{\Theta|\underline{X}}(\theta|\underline{x}) \propto p_{\underline{X}|\Theta}(\underline{x}|\theta) f_{\Theta}(\theta)$$

ซึ่งได้ $$f_{\Theta|\underline{X}}(\theta|\underline{x}) \propto \theta^{\sum_{i=1}^{n} x_i + \alpha - 1} e^{-(n+\lambda)\theta}$$

เข้าลักษณะของฟังก์ชันความหนาแน่นของการแจกแจงแกมมา $G(\sum_{i=1}^{n} x_i + \alpha, n+\lambda)$ ฉะนั้น

$$\Theta|\underline{X}=\underline{x} \sim G\left(\sum_{i=1}^{n} x_i + \alpha, n+\lambda\right) \text{ [ในตัวอย่างนี้ } c(\underline{x}) = \frac{(n+\lambda)^{\sum_{i=1}^{n} x_i + \alpha}}{\Gamma\left(\sum_{i=1}^{n} x_i + \alpha\right)} \text{]}$$

ตัวอย่างที่ 2 สมมติค่าสินไหมทดแทนรวมต่อปีมีการแจกแจงแบบปกติ โดยมีค่าเฉลี่ย θ และความแปรปรวน σ^2 เป็นค่าคงที่ทราบค่า แต่ค่าเฉลี่ยเป็นตัวแปรสุ่ม ซึ่ง θ เป็นค่าเป็นไปได้ของพารามิเตอร์สุ่ม (ตัวแปรสุ่ม) Θ สมมติว่า Θ มีการแจกแจงปกติด้วยค่าเฉลี่ย μ และความแปรปรวน λ และทราบค่า μ และ λ (อาจเป็นค่าประมาณโดยใช้ข้อมูลในอดีต) จากตัวอย่างสุ่มค่าสินไหมทดแทนรวม $X_1, X_2, \ldots, X_n$ ใน n ปีที่ผ่านมา จะนำมาใช้ประโยชน์ในการประมาณค่าสินไหมทดแทนรวม X_{n+1} ในปีที่ $n+1$ ด้วยค่าคาดหมาย $E(X_{n+1}|\Theta=\theta)$ ซึ่งจะขึ้นอยู่กับค่าพารามิเตอร์ θ ไม่ทราบค่า ดังนั้น จะประมาณ θ ด้วยวิธีการเบส์ เมื่อกำหนดฟังก์ชันความเสียหายเป็นฟังก์ชันค่าคลาดเคลื่อนสัมบูรณ์

จากสมการ (1) เขียนในรูปสัดส่วน:

$$f_{\Theta|\bar{X}}(\theta|\bar{x}) \propto f_{\bar{X}|\Theta}(\bar{x}|\theta) f_{\Theta}(\theta)$$

เนื่องจาก $X_1, X_2, \ldots, X_n$ เป็นตัวอย่างสุ่มจาก $X|\Theta=\theta \sim N(\theta, \sigma^2)$ และกำหนด $\Theta \sim N(\mu, \lambda)$ เพราะฉะนั้น

$$\begin{aligned} f_{\bar{X}|\Theta}(\bar{x}|\theta) f_{\Theta}(\theta) &= \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\theta)^2} \times \frac{1}{\sqrt{2\pi}\sqrt{\lambda}} e^{-\frac{1}{2\lambda}(\theta-\mu)^2} \\ &\propto \exp\left[-\frac{1}{2\sigma^2}\left(\sum_{i=1}^{n} x_i^2 - 2\theta\sum_{i=1}^{n} x_i + n\theta^2\right) - \frac{1}{2\lambda}\left(\theta^2 - 2\mu\theta + \mu^2\right)\right] \\ &\propto \exp\left[-\frac{1}{2\sigma^2}\left(-2n\bar{x}\theta + n\theta^2\right) - \frac{1}{2\lambda}\left(\theta^2 - 2\mu\theta\right)\right] \end{aligned}$$

$$\propto \ \exp\left[-\frac{1}{2}\left(\frac{n}{\sigma^2} + \frac{1}{\lambda}\right)\theta^2 + \left(\frac{n\bar{x}}{\sigma^2} + \frac{\mu}{\lambda}\right)\theta\right]$$

หรือ

$$f_{\bar{X}|\Theta}(\bar{x}|\theta)\, f_{\Theta}(\theta) \ \propto \ \exp\left[-\frac{1}{2}\left(\frac{\sigma^2 + n\lambda}{\lambda\sigma^2}\right)\left\{\theta^2 - 2\left(\frac{\mu\sigma^2 + n\lambda\,\bar{x}}{\sigma^2 + n\lambda}\right)\theta\right\}\right]$$

$$\propto \ \exp\left[-\frac{1}{2}\left(\frac{\sigma^2 + n\lambda}{\lambda\sigma^2}\right)\left(\theta - \frac{\mu\sigma^2 + n\lambda\bar{x}}{\sigma^2 + n\lambda}\right)^2\right]$$

เพราะฉะนั้นได้

$$f_{\Theta|\bar{X}}(\theta|\bar{x}) \ \propto \ \exp\left[-\frac{1}{2}\left(\frac{\sigma^2 + n\lambda}{\lambda\sigma^2}\right)\left(\theta - \frac{\mu\sigma^2 + n\lambda\bar{x}}{\sigma^2 + n\lambda}\right)^2\right]$$

ซึ่งเข้าลักษณะฟังก์ชันความหนาแน่นของการแจกแจงปกติ ดังนั้น สรุปได้ว่าการแจกแจงภายหลังของ Θ เมื่อกำหนด $\bar{X} = \bar{x}$ เป็นการแจกแจงแบบปกติโดยมีค่าเฉลี่ย

$$E(\Theta|\bar{X} = \bar{x}) = \frac{\mu\sigma^2 + n\lambda\bar{x}}{\sigma^2 + n\lambda} \quad \text{และความแปรปรวน} \quad Var(\Theta|\bar{X} = \bar{x}) = \frac{\lambda\sigma^2}{\sigma^2 + n\lambda}$$

เมื่อกำหนดฟังก์ชันความเสียหายเป็นฟังก์ชันของค่าเคลื่อนสัมบูรณ์ได้ตัวประมาณเบส์ของ θ เป็นมัธยฐานของ $\Theta|\bar{X}$ ซึ่งเท่ากับค่าเฉลี่ยของ $\Theta|\bar{X}$ (กรณีการแจกแจงปกติ) คือ

$$\delta(\underline{X}) = \hat{\theta} = E(\Theta|\bar{X}) = \frac{\mu\sigma^2 + n\lambda\bar{X}}{\sigma^2 + n\lambda}$$

◀

การหาค่าประมาณเบส์ในปัญหาตัวอย่างที่ 1 และ 2 ใช้วิธีการเชิงวิเคราะห์ (ใช้แคลคูลัส) แต่บางปัญหาจะยากมากที่จะใช้วิธีการเชิงวิเคราะห์ ในกรณีนี้อาจแก้ปัญหาโดยใช้การจำลองมอนติคาร์โลในการหาค่าประมาณเบส์ ซึ่งในบทความนี้จะกล่าวถึงการใช้การจำลองมอนติคาร์โลที่ใช้วิธีการสุ่มหรือการจำลองข้อมูลที่มีชื่อเรียกว่า **การสุ่มตัวอย่างแบบกิบส์** (Gibbs sampling) หรือ **กิบส์แซมเพลอร์** (Gibbs sampler) ซึ่งเป็นวิธีการหนึ่งใน **วิธีมอนติคาร์โลลูกโซ่มาร์คอฟ** (Markov chain Monte Carlo (MCMC) methods)

3. การสุ่มตัวอย่างแบบกิบส์

หลักการของวิธีการสุ่มตัวอย่างแบบกิบส์ คือ การจำลองตัวแปรสุ่มจากการแจกแจงแบบมีเงื่อนไข เริ่มจากแนวคิดการจำลองตัวแปรสุ่มตามขั้นตอนวิธีที่กล่าวเป็นทฤษฎีบทต่อไปนี้

ทฤษฎีบท สมมติจำลองตัวแปรสุ่มตามขั้นตอนวิธีดังนี้

(1) จำลอง $Y \sim f_Y(y)$

(2) จำลอง $X \sim f_{X|Y}(x|Y)$

ดังนั้น X มีฟังก์ชันความหนาแน่นหรือฟังก์ชันความน่าจะเป็น $f_X(x)$

◀

การสุ่มตัวอย่างแบบกิบส์มีแนวทางทำนองเดียวกันกับทฤษฎีบทข้างต้น

สำหรับการจำลองตัวแปรสุ่มสองตัว (X, Y) และ $(X, Y) \sim f(x, y)$ มีขั้นตอนวิธีดังต่อไปนี้ โดยสมมติว่า มีตัวแบบจำลองหรือขั้นตอนวิธีจำลองตัวแปรสุ่ม X จาก $f_{X|Y}$ และจำลองตัวแปรสุ่ม Y จาก $f_{Y|X}$

(1) ให้ n เป็นจำนวนเต็มบวก และ Y_0 เป็นค่าเริ่มต้นที่เป็นไปได้ของตัวแปรสุ่ม Y

(2) สำหรับ $i = 1, 2, \ldots, n$ ทำขั้นตอน (3) และ (4)

(3) จำลอง $X_i \sim f_{X|Y}(x|Y_{i-1})$

(4) จำลอง $Y_i \sim f_{Y|X}(y|X_i)$

นั่นคือ จากค่า $Y_0 = y_0$ จำลอง X_1 จาก $f_{X|Y}(x|y_0)$ สมมติได้ $X_1 = x_1$ นำไปจำลอง Y_1 จาก $f_{Y|X}(y|x_1)$ สมมติได้ $Y_1 = y_1$ นำไปจำลอง X_2 จาก $f_{X|Y}(x|y_1)$ สมมติได้ $X_2 = x_2$ นำไปจำลอง Y_2 จาก $f_{Y|X}(y|x_2)$ เวียนซ้ำเช่นนี้ไปเรื่อย ๆ จะได้ลำดับ $(X_1, Y_1), (X_2, Y_2), \ldots$ ซึ่งได้ว่า $X_1, X_2, \ldots$ เป็นอิสระกัน และ $Y_1, Y_2, \ldots$ เป็นอิสระกัน และสำหรับ i ค่าใหญ่ X_i จะมีการแจกแจงลู่เข้าสู่การแจกแจงของ $X \sim f_X(x)$ และ Y_i จะมีการแจกแจงลู่เข้าสู่การแจกแจงของ $Y \sim f_Y(y)$ และด้วยกฎอ่อนของจำนวนใหญ่ (weak law of large numbers) จะได้ ค่าเฉลี่ยตัวอย่าง $\frac{1}{n}\sum_{i=1}^{n} h(X_i) \xrightarrow{p} E[h(X)]$ (ลู่เข้าในความน่าจะเป็น) ขณะที่ $n \to \infty$

เนื่องจากการจำลอง (X_{k+1}, Y_{k+1}) จะใช้ (X_k, Y_k) เท่านั้น ไม่ขึ้นอยู่กับ (X_{k-1}, Y_{k-1}), $(X_{k-2}, Y_{k-2}), \ldots$ แสดงว่า สถานะอนาคตของลำดับขึ้นอยู่กับสถานะปัจจุบันของลำดับเท่านั้น ลักษณะเช่นนี้สอดคล้องกระบวนการเฟ้นสุ่ม (stochastic process) ที่เรียกว่า "ลูกโซ่มาร์คอฟ" (Markov chain) และด้วยคุณสมบัติการลู่เข้าของความน่าจะเป็นของการเปลี่ยนสถานะ (convergence of transition probabilities) ในลูกโซ่มาร์คอฟ จะได้ว่าลำดับ (X_i, Y_i) มีการแจกแจงลู่เข้าสู่ $f(x, y)$ สำหรับ i ค่าใหญ่ หรือ (X_i, Y_i) มีการแจกแจง $f(x, y)$ โดยประมาณ ด้วยความคลาดเคลื่อนน้อยเมื่อ i ค่าใหญ่

ในทางปฏิบัติ i ค่าใหญ่นั้น จะเริ่มด้วยค่าใหญ่ให้เป็น $m+1$ เช่น เริ่มที่ 1001 (กำหนด $m = 1000$) เป็นต้น โดยไม่ใช้ค่าของ $X_1, X_2, \ldots, X_m$ และไม่ใช้ค่าของ $Y_1, Y_2, \ldots, Y_m$ (จำลองทิ้ง m ค่าแรก) จะเริ่มใช้ค่าของ X_i และ Y_i ตั้งแต่ X_{m+1} และ Y_{m+1} เป็นต้นไป

ในกรณีเวกเตอร์สุ่มของตัวแปรสุ่มมากกว่าสองตัว $(p > 2)$ การจำลอง $\underline{X} = (X_1, X_2, X_3, \ldots, X_p)$ ด้วยวิธีการสุ่มตัวอย่างแบบกิบส์ จะมีกรรมวิธีทำนองเดียวกันกับกรณีตัวแปรสุ่มสองตัว โดยสมมติว่า มีตัวแบบจำลองหรือขั้นตอนวิธีจำลองตัวแปรสุ่ม X_j จากการแจกแจงแบบมีเงื่อนไข และสมมติว่า มีฟังก์ชันความหนาแน่นแบบมีเงื่อนไขหรือฟังก์ชันความน่าจะเป็นแบบมีเงื่อนไข

$$f_{j|\cdot}(x_j \,|\, x_1, x_2, \ldots, x_{j-1}, x_{j+1}, \ldots, x_p)$$

เมื่อกำหนด $X_1 = x_1, X_2 = x_2, \ldots, X_{j-1} = x_{j-1}, X_{j+1} = x_{j+1}, \ldots, X_p = x_p$ สำหรับ $j = 1, 2, 3, \ldots, p$ เพราะฉะนั้น ขั้นตอนวิธีของการสุ่มตัวอย่างแบบกิบส์ กรณีตัวแปรสุ่ม p ตัว $(X_1, X_2, \ldots, X_p)$ มีดังนี้ โดยให้ $X_{j,i}$ แทนค่าของ X_j ในการจำลองครั้งที่ i และในทำนองเดียวกันกับกรณีสองตัวแปรสุ่ม ควรใช้ค่า $X_{j,i}$ เมื่อ i มีค่าใหญ่

(1) ให้ n เป็นจำนวนเต็มบวก และ $(X_{2,0}, X_{3,0}, \ldots, X_{p,0})$ เป็นค่าเริ่มต้นของ $(X_2, X_3, \ldots, X_p)$

(2) สำหรับ $i = 1, 2, \ldots, n$ ทำขั้นตอน (3) ถึง $(2+p)$

(3) จำลอง $X_{1,i} \sim f_{1|\cdot}(x_1 \,|\, X_{2,i-1}, X_{3,i-1}, \ldots, X_{p,i-1})$

(4) จำลอง $X_{2,i} \sim f_{2|\cdot}(x_2 \,|\, X_{1,i}, X_{3,i-1}, \ldots, X_{p,i-1})$

(5) จำลอง $X_{3,i} \sim f_{3|\cdot}(x_3 \,|\, X_{1,i}, X_{2,i}, X_{4,i-1}, \ldots, X_{p,i-1})$

$\vdots$

$(2+p)$ จำลอง $X_{p,i} \sim f_{p|\cdot}(x_p \,|\, X_{1,i}, X_{2,i}, \ldots, X_{p-1,i})$

4. กรณีศึกษา: การหาค่าประมาณเบส์โดยใช้การสุ่มตัวอย่างแบบกิบส์

ให้ $X_1, X_2, \ldots, X_n$ เป็นตัวอย่างสุ่มจาก $N(\theta, \sigma^2)$ ให้ $Y = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ กำหนด

$$Y \,|\, \Theta = \theta \sim N\left(\theta, \frac{\sigma^2}{n}\right), \quad \Theta \,|\, V^2 = v^2 \sim N(0, v^2) \quad \text{และ} \quad V^2 \sim g(v^2) = \frac{\lambda^\alpha}{\Gamma(\alpha)} \frac{1}{(v^2)^{\alpha+1}} e^{-\lambda/v^2},$$

$v^2 > 0$ จะหาค่าประมาณเบส์ของ θ ด้วยวิธีการจำลอง กำหนดฟังก์ชันความเสียหายในรูปแบบกำลังสอง และใช้การสุ่มตัวอย่างหรือการจำลองข้อมูลแบบกิบส์ โดยที่ทราบค่าของ $n,\ y,\ \sigma^2, \alpha$ และ λ

ด้วยวิธีการจำลองแบบกิบส์เพื่อหาค่าประมาณเบส์ของ θ จะจำลอง (Θ_i, V_i^2) ดังนั้น ต้องการฟังก์ชันความหนาแน่นแบบมีเงื่อนไข $f_{\Theta|Y,V^2}(\theta|y,v^2)$ และ $f_{V^2|Y,\Theta}(v^2|y,\theta)$

$$
\begin{aligned}
f_{\Theta|Y,V^2}(\theta|y,v^2) &= \frac{f_1(\theta,y,v^2)}{f_2(y,v^2)} \\
&\propto f_1(\theta,y,v^2) = g(v^2)h(\theta|v^2)k(y|\theta,v^2) \\
&\propto f(y|\theta)h(\theta|v^2) \quad [\text{การแจกแจงแบบมีเงื่อนไขของ } Y|\theta \text{ ไม่ขึ้นอยู่กับ } v^2]
\end{aligned}
$$

ด้วยวิธีการทำนองเดียวกันกับตัวอย่างที่ 2 ได้ว่า

$$
\begin{aligned}
f(y|\theta)h(\theta|v^2) &= \frac{\sqrt{n}}{\sigma\sqrt{2\pi}} e^{-\frac{n}{2\sigma^2}(y-\theta)^2} \times \frac{1}{v\sqrt{2\pi}} e^{-\frac{1}{2v^2}\theta^2} \\
&\propto \exp\left[-\frac{1}{2}\left(\frac{nv^2+\sigma^2}{v^2\sigma^2} \right)\left(\theta - \frac{nv^2 y}{nv^2+\sigma^2} \right)^2 \right]
\end{aligned}
$$

เพราะฉะนั้น

$$
\Theta|Y=y, V^2=v^2 \sim N\left(\frac{nv^2 y}{nv^2+\sigma^2}, \frac{v^2\sigma^2}{nv^2+\sigma^2} \right)
$$

ในทำนองเดียวกัน สำหรับการหา $f_{V^2|Y,\Theta}(v^2|y,\theta)$ ได้ว่า

$$
\begin{aligned}
f_{V^2|Y,\Theta}(v^2|y,\theta) &\propto f(y|\theta)h(\theta|v^2)g(v^2) \\
&\propto h(\theta|v^2)g(v^2) \\
&\propto \frac{1}{\sqrt{u}} e^{-\frac{1}{2u}\theta^2} \frac{1}{u^{\alpha+1}} e^{-\frac{\lambda}{u}} \quad (\text{ให้ } u \text{ แทน } v^2) \\
&\propto \frac{1}{u^{\alpha+3/2}} e^{-\left(\frac{\theta^2}{2}+\lambda\right)\frac{1}{u}}
\end{aligned}
$$

จากนี้จะหารูปแบบการแจกแจงที่ง่ายขึ้นโดยทำการแปลงตัวแปรดังนี้ ให้ $w = \frac{1}{u}$ ฉะนั้น ได้จาโคเบียน (Jacobian) ของการแปลงเป็น $J = \frac{du}{dw} = -\frac{1}{w^2}$ และได้ฟังก์ชันความหนาแน่นของ W เมื่อกำหนด $Y = y, \Theta = \theta$ ดังนี้

$$f_{W|Y,\Theta}(w|y,\theta) \;\propto\; w^{(\alpha+1/2)-1}e^{-(\theta^2/2+\lambda)w}$$

ซึ่งเข้าลักษณะฟังก์ชันความหนาแน่นของการแจกแจงแกมมา $G\left(\alpha+\frac{1}{2}, \frac{\theta^2}{2}+\lambda\right)$ ดังนั้น ได้ว่า

$$W|Y=y,\Theta=\theta \;\sim\; G\left(\alpha+\frac{1}{2}, \frac{\theta^2}{2}+\lambda\right)$$

เพราะฉะนั้น จำลอง (Θ_i, V_i^2) ด้วยขั้นตอนวิธีดังนี้

(1) กำหนดค่าเริ่มต้น V_0^2 และค่าของ n, y, σ^2, α และ λ

(2) สำหรับ $i = 1,2,\ldots,m$ ทำขั้นตอน (3) ถึง (5)

(3) จำลอง Θ_i จาก $f_{\Theta|Y,V^2}(\theta|y,V_{i-1}^2)$ [จาก $N\left(\frac{nV_{i-1}^2 y}{nV_{i-1}^2+\sigma^2}, \frac{V_{i-1}^2\sigma^2}{nV_{i-1}^2+\sigma^2}\right)$]

(4) จำลอง W_i จาก $f_{W|Y,\Theta}(w|y,\Theta_i)$ [จาก $G\left(\alpha+\frac{1}{2}, \frac{\Theta_i^2}{2}+\lambda\right)$]

(5) ให้ $V_i^2 = \frac{1}{W_i}$

ผลการจำลอง เมื่อ $i \to \infty$ จะได้ว่า

$$\Theta_i \xrightarrow{d} f_{\Theta|Y}(\theta|y) \text{ (ลู่เข้าสู่ในการแจกแจง)}$$

$$V_i^2 \xrightarrow{d} f_{V^2|Y}(v^2|y)$$

เพราะฉะนั้น ค่าประมาณเบส์เมื่อกำหนดฟังก์ชันความเสียหายในรูปแบบกำลังสอง คือค่าเฉลี่ยตัวอย่างของ Θ_i เท่ากับ $\frac{1}{m-k}\sum_{i=k+1}^{m}\Theta_i$ เมื่อตัดค่า Θ_i จำนวน k ค่าแรก

ตัวอย่างผลการจำลองแสดงในตารางที่ 1 และ 2 กำหนด $n = 15$, $y = 20$, $\lambda = 2$ และค่าเริ่มต้น $V_0^2 = 1$ และแปรเปลี่ยนค่า (เพิ่มขึ้น 25 %) ของพารามิเตอร์ σ^2 (พารามิเตอร์ในระดับชั้นต้น) และ α (พารามิเตอร์ในระดับชั้นปลาย) เพื่อพิจารณาผลกระทบต่อค่าประมาณเบส์ $\hat{\theta}$ ของ θ ในแต่ละกรณีทำการจำลองโดยใช้โปรแกรม Fortran รูปที่ 1 กำหนดจำนวนรอบจำลอง $m = 6000$ รอบ เริ่มต้นด้วยเลขสุ่ม (seed) 4567 และตัดค่า Θ_i จำนวน $k = 1000$ ค่าแรก ผลการจำลองแสดงค่าประมาณเบส์ $\hat{\theta}$ ของ θ ในแต่ละกรณีในตารางที่ 1 และ 2

[หมายเหตุ : โปรแกรมในรูปที่ 1 ใช้สำหรับ $\alpha \geq 0.5$]

```
C      *** THIS PROGRAM USES GIBBS SAMPLING TO COMPUTE THE BAYES
C          ESTIMATE.

      REAL LAMBDA
    5 PRINT*, '  '
      PRINT*, '  ENTER THE NUMBER OF SIMULATION RUNS M:'
      READ*, M
      IF(M .LE. 0) STOP
      PRINT*, '  ENTER THE NUMBER OF OMITTED VALUES K:'
      READ*, K
      PRINT*, '  ENTER THE STARTING VALUE OF V SQUARED:'
      READ*, VI0
      PRINT*, '  ENTER THE SAMPLE SIZE N:'
      READ*, N
      PRINT*, '  ENTER THE SAMPLE MEAN Y:'
      READ*, Y
      PRINT*, '  ENTER THE SIGMA SQUARED:'
      READ*, SIGMA2
      PRINT*, '  ENTER THE ALPHA VALUE:'
      READ*, ALPHA
      PRINT*, '  ENTER THE LAMBDA VALUE:'
      READ*, LAMBDA
      PRINT*, '  ENTER THE DESIRED SEED:'
      READ*, IX
      VI = VI0
      ALPHAP = ALPHA + 0.5
      NUMRUNS = M - K
      A = SQRT(2*ALPHAP - 1.0)
      B = 2.0*ALPHAP - 1.386294 + (1.0/A)
      DO 10 J = 1,K
            CALL
GIBBS(IX,VI,N,Y,SIGMA2,ALPHAP,LAMBDA,THETA,A,B)
   10 CONTINUE
      SUM = 0.0
      DO 20 J = 1,NUMRUNS
            CALL
GIBBS(IX,VI,N,Y,SIGMA2,ALPHAP,LAMBDA,THETA,A,B)
            SUM = SUM + THETA
   20 CONTINUE
      AVERAGE = SUM/FLOAT(NUMRUNS)
      PRINT*, '  '
      PRINT*, '  THE BAYES ESTIMATE IS ',AVERAGE
      GO TO 5
      END
```

รูปที่ 1

```
C     *** GIBBS ALGORITHM
      SUBROUTINE GIBBS(IX,VI,N,Y,SIGMA2,ALPHAP,LAMBDA,THETA,A,B)
      REAL   LAMBDA,LAMBDAP,MEAN
C     *** THIS SECTION SIMULATES Θi .
      DEN = N*VI + SIGMA2
      MEAN = N*VI*Y/DEN
      VAR = VI*SIGMA2/DEN
      SD = SQRT(VAR)
      R1 = URAND(IX)
      R2 = URAND(IX)
      Z = SQRT(-2.0*ALOG(R1))*COS(6.283185*R2)
      THETA = MEAN + SD*Z
C     *** THIS SECTION SIMULATES Vi^2 .
      LAMBDAP     = 0.5*THETA*THETA + LAMBDA
    5 R1 = URAND(IX)
      R2 = URAND(IX)
      U = ALPHAP*((R1/(1.0-R1))**A)
      W = B - ALOG(R1*R1*R2)
      IF(U .GT. W) GO TO 5
      VI = LAMBDAP/U
      RETURN
      END
C     *** THIS FUNCTION SIMULATES A RANDOM NUMBER U(0,1).
      FUNCTION URAND(IX)
      IX = DMOD(16807.0D0*IX, 2147483647.0D0)
      URAND = IX/2147483647.0
      RETURN
      END
```

รูปที่ 1 (ต่อ)

ตารางที่ 1 $\alpha = 2, \lambda = 2$

σ^2	49.00	61.25	76.56	95.70	119.63	149.54
$\hat{\theta}$	19.61	19.52	19.39	19.24	19.04	18.79

ตารางที่ 2 $\sigma^2 = 49, \lambda = 2$

α	2.00	2.50	3.13	3.91	4.89	6.11
$\hat{\theta}$	19.61	19.59	19.56	19.51	19.48	19.44

ผลในตารางที่ 1 และ 2 พบว่าค่าพารามิเตอร์ความแปรปรวน σ^2 จะมีผลกระทบต่อค่าประมาณเบส์มากกว่าผลกระทบจากค่าของพารามิเตอร์ α (และ λ เมื่อทดลองกับ λ)

◀

5. บทสรุป

การจำลองตัวแปรสุ่มหลายตัวที่ไม่เป็นอิสระกันจากการแจกแจงความน่าจะเป็นร่วมโดย ตรงอาจจะยากเกินไปที่จะทำได้ หรืออาจไม่มีประสิทธิภาพ หนทางหนึ่งในการแก้ปัญหาดังกล่าว คือ การใช้วิธีการประมาณ เช่น การจำลองตัวแปรสุ่มด้วยวิธีการที่เรียกว่าการสุ่มตัวอย่างแบบกิบส์ จะเป็นวิธีหนึ่งที่อาจช่วยแก้ปัญหาได้ดังเช่นกรณีศึกษาข้างต้น

บทความนี้คงจะเป็นประโยชน์บ้างแก่นักจำลอง นักสถิติ นักคณิตศาสตร์ นักวิจัยการดำเนินการ และผู้สนใจทั่วไปที่จะนำไปใช้ประโยชน์ในงานของท่านหรือเพื่อศึกษาค้นคว้าต่อไป

บรรณานุกรม

Casella, G. and George, E. I. (1992). Explaining the Gibbs sampler, **The American Statistician,** 46, 167-174.

Cheng, R. C. H. (1977). The generation of gamma variables. Appl. Stat., 26, 71 - 75.

Fishman, G. S. (2006). **A First Course in Monte Carlo.** Belmont, CA: Thomson Brooks/Cole.

Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications, **Biometrika,** 57, 92-109.

Robert, C. P. and Casella, G. (2004). **Monte Carlo Statistical Methods.** 2nd ed. New York: Springer.

Ross, S. M. (1997). **Simulation.** 2nd ed. New York: Academic Press.

มานพ วราภักดิ์. (2550). **การจำลอง.** กรุงเทพมหานคร : สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

_________ . (2550). การสุ่มตัวอย่างแบบกิบส์ใน MCMC. **จุฬาลงกรณ์ธุรกิจปริทัศน์** 114, 103 - 115.