วิธีประมาณความแปรปรวนของตัวประมาณด้วยความแปรปรวนแบบมีเงื่อนไข :
กรณีศึกษาความแปรปรวนของมัธยฐาน

โดย
นางสาวกัญญารัตน์ หัสมา

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาสถิติประยุกต์
ภาควิชาสถิติ
บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร
ปีการศึกษา 2552

วิธีประมาณความแปรปรวนของตัวประมาณด้วยความแปรปรวนแบบมีเงื่อนไข :
กรณีศึกษาความแปรปรวนของมัธยฐาน

โดย
นางสาวกัญญารัตน์ หัสมา

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาสถิติประยุกต์
ภาควิชาสถิติ
บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร
ปีการศึกษา 2552

ESTIMATION OF VARIANCE OF AN ESTIMATOR BY CONDITIONAL VARIANCE :

A CASE STUDY THE VARIANCE OF MEDIAN

By

Kanyarat Husma

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

Department of Statistics

Graduate School

SILPAKORN UNIVERSITY

2009

บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร อนุมัติให้วิทยานิพนธ์เรื่อง "วิธีประมาณความแปรปรวนของตัวประมาณด้วยความแปรปรวนแบบมีเงื่อนไข : กรณีศึกษาความแปรปรวนของมัธยฐาน" เสนอโดย นางสาวกัญญารัตน์ หัสมา เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตร์มหาบัณฑิต สาขาวิชาสถิติประยุกต์

....

(รองศาสตราจารย์ ดร. ศิริชัย ชินะตังกูร)

คณบดีบัณฑิตวิทยาลัย

วันที่........เดือน.......................พ.ศ. ....

อาจารย์ที่ปรึกษาวิทยานิพนธ์

รองศาสตราจารย์ไพบูลย์ รัตนประเสริฐ

คณะกรรมการตรวจสอบวิทยานิพนธ์

....ประธานกรรมการ

(ผู้ช่วยศาสตราจารย์ ดร.กมลชนก พานิชการ)

............../ ..../...................

....กรรมการ

(ผู้ช่วยศาสตราจารย์ ดร.บุญอ้อม โฉมที)

............../........................../...................

....กรรมการ

(รองศาสตราจารย์วีรานันท์ พงศาภักดี)

............./......................./.............

....กรรมการ

(ผู้ช่วยศาสตราจารย์ ดร.กุศยา ปลั่งพงษ์พันธ์)

............../........................../...................

....กรรมการ

(รองศาสตราจารย์ไพบูลย์ รัตนประเสริฐ)

............./......................./.............

47304201 : สาขาวิชาสถิติประยุกต์
คำสำคัญ : การประมาณความแปรปรวน / ความแปรปรวนแบบมีเงื่อนไข / แจ็คไนฟ์ / บูตสแตรพ

กัญญารัตน์ หัสมา : วิธีประมาณความแปรปรวนของตัวประมาณด้วยความแปรปรวนแบบมีเงื่อนไข : กรณีศึกษาความแปรปรวนของมัธยฐาน. อาจารย์ที่ปรึกษาวิทยานิพนธ์ : รศ. ไพบูลย์ รัตนประเสริฐ. 70 หน้า.

ในการวิจัยครั้งนี้มีวัตถุประสงค์ เสนอหลักและวิธีความแปรปรวนแบบมีเงื่อนไข เพื่อใช้ในการประมาณความแปรปรวนของมัธยฐานตัวอย่าง วิธีการประมาณความแปรปรวนมีหลายวิธีด้วยกัน วิธีที่ใช้กันแพร่หลายได้แก่ วิธีแจ็คไนฟ์ และวิธีบูตสแตรพ วิธีการ 2 วิธีนี้มีข้อจำกัดคือ อาศัยการสุ่มตัวอย่างซ้ำจากตัวอย่างเพียงชุดเดียว กรณีตัวอย่างที่ได้ไม่ใช่ตัวแทนที่ดีของประชากรอาจทำให้ค่าบางค่าไม่มีโอกาสที่จะเกิดขึ้นในการสุ่มตัวอย่างซ้ำ ดังนั้น การประมาณความแปรปรวนของมัธยฐานตัวอย่างด้วยวิธีดังกล่าว อาจไม่ถูกต้องเพียงพอ ในการวิจัยได้ทำการจำลองแบบ เพื่อศึกษาเปรียบเทียบประสิทธิภาพของวิธีความแปรปรวนแบบมีเงื่อนไข วิธีแจ็คไนฟ์ และวิธีบูตสแตรพ ในการประมาณความแปรปรวนของค่ามัธยฐานตัวอย่าง โดยมีเกณฑ์ที่ใช้ในการเปรียบเทียบคือ พิจารณาว่าวิธีการประมาณวิธีใดให้ค่าประมาณความแปรปรวนของค่ามัธยฐานตัวอย่างใกล้เคียงกับความแปรปรวนของค่ามัธยฐานตัวอย่างมากกว่าย่อมจะมีประสิทธิภาพมากกว่า ซึ่งวัดจากความแปรปรวนของตัวประมาณความแปรปรวนของค่ามัธยฐานตัวอย่าง สำหรับประชากรที่ใช้ในการศึกษาผู้วิจัยได้ศึกษาประชากรที่มีการแจกแจงแบบปกติมาตรฐาน, การแจกแจงแบบเอกรูป และการแจกแจงแบบปลอมปน

ผลการวิจัยพบว่า

การประมาณความแปรปรวนของค่ามัธยฐานตัวอย่างด้วยวิธีความแปรปรวนแบบมีเงื่อนไข มีประสิทธิภาพมากกว่าการประมาณความแปรปรวนของค่ามัธยฐานตัวอย่างด้วยวิธีบูตสแตรพ และวิธีแจ็คไนฟ์ในกรณีที่ประชากรมีการแจกแจงแบบปกติปลอมปนทั้ง $\mathrm{CN}(0 , 1 , p , 10 , 4)$ เมื่อ $p = 0.05$ และ $p = 0.10$ สำหรับในการศึกษาการแจกแจงแบบปกติมาตรฐานและการแจกแจงแบบเอกรูป การประมาณความแปรปรวนของค่ามัธยฐานตัวอย่างด้วยวิธีความแปรปรวนแบบมีเงื่อนไข มีประสิทธิภาพไม่แตกต่างจากการประมาณความแปรปรวนของค่ามัธยฐานตัวอย่างด้วยวิธีบูตสแตรพ และดีกว่าการประมาณความแปรปรวนของค่ามัธยฐานตัวอย่างด้วยวิธีแจ็คไนฟ์

ภาควิชาสถิติ บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร ปีการศึกษา 2552
ลายมือชื่อนักศึกษา..
ลายมือชื่ออาจารย์ที่ปรึกษาวิทยานิพนธ์..

47304201 : MAJOR : APPLIED STATISTICS
KEY WORDS : ESTIMATE FOR VARIANCE / CONDITIONAL VARIANCE / JACKKNIFE / BOOTSTRAP

KANYARAT HUSMA : ESTIMATION OF VARIANCE OF AN ESTIMATOR BY CONDITIONAL VARIANCE : A CASE STUDY THE VARIANCE OF MEDIAN. THESIS ADVISOR : ASSOC. PROF. PAIBOOL RATANAPRASERT. 70 pp.

The aim of this research is to present the principle and the method of Conditional Variance for estimating variance of the sample median. There are many well-known methods of estimating variance of an estimator such as Jackknife method and Bootstrap methods. However , these two methods have some limitations that depend on the resampling from only one available sample. If the sample is not a good representative , some characteristics have no chance of occurring in the resampling , and the estimator of the variance may not accurate enough. In this research , simulation had been done to compare the efficiency of Conditional Variance method with Jackknife and Bootstrap methods for estimating the variance of the sample median. The criteria that uses in comparison is : any method which has the estimated variance of the sample median closer to the empirical variance of the sample median is said to be more efficient than the others. The populations that were used in the study are standard normal distribution , uniform distribution and contaminated normal distributions.

The results of the study indicate that :

Conditional Variance method for estimating variance of the sample median is more efficient than the Bootstrap and the Jackknife method in cases of the contaminated normal distributions both $CN(0,1,p,10,4)$ when $p = 0.10$ and $CN(0,1,p,10,4)$ when $p = 0.05$. In other case studies , The efficient of the Conditional Variance method for estimating variance of the sample median is not different from the Bootstrap method but better than the Jackknife method.

Department of Statistics Graduate School , Silpakorn University Academic Year 2009

Student's signature..

Thesis Advisor's signature..

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จสมบูรณ์ได้ เพราะ ได้รับความเมตตาอย่างสูง จาก รองศาสตราจารย์ไพบูลย์ รัตนประเสริฐ อาจารย์ที่ปรึกษาวิทยานิพนธ์ ที่ได้ให้คำแนะนำช่วยเหลือตลอดจนตรวจแก้ไขอย่างละเอียดทุกขั้นตอน ผู้วิจัยขอกราบขอบพระคุณอาจารย์เป็นอย่างสูงไว้ ณ ที่นี้

กราบขอบพระคุณ ผู้ช่วยศาสตราจารย์ ดร.กมลชนก พานิชการ ประธานคณะกรรมการตรวจสอบวิทยานิพนธ์ ผู้ประสิทธิ์ประสาทวิชาความรู้อันมีคุณค่ายิ่งแก่ผู้วิจัย กราบขอบพระคุณ ผู้ช่วยศาสตราจารย์ ดร.บุญอ้อม โฉมที ผู้ช่วยศาสตราจารย์ ดร.กุศยา ปลั่งพงษ์พันธ์ และรองศาสตราจารย์วีรานันท์ พงศาภักดี ที่สละเวลาในการอ่าน ซักถาม และให้คำแนะนำอันมีค่ายิ่ง สำหรับการวิจัยครั้งนี้

ท้ายสุดขอกราบขอบพระคุณ บิดาและมารดา ที่ให้ความกรุณาช่วยเหลือทุนทรัพย์แก่ผู้วิจัย พร้อมทั้งสนับสนุนและให้กำลังใจเป็นอย่างดี

คุณค่าอันพึงมีจากวิทยานิพนธ์ฉบับนี้ ผู้วิจัยขอมอบเป็นเครื่องบูชาพระคุณบิดา มารดา ครู อาจารย์ และผู้มีพระคุณทุกท่านที่ได้เมตตาอบรมสั่งสอน ให้ความอนุเคราะห์แก่ผู้วิจัยโดยเสมอมา

สารบัญ

หน้า

บทที่ 1
คำนำ

ความเป็นมาและความสำคัญของปัญหา

ในทางสถิติเราใช้ข้อมูลหรือสารสนเทศจากตัวอย่าง เพื่อวิเคราะห์และประเมินผลถึงสารสนเทศในประชากร ดังนั้นตัวอย่างที่เลือกมาจึงมีผลต่อข้อสรุปเกี่ยวกับประชากร ซึ่งการเลือกตัวอย่างจะมีหลักการสำคัญ 2 ประการ ประการแรกคือต้องการที่จะหลีกเลี่ยงความเอนเอียง (bias) ที่เกิดจากวิธีการเลือกตัวอย่าง ประการที่สองคือต้องการที่จะได้ผลลัพธ์ที่ถูกต้อง มีความแม่นยำสูง และมีประโยชน์ สามารถทำให้แล้วเสร็จได้ภายในเวลาที่เหมาะสม โดยใช้งบประมาณน้อยที่สุด บางครั้งวิธีการเลือกตัวอย่างที่ใช้เป็นวิธีที่สร้างขึ้นโดยนำวิธีแม่บท เช่น การเลือกตัวอย่างสุ่มแบบง่าย การเลือกตัวอย่างแบบมีระบบ การเลือกตัวอย่างแบบมีชั้นภูมิ และการเลือกตัวอย่างแบบกลุ่มมาผสมผสานกัน เพื่อสร้างเป็นแผนแบบการเลือกตัวอย่างสำหรับแต่ละกรณี อาจทำให้ตัวประมาณค่าจากตัวอย่างเป็นฟังก์ชันที่ซับซ้อนจนไม่อาจหาค่าประมาณความคลาดเคลื่อนมาตรฐานของฟังก์ชันเหล่านั้นได้โดยง่าย และถึงแม้ว่าตัวประมาณค่าพารามิเตอร์อาจจะไม่อยู่ในรูปที่ซับซ้อนมากนัก แต่บางครั้งก็ติดค่าพารามิเตอร์บางตัว ซึ่งแม้จะหาค่าได้แต่ก็ไม่อาจดำเนินการได้อย่างง่ายๆ ตัวอย่างเช่น การประมาณค่ามัธยฐานของประชากรด้วยมัธยฐานจากตัวอย่างที่มีรูปของความแปรปรวนคือ $\frac{1}{4nf^2(0)}$ (Paibool Ratanaprasert 1987 : 26) ซึ่งถึงแม้จะสามารถประมาณฟังก์ชันความน่าจะเป็นได้ แต่ก็ทำได้ยาก ดังนั้นจึงมีความจำเป็นที่จะต้องหาวิธีการที่จะสามารถใช้ประมาณค่าความคลาดเคลื่อนมาตรฐาน หรือฟังก์ชันในลักษณะอื่นจากข้อมูลตัวอย่างได้ง่ายๆ ซึ่งทางเลือกหนึ่งก็คือการใช้วิธีสุ่มตัวอย่างซ้ำ โดยนำข้อมูลตัวอย่างชุดหนึ่งที่สุ่มได้ มาทำการเลือกตัวอย่างย่อยซ้ำกันหลายๆ ครั้ง (Resampling) โดยถือเสมือนว่า ตัวอย่างที่ได้จากการสุ่มซ้ำจากตัวอย่างชุดเดียวนี้เป็นตัวอย่างสุ่มซ้ำๆ จากประชากรจริง แล้วทำการหาค่าสถิติจากตัวอย่างซ้ำแต่ละตัวอย่างที่สร้างขึ้น จากนั้นจึงหาความแปรปรวนของค่าสถิติที่ได้เหล่านี้ และใช้เป็นค่าประมาณความคลาดเคลื่อนมาตรฐานหรือค่าแปรปรวนของตัวประมาณ ซึ่งวิธีการเลือกตัวอย่างซ้ำที่มักใช้กันโดยทั่วไปได้แก่ วิธีแจ็คไนฟ์ (Jackknife) และวิธีบูตสแตรพ (Bootstrap)

อย่างไรก็ตาม วิธีการประมาณค่าความแปรปรวนของตัวสถิติโดยวิธีแจ็คไนฟ์ (Jackknife) และวิธีบูตสแตรพ (Bootstrap) มีข้อจำกัด คือ การสุ่มตัวอย่างซ้ำขึ้นอยู่กับตัวอย่างสุ่มเพียงชุดเดียวที่สุ่มได้จากประชากร ดังนั้นอาจทำให้ค่าบางค่าไม่มีโอกาสที่จะเกิดขึ้นในการสุ่ม

ตัวอย่างซ้ำได้เลย ดังนั้นในกรณีที่ตัวอย่างสุ่มที่ได้ไม่ใช่ตัวแทนที่ดีของประชากร การประมาณค่าความแปรปรวนของตัวสถิติที่สนใจก็อาจมีความถูกต้องไม่เพียงพอ นอกจากนั้นการสุ่มตัวอย่างซ้ำโดยวิธีบูตสแตรพ มีหลักว่าจะต้องสุ่มจำนวนมากๆ แต่ไม่สามารถระบุได้ว่าแค่ไหนถึงจะมาก จึงทำให้ผู้ใช้วิธีการนี้มักไม่ค่อยสบายใจนักเพราะไม่แน่ใจว่าจำนวนที่ใช้นั้นมากเพียงพอหรือยัง

ถ้ากำหนดให้ $\hat{\theta}_\phi$ เป็นตัวประมาณค่าพารามิเตอร์ของประชากรที่ขึ้นอยู่กับฟังก์ชัน ϕ และ $E(\hat{\theta}_\phi \mid \Delta)$ คือ ค่าเฉลี่ยแบบมีเงื่อนไขของ $\hat{\theta}_\phi$ เมื่อกำหนดเงื่อนไข Δ จากทฤษฎีของการหาความแปรปรวนแบบมีเงื่อนไข จะได้ว่า

$$\operatorname{var}\left(\hat{\theta}_\phi\right) = \operatorname{var}\left(E\left(\hat{\theta}_\phi \mid \Delta\right)\right) + E\left(\operatorname{var}\left(\hat{\theta}_\phi \mid \Delta\right)\right)$$

และถ้า $\hat{\theta}_\phi$ มีการแจกแจงแบบสมมาตรรอบ θ ซึ่งเป็นพารามิเตอร์ที่ต้องการประมาณ เราจะได้ว่า $E(\hat{\theta}_\phi \mid \Delta) = \theta$ และ $\operatorname{var}\left(E\left(\hat{\theta}_\phi \mid \Delta\right)\right) = \operatorname{var}(\theta) = 0$ ดังนั้นจึงได้

$$\operatorname{var}\left(\hat{\theta}_\phi\right) = E\left(\operatorname{var}\left(\hat{\theta}_\phi \mid \Delta\right)\right)$$

นั่นคือ ความแปรปรวนแบบไม่มีเงื่อนไขของตัวประมาณค่าเท่ากับค่าคาดหวังของความแปรปรวนแบบมีเงื่อนไขของตัวประมาณค่า ซึ่งจะได้ว่าความแปรปรวนแบบมีเงื่อนไขของตัวประมาณค่าเป็นตัวประมาณค่าที่ไม่เอนเอียงของความแปรปรวนแบบไม่มีเงื่อนไข

ดังนั้น เราจึงใช้ความแปรปรวนแบบมีเงื่อนไขของตัวประมาณค่าเพื่อประมาณความแปรปรวนแบบไม่มีเงื่อนไข โดยเงื่อนไขที่ใช้อาจใช้ขนาดของค่าสังเกตที่ได้จากตัวอย่างที่แตกต่างไปจากค่ากลางของตัวอย่างซึ่งสอดคล้องกับฟังก์ชัน

$$\mathrm{T}_\varphi(\mathbf{X},\hat{\theta}) = \sum_{i=1}^{n} sign\left(x_i - \hat{\theta}\right)\phi\left(\left|x_i - \hat{\theta}\right|\right)$$

เมื่อ $\mathbf{X} = \mathrm{X}_1 , \mathrm{X}_2 , \ldots , \mathrm{X}_\mathrm{n}$ เป็นตัวอย่างขนาดสุ่มขนาด n จากประชากร ซึ่งเมื่อเรากำหนดฟังก์ชัน $\phi\left(\left|x_i - \hat{\theta}\right|\right)$ ในลักษณะต่างๆ ก็จะได้ ค่ากลางแตกต่างกันออกไป ตัวอย่างเช่น $\phi(\mathrm{u}) = \mathrm{u}$ จะได้ว่าค่ากลางที่สอดคล้องกันคือค่าเฉลี่ยของตัวอย่าง และถ้ากำหนด $\phi(\mathrm{u}) = 1$ จะได้ $\mathrm{T}_\varphi(\mathbf{X},\hat{\theta}) = \sum_{i=1}^{n} sign(x_i - \hat{\theta})$ คือตัวสถิติเครื่องหมาย ค่ากลางที่สอดคล้องกับสมการคือค่ามัธยฐานของตัวอย่าง

ซึ่งความแปรปรวนแบบมีเงื่อนไขที่สอดคล้องกับฟังก์ชัน

$$\mathrm{T}_{\varphi}(\mathbf{X},\hat{\theta}) = \sum_{i=1}^{n} sign\left(x_i - \hat{\theta}\right)\phi\left(\left|x_i - \hat{\theta}\right|\right)$$

นิยามได้ดังต่อไปนี้

บทนิยาม ความแปรปรวนแบบมีเงื่อนไข คือความแปรปรวนของเซตของคำตอบที่ได้จากฟังก์ชัน

$$\mathrm{T}_{\varphi}(\mathbf{X},\hat{\theta}) = \sum_{i=1}^{n} sign(x_i - \hat{\theta})\phi(\left|x_i - \hat{\theta}\right|)$$

ที่ถูกสร้างโดยเซตของค่าสังเกตที่เป็นไปได้จำนวน 2^n ชุด ในรูป $\theta \pm \Delta_1, \theta \pm \Delta_2, ..., \theta \pm \Delta_n$ ที่มีเงื่อนไขบน $\Delta_1, \Delta_2 ..., \Delta_n$ เมื่อ $\Delta_i = \left|x_i - \theta\right|$, $i = 1,2,...,n$

ตัวอย่างเช่น n = 3 ค่าสังเกตแบบมีเงื่อนไข จำนวน 2^n ชุด ที่เป็นไปได้บนเงื่อนไข $\Delta_i = \left|x_i - \theta\right|$, $i = 1,2,3$ คือ

ชุดที่ 1 $\theta + \Delta_1$, $\theta + \Delta_2$, $\theta + \Delta_3$

ชุดที่ 2 $\theta - \Delta_1$, $\theta + \Delta_2$, $\theta + \Delta_3$

ชุดที่ 3 $\theta + \Delta_1$, $\theta - \Delta_2$, $\theta + \Delta_3$

ชุดที่ 4 $\theta - \Delta_1$, $\theta - \Delta_2$, $\theta + \Delta_3$

ชุดที่ 5 $\theta + \Delta_1$, $\theta + \Delta_2$, $\theta - \Delta_3$

ชุดที่ 6 $\theta - \Delta_1$, $\theta + \Delta_2$, $\theta - \Delta_3$

ชุดที่ 7 $\theta + \Delta_1$, $\theta - \Delta_2$, $\theta - \Delta_3$

ชุดที่ 8 $\theta - \Delta_1$, $\theta - \Delta_2$, $\theta - \Delta_3$

ผู้วิจัยสนใจที่จะใช้ความแปรปรวนแบบมีเงื่อนไข (Conditional Variance) เพื่อประมาณความแปรปรวนของมัธยฐาน และทำการเปรียบเทียบประสิทธิภาพของการประมาณค่าความแปรปรวนของมัธยฐานที่ได้จากวิธีประมาณที่ได้นำเสนอขึ้นมาใหม่นี้ กับวิธีการประมาณความแปรปรวนที่มีใช้อยู่ในปัจจุบัน 2 วิธี คือวิธีแจ็คไนฟ์ (Jackknife) และวิธีบูตสแตรพ (Bootstrap)

วัตถุประสงค์ของการวิจัย

เพื่อเปรียบเทียบประสิทธิภาพการประมาณค่าความแปรปรวนของมัธยฐาน ด้วยวิธีแจ็คไนฟ์ (Jackknife) วิธีบูตสแตรพ (Bootstrap) กับวิธีประมาณความแปรปรวนแบบมีเงื่อนไข (Conditional Variance) ที่ได้เสนอขึ้นใหม่

ขอบเขตของการวิจัย

การศึกษาเป็นการใช้การจำลองแบบ โดยประชากรที่นำมาศึกษาจะประกอบด้วยประชากรที่มีลักษณะดังนี้

1. การแจกแจงแบบปกติมาตรฐาน
2. การแจกแจงแบบเอกรูป (Uniform distribution) ในช่วง $(0,1)$
3. การแจกแจงแบบปกติปลอมปน (Contaminated Normal) ซึ่งถือว่าเป็นการแจกแจงที่เกิดค่าผิดปกติ คือเป็นการแจกแจงผสมระหว่างการแจกแจงแบบปกติ 2 การแจกแจง ที่มีค่าเฉลี่ยและความแปรปรวนไม่เท่ากันผสมกัน โดยใช้สัญลักษณ์ $CN(\mu_1,\sigma_1^2,p,\mu_2,\sigma_2^2)$ ซึ่งในการศึกษาครั้งนี้จะกำหนดให้ข้อมูลส่วนหนึ่งมี $\mu_1 = 0$ และ $\sigma_1^2 = 1$ ส่วนอีกส่วนหนึ่งมี $\mu_2 = 10$ และ $\sigma_2^2 = 4$ และมีสัดส่วนของการแจกแจงผิดปกติที่มาผสมกัน คือ p = 0.05 และ p = 0.10

   $CN(0,1,p,10,4)$ เมื่อ p = 0.05 และ p = 0.10 หมายความว่า การแจกแจงดังกล่าวเกิดจากข้อมูล 2 ส่วน คือเกิดจากการแจกแจง $N(\mu_1,\sigma_1^2)$ ในสัดส่วน (1-p) และการแจกแจง $N(\mu_2,\sigma_2^2)$ ในสัดส่วน p เราจึงอาจกล่าวว่าการแจกแจง $N(\mu_1,\sigma_1^2)$ ปนกับ p% ของการแจกแจง $N(\mu_2,\sigma_2^2)$ โดยที่ p คือสัดส่วนของตัวอย่างสุ่มจากการแจกแจงแบบ $N(\mu_2,\sigma_2^2)$ ที่นำไปผสมกับตัวอย่างที่สุ่มจากการแจกแจงแบบ $N(\mu_1,\sigma_1^2)$
4. เนื่องจากวิธีบูตสแตรพการประมาณความแปรปรวนของตัวสถิติขึ้นกับจำนวนครั้งของการสุ่มตัวอย่างซ้ำ ในการวิจัยครั้งนี้จะกำหนดจำนวนครั้งของการสุ่มตัวอย่างซ้ำสำหรับวิธีบูตสแตรพให้เท่ากับ 500 ครั้ง
5. สุ่มตัวอย่างขนาด n จากประชากรที่มีการแจกแจงในลักษณะต่างๆ ที่กำหนด โดยกำหนดให้ขนาดตัวอย่าง $n = 6$, 10 และ 15

ประโยชน์ที่คาดว่าจะได้รับ

1. เพื่อทราบว่าวิธีใดให้ประสิทธิภาพการประมาณค่าความแปรปรวนของมัธยฐานได้ดีกว่ากัน ภายใต้เงื่อนไขที่กำหนด

2. เพื่อเป็นแนวทางใหม่ในการประมาณความแปรปรวนของตัวสถิติให้มีประสิทธิภาพมากขึ้น

นิยามศัพท์เฉพาะ

สมมติว่าเราประมาณ θ โดย $\hat{\theta}$ ซึ่งเป็นคำตอบของฟังก์ชัน

$$\mathrm{T}_{\varphi}(\mathbf{X},\hat{\theta}) = \sum_{i=1}^{n} sign(x_i - \hat{\theta})\phi(\left|x_i - \hat{\theta}\right|)$$

แล้วจะได้ว่าความแปรปรวนแบบมีเงื่อนไข นิยามได้ดังต่อไปนี้

บทนิยาม ความแปรปรวนแบบมีเงื่อนไข คือความแปรปรวนของเซตของคำตอบที่ได้จากฟังก์ชัน

$$\mathrm{T}_{\varphi}(\mathbf{X},\hat{\theta}) = \sum_{i=1}^{n} sign(x_i - \hat{\theta})\phi(\left|x_i - \hat{\theta}\right|)$$

ที่ถูกสร้างโดยเซตของค่าสังเกตที่เป็นไปได้จำนวน 2^n ชุด ในรูป $\theta \pm \Delta_1, \theta \pm \Delta_2, ..., \theta \pm \Delta_n$ ที่มีเงื่อนไขบน $\Delta_1, \Delta_2 ..., \Delta_n$ เมื่อ $\Delta_i = \left|x_i - \theta\right|$, $i = 1,2,...,n$

บทที่ 2
ทฤษฎีและวรรณกรรมที่เกี่ยวข้อง

ในบทนี้จะเป็นส่วนของทฤษฎีต่างๆ ที่เกี่ยวข้องกับงานวิจัยตลอดจนนำเสนอผลงานวิจัยที่เกี่ยวข้องที่ได้ศึกษาค้นคว้า การนำเสนอแบ่งออกเป็น 4 หัวข้อ ดังนี้

1. วิธีประมาณความแปรปรวนด้วยวิธีแจ็คไนฟ์
2. วิธีประมาณความแปรปรวนด้วยวิธีบูตสแตรพ
3. วิธีประมาณความแปรปรวนแบบมีเงื่อนไข
4. งานวิจัยที่เกี่ยวข้องกับการประมาณค่าความแปรปรวน

1. วิธีประมาณความแปรปรวนด้วยวิธีแจ็คไนฟ์

วิธีแจ็คไนฟ์ (Jackknife) ได้รับการพัฒนาขึ้นโดย Quenouille ในปี ค.ศ. 1949 ต่อมาในปี ค.ศ. 1958 Tukey เป็นผู้ตั้งชื่อให้แก่วิธีดังกล่าว เนื่องจากวิธีดังกล่าวเปรียบเสมือนกับมีดพับของลูกเสือซึ่งมีความสมบุกสมบันและพร้อมเสมอที่จะถูกใช้ภายใต้สถานการณ์ต่างๆ วิธีแจ็คไนฟ์ (Jackknife) ถูกพัฒนาขึ้นเพื่อใช้เป็นเทคนิคในการลดความเอนเอียง (bias) ในการประมาณค่าพารามิเตอร์ ซึ่งวิธีการดังกล่าวนี้ สามารถหาค่าประมาณความแปรปรวนของตัวประมาณใดๆ ที่มีวิธีการหาค่าประมาณโดยใช้ข้อมูลตัวอย่างในลักษณะเดียวกับการหาค่าพารามิเตอร์โดยใช้ข้อมูลจากประชากรตัวอย่าง เช่น ค่าเฉลี่ยของตัวอย่างมีวิธีการคำนวณเหมือนกับค่าเฉลี่ยของประชากร หรือสัมประสิทธิ์สหสัมพันธ์ของข้อมูลตัวอย่างมีวิธีการคำนวณเช่นเดียวกับสัมประสิทธิ์สหสัมพันธ์ของประชากร เป็นต้น ซึ่งวิธีแจ็คไนฟ์ (Jackknife) จัดเป็นวิธีที่ใช้การเลือกตัวอย่างซ้ำอีกวิธีหนึ่ง ซึ่งเป็นการเลือกตัวอย่างใหม่จากตัวอย่างสุ่มเพียงชุดเดียว โดยมีวิธีการดำเนินการดังนี้

สมมติว่า $X_1 , X_2 , \ldots , X_n$ เป็นตัวอย่างสุ่มขนาด n จากประชากรที่มีฟังก์ชันความหนาแน่น f และฟังก์ชันการแจกแจง F โดย $x_1 , x_2 , \ldots , x_n$ เป็นค่าของ $X_1 , X_2 , \ldots , X_n$ ให้ θ เป็นพารามิเตอร์ที่ต้องการประมาณในประชากรดังกล่าวนี้ และให้ $\hat{\theta}$ เป็นค่าประมาณของ θ ที่คำนวณจากข้อมูลตัวอย่างขนาด n ในการหาค่าประมาณแบบแจ็คไนฟ์จะเริ่มจากตัดค่า x_1 ออกจากตัวอย่าง แล้วคำนวณค่าประมาณของ θ จากตัวอย่างที่ตัด x_1 ออกจากตัวอย่าง สมมติว่า ได้ค่าประมาณ คือ $\hat{\theta}_{(1)}$ ต่อไปนำ x_1 กลับคืนเข้าไปในตัวอย่าง แล้วตัด

x_2 ออกจากตัวอย่าง จากนั้นคำนวณค่าประมาณของ θ จากตัวอย่างที่ตัด x_2 ออก จะได้ค่าประมาณของ θ คือ $\hat{\theta}_{(2)}$ ทำเช่นนี้ซ้ำ n ครั้ง โดยครั้งสุดท้ายตัด x_n ออกจากตัวอย่าง แล้วคำนวณค่าประมาณของ θ จากตัวอย่างที่ตัด x_n ออก จะได้ค่าประมาณของ $\hat{\theta}_{(i)}$ จำนวน n ตัว คือ $\hat{\theta}_{(1)}$, $\hat{\theta}_{(2)}$, … , $\hat{\theta}_{(\mathrm{n})}$ ซึ่งค่า $\hat{\theta}_{(i)}$ นี้ จะเรียกว่าตัวสถิติแจ็คไนฟ์ (Jackknife Statistic) เมื่อได้ค่า $\hat{\theta}_{(i)}$ แล้ว จากนั้นให้คำนวณค่าแฝง (Pseudovalues) ดังสมการ

$$J_i = n\hat{\theta} - (n-1)\hat{\theta}_{(i)} \ ; \ i = 1,2,...,n$$

ซึ่งตัวประมาณแจ็คไนฟ์ของพารามิเตอร์ θ (Jackknife Estimator of θ) คือ

$$J\left(\hat{\theta}\right) = \frac{1}{n}\sum_{i=1}^{n} J_i = \frac{1}{n}\sum_{i=1}^{n}\left(n\hat{\theta} - (n-1)\hat{\theta}_{(i)}\right)$$
$$= n\hat{\theta} - (n-1)\overline{\hat{\theta}}$$

เมื่อ $\overline{\hat{\theta}}$ เป็นค่าเฉลี่ยของ $\hat{\theta}_{(1)}$, $\hat{\theta}_{(2)}$, … , $\hat{\theta}_{(\mathrm{n})}$ นั่นคือ

$$\overline{\hat{\theta}} = \frac{\hat{\theta}_{(1)} + \hat{\theta}_{(2)} + ... + \hat{\theta}_{(n)}}{n} = \frac{\sum_{i=1}^{n}\hat{\theta}_{(i)}}{n}$$

สำหรับการประมาณความแปรปรวนของ $\hat{\theta}$ Tukey (1958) ได้เสนอวิธีการประมาณความแปรปรวนของ θ จากตัวอย่างด้วยวิธีแจ็คไนฟ์ ดังนี้

$$Var(\hat{\theta}) = \frac{(n-1)}{n}\sum_{i=1}^{n}(\hat{\theta}_{(i)} - \hat{\theta}_{(.)})^2$$

ในกรณีที่ $\hat{\theta}$ คือ $\bar{\mathrm{x}}$ จะได้ว่า $\hat{\mathrm{V}}\mathrm{ar}_{\mathrm{J}}(\bar{\mathrm{x}})$ โดยวิธีแจ็คไนฟ์ จะอยู่ในรูปของ

$$\hat{\mathrm{V}}\mathrm{ar}_{\mathrm{J}}(\bar{\mathrm{x}}) = \frac{(\mathrm{n}-1)}{\mathrm{n}}\sum_{\mathrm{i}=1}^{\mathrm{n}}(\bar{\mathrm{x}}_{(\mathrm{i})} - \bar{\mathrm{x}}_{(.)})^2$$

เมื่อ

$$\bar{x}_{(i)} = \frac{\sum_{k=1}^{n} x_k - x_i}{n-1}$$

เป็นค่าเฉลี่ยของตัวอย่างสุ่มชุดที่ i , $i = 1,2,...,n$ ซึ่งตัดค่า x_i ออก และ

$$\bar{x}_{(.)} = \frac{\bar{x}_{(1)} + \bar{x}_{(2)} + ... + \bar{x}_{(n)}}{n} = \frac{\sum_{i=1}^{n} \bar{x}_{(i)}}{n}$$

สำหรับ $\hat{\mathrm{V}}\mathrm{ar}_{\mathrm{J}}(\bar{\mathrm{x}})$ เมื่อ $\hat{\theta}_{\mathrm{J}} = \bar{\mathrm{x}}$ สามารถพิสูจน์ได้ว่า $\hat{\mathrm{V}}\mathrm{ar}_{\mathrm{J}}(\bar{\mathrm{x}})$ ที่ได้จากวิธีแจ็คไนฟ์มีค่าเท่ากับ $\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n(n-1)}$ ซึ่งเป็นรูปของค่าประมาณความแปรปรวนของ $\bar{\mathrm{x}}$ ที่ไม่เอนเอียงตามปกติที่ใช้กัน โดยการพิสูจน์ทำได้ดังนี้

$$\hat{\mathrm{V}}\mathrm{ar}_{\mathrm{J}}(\bar{\mathrm{x}}) = \frac{(\mathrm{n}-1)}{\mathrm{n}} \sum_{\mathrm{i}=1}^{\mathrm{n}} (\bar{\mathrm{x}}_{(\mathrm{i})} - \bar{\mathrm{x}}_{(.)})^2$$

เนื่องจาก

$$\bar{x}_{(i)} = \frac{\sum_{k=1}^{n} x_k - x_i}{n-1}$$

และ

$$\begin{aligned}
\bar{x}(.) &= \frac{\bar{x}_{(1)} + \bar{x}_{(2)} + ... + \bar{x}_{(n)}}{n} \\
&= \frac{1}{n(n-1)}\left[\sum_{i=1}^{n} x_i - x_1 + \sum_{i=1}^{n} x_i - x_2 + ... + \sum_{i=1}^{n} x_i - x_n\right] \\
&= \frac{1}{n(n-1)}\left[n\sum_{i=1}^{n} x_i - \sum_{i=1}^{n} x_i\right] \\
&= \frac{(n-1)}{n(n-1)} \sum_{i=1}^{n} x_i
\end{aligned}$$

$$= \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$= \bar{x}$$

นั่นคือ $\bar{x}_{(.)} = \bar{x}$

เมื่อแทน $\bar{x}_{(i)}$ ด้วย $\dfrac{\sum_{k=1}^{n} x_k - x_i}{n-1}$ และ $\bar{x}_{(.)} = \dfrac{n\sum_{i=1}^{n} x_i - \sum_{i=1}^{n} x_i}{n(n-1)}$ ลงในสูตร $\hat{\mathrm{V}}\mathrm{ar}_{\mathrm{J}}(\bar{\mathrm{x}})$

จะได้

$$\hat{\mathrm{V}}\mathrm{ar}_{\mathrm{J}}(\bar{\mathrm{x}}) = \frac{(n-1)}{n}\sum_{i=1}^{n}\left(\frac{\sum_{k=1}^{n} x_k - x_i}{n-1} - \bar{x}_{(.)}\right)^2$$

$$= \frac{(n-1)}{n}\sum_{i=1}^{n}\left(\frac{\sum_{k=1}^{n} x_k - x_i}{n-1} - \frac{\sum_{i=1}^{n} x_i}{n}\right)^2$$

$$= \frac{(n-1)}{n}\sum_{i=1}^{n}\left(\frac{n\sum_{k=1}^{n} x_k - nx_i}{n(n-1)} - \frac{n\sum_{i=1}^{n} x_i - \sum_{i=1}^{n} x_i}{n(n-1)}\right)^2$$

$$= \frac{(n-1)}{n}\sum_{i=1}^{n}\left(\frac{n\bar{x} - nx_i}{n(n-1)}\right)^2$$

$$= \frac{(n-1)}{n} n \frac{\sum_{i=1}^{n}(\bar{x} - x_i)^2}{n(n-1)}$$

$$= \frac{n-1}{n}\sum_{i=1}^{n}\frac{\left(-x_i + \bar{x}\right)^2}{n-1}$$

$$= \frac{n-1}{n}\sum_{i=1}^{n}\frac{\left(x_i - \bar{x}\right)^2}{n-1}$$

ดังนั้น $$\hat{\mathrm{V}}\mathrm{ar}_{\mathrm{J}}(\bar{\mathrm{x}}) = \frac{n-1}{n}\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}$$

จากที่กล่าวมาแล้วว่า วิธีประมาณความแปรปรวนของตัวประมาณโดยวิธีแจ็คไนฟ์ เป็นการสร้างตัวอย่างใหม่จากตัวอย่างขนาด n เพียงชุดเดียว โดยมีการตัดหน่วยตัวอย่างของตัวอย่างครั้งละ 1 หน่วย จะได้ตัวอย่างใหม่จำนวน n ชุด ซึ่งเรียกวิธีการประมาณความแปรปรวนของตัวประมาณแบบนี้ว่า Delete-1 jackknife variance estimation และเรียกตัวประมาณที่ได้จากวิธีนี้ว่า Delete-1 jackknife variance estimators นอกจากนี้ยังมีวิธีการประมาณความแปรปรวนของตัวประมาณโดยวิธีแจ็คไนฟ์โดยการตัดหน่วยของตัวอย่างครั้งละ d หน่วย เมื่อ d คือ ขนาดของตัวอย่างสุ่มที่ตัดออกจากตัวอย่างสุ่มขนาด n ที่สุ่มจากประชากร ซึ่งจำนวนชุดของตัวอย่างใหม่ที่เป็นไปได้ คือ $\binom{n}{d}$ ชุด เมื่อ $d < n$ สมมติว่ามีตัวอย่างใหม่ที่สร้างได้จำนวน s ชุด ในส่วนของการประมาณความแปรปรวนของตัวประมาณ จะดำเนินการเช่นเดียวกับวิธี Delete-1 jackknife variance estimation แต่เป็นการประมาณความแปรปรวนของตัวประมาณจากตัวอย่างใหม่จำนวน s ชุด โดยเรียกการประมาณความแปรปรวนของตัวประมาณโดยวิธีแจ็คไนฟ์แบบนี้ว่า Delete-d jackknife variance estimation

2. วิธีประมาณความแปรปรวนด้วยวิธีบูตสแตรพ

วิธีบูตสแตรพจัดเป็นวิธีการประมาณความแปรปรวนโดยใช้การสุ่มตัวอย่างซ้ำๆ จากตัวอย่างที่มีอยู่ชุดเดียวเช่นเดียวกับวิธีแจ็คไนฟ์ แต่วิธีการสุ่มตัวอย่างซ้ำโดยวิธีบูตสแตรพจะใช้การสร้างตัวอย่างชุดใหม่จากตัวอย่างสุ่มที่มีเพียงชุดเดียว โดยใช้การสุ่มตัวอย่างแบบแทนที่ วิธีการนี้ถูกเสนอโดย Efron (1979) และได้พัฒนาต่อมาโดย Efron (1982)

เราสามารถประมาณค่าความแปรปรวนของตัวสถิติต่างๆ โดยการสุ่มตัวอย่างขนาด n ที่เป็นไปได้ทั้งหมดจากประชากร แล้วนำค่าของตัวสถิติที่ได้จากตัวอย่างสุ่มแต่ละชุดมาหาความแปรปรวน ก็จะได้ค่าความแปรปรวนของตัวสถิติ แต่เนื่องจากในทางปฏิบัติแล้วสารสนเทศที่มีจากประชากรที่ต้องการศึกษาคือ ตัวอย่างสุ่มเพียง 1 ชุด ดังนั้นจึงไม่อาจหาความแปรปรวนของตัวประมาณด้วยวิธีการดังกล่าว Efron (1979) จึงเสนอให้ใช้วิธีการสุ่มตัวอย่างแบบแทนที่ขนาด n ที่เป็นไปได้ นั่นคือ แทนที่จะสุ่มตัวอย่างซ้ำๆ จากประชากรที่มีฟังก์ชันการแจกแจง F โดยตรง จะใช้การสุ่มตัวอย่างจาก Empirical distribution function (F_n) ของข้อมูลตัวอย่าง $\mathbf{x} = \mathrm{x}_1, \mathrm{x}_2, \ldots, \mathrm{x}_\mathrm{n}$ โดยมีวิธีการดำเนินการดังนี้

ให้ $\mathbf{X} = X_1 , X_2 , \ldots , X_n$ เป็นชุดของตัวแปรสุ่ม และ X_1 เป็นตัวอย่างสุ่มขนาด n ที่เป็นอิสระกัน มาจากประชากรที่มีฟังก์ชันการแจกแจง F มีฟังก์ชันความหนาแน่นเป็น f ให้ $\mathbf{x} = (x_1 , x_2 , x_n)$ เป็นชุดของตัวแปรสุ่ม โดย $\mathbf{x} = x_1 , x_2 , x_n$ เป็นค่าของ $\mathbf{X} = X_1 , X_2 , \ldots , X_n$ และให้ $\hat{\theta}(F)$ เป็นตัวประมาณของพารามิเตอร์ $\theta(F)$ ซึ่งขึ้นอยู่กับฟังก์ชันการแจกแจง F

จากตัวอย่างสุ่มขนาด n $X_1 , X_2 , \ldots , X_n$ จากประชากรที่มีฟังก์ชันการแจกแจง F สร้าง Empirical distribution function โดยให้ความน่าจะเป็นในการเกิดค่าต่างๆ ของ X_i , $i = 1 , 2 , \ldots , n$ คือ $\dfrac{1}{n}$

ทำการสุ่มตัวอย่างขนาด n แบบแทนที่จาก Empirical distribution function ที่ได้ นั่นคือ จะทำการสุ่มตัวอย่างทีละ 1 ค่า จำนวน n ครั้ง จากชุดของตัวอย่าง $\mathbf{x} = (x_1 , x_2 , x_n)$ โดยค่าที่ได้จะคืนกลับไปในชุดตัวอย่างก่อนที่จะมีการสุ่มตัวอย่างครั้งต่อไป ดังนั้น ในตัวอย่างขนาด n ชุดหนึ่ง ค่าของ X_i , $i = 1 , 2 , \ldots , n$ อาจเกิดได้มากกว่า 1 ครั้ง ให้ $\mathbf{X}^* = (X_1^* , X_2^* , \ldots , X_n^*)$ เป็นชุดของตัวอย่างขนาด n ที่สุ่มได้ ซึ่งจะเรียกชุดของตัวอย่างดังกล่าวนี้ว่าตัวอย่างบูตสแตรพ (Bootstrap sample) ดำเนินวิธีการดังกล่าวข้างต้นซ้ำๆ กัน ซึ่งแต่ละครั้งจะได้ตัวอย่างบูตสแตรพ 1 ชุดเสมอ สมมติว่าเรามีตัวอย่างบูตสแตรพขนาด n จำนวน B ชุด โดยที่ B มีขนาดใหญ่ นั่นคือ $X_i^* = 1 , 2 , \ldots , B$ เป็นตัวอย่างบูตสแตรพจำนวน B ชุด โดยที่แต่ละชุดสามารถคำนวณ $\hat{\theta}^*$ ซึ่งเป็นค่าประมาณของ $\theta(F)$ ได้ ดังนั้นจะมีค่าประมาณ $\theta(F)$ จำนวน B ตัว คือ $\hat{\theta}_1^*$, $\hat{\theta}_2^*$, … , $\hat{\theta}_n^*$ สามารถหาค่าประมาณความแปรปรวนของตัวประมาณ $(Var(\hat{\theta}))$ โดยวิธีบูตสแตรพได้ดังนี้

$$Var(\hat{\theta}) = \left\{ \frac{\sum_{i=1}^{B} (\hat{\theta}_i^* - \hat{\theta}_{(.)}^*)^2}{B-1} \right\} \qquad \text{เมื่อ } \hat{\theta}_{(.)}^* = \frac{\sum_{i=1}^{B} \hat{\theta}_i^*}{B}$$

สำหรับ $Var(\hat{\theta})$ เมื่อ $\hat{\theta} = \overline{x}$ จะได้ว่า $\hat{V}ar(\overline{x})$ โดยวิธีบูตสแตรพจะอยู่ในรูป

$$\hat{V}ar_B(\overline{x}^*) = \left\{ \frac{\sum_{i=1}^{B} (\overline{x}_i^* - \overline{x}_{(.)}^*)^2}{B-1} \right\}$$

โดยที่ $\overline{x}_i^* = \dfrac{\sum_{i=1}^{B} X_i^*}{n}$ คือ ค่าเฉลี่ยตัวอย่างบูตสแตรพ ชุดที่ i , i = 1 , 2 , … , B

และ $\overline{x}_{(.)}^* = \dfrac{\sum_{i=1}^{B} \overline{x}_i^*}{B}$ เป็นค่าเฉลี่ยของ $\overline{x}_1^*, \overline{x}_2^*, \ldots, \overline{x}_B^*$

ได้มีผู้วิจัยหลายท่านนำวิธีบูตสแตรพ ไปหาความแปรปรวนของตัวประมาณในลักษณะต่างๆ เช่น

1. Efron (1981) ได้เปรียบเทียบการประมาณความคลาดเคลื่อนของตัวประมาณสัมประสิทธิ์สหสัมพันธ์จากตัวอย่าง โดยใช้วิธีแจ็คไนฟ์ วิธีบูตสแตรพ วิธี Half-sampling วิธี Subsampling และวิธี Balanced repeated replication พบว่า การใช้วิธีแจ็คไนฟ์ วิธีบูตสแตรพ และวิธีการสุ่มซ้ำแบบอื่นๆ สามารถประมาณความคลาดเคลื่อนของตัวประมาณสัมประสิทธิ์สหสัมพันธ์ของตัวอย่างได้ดี ในตัวอย่างที่มีขนาดเล็ก (ขนาดตัวอย่างสุ่มเท่ากับ 14) โดยเฉพาะในกรณีวิธีแจ็คไนฟ์และวิธีบูตสแตรพ พบว่า การประมาณค่าความคลาดเคลื่อนของตัวประมาณสัมประสิทธิ์สหสัมพันธ์จากตัวอย่างโดยวิธีบูตสแตรพ มีความแตกต่างจากความคลาดเคลื่อนมาตรฐานของสัมประสิทธิ์สหสัมพันธ์ของประชากร น้อยกว่าการประมาณค่าความคลาดเคลื่อนของตัวประมาณสัมประสิทธิ์สหสัมพันธ์จากตัวอย่างโดยวิธีแจ็คไนฟ์ และวิธีการสุ่มตัวอย่างแบบอื่นๆ ซึ่งแสดงว่า การประมาณความคลาดเคลื่อนของตัวประมาณสัมประสิทธิ์สหสัมพันธ์โดยวิธีบูตสแตรพมีประสิทธิภาพมากกว่าการประมาณความคลาดเคลื่อนของตัวประมาณสัมประสิทธิ์สหสัมพันธ์โดยวิธีแจ็คไนฟ์ และวิธีการสุ่มซ้ำแบบอื่นๆ

2. Bickel and Freedman (1981) ได้เสนอ การใช้วิธีบูตสแตรพประมาณความแปรปรวนของตัวประมาณค่าเฉลี่ยจากตัวอย่าง โดยทำการศึกษาตัวอย่างที่มาจากประชากรที่มีการแจกแจงแบบปกติมาตรฐาน พบว่า ค่าประมาณความแปรปรวนของค่าเฉลี่ยตัวอย่างเข้าใกล้ความแปรปรวนของค่าเฉลี่ยตัวอย่างที่แท้จริง เมื่อใช้ขนาดตัวอย่าง n เพิ่มขึ้น

3. Efron (1982) ได้ใช้วิธีบูตสแตรพประมาณความแปรปรวนของค่าเฉลี่ย Trimmed mean จากตัวอย่างเปรียบเทียบกับวิธีแจ็คไนฟ์ สำหรับประชากรที่มีการแจกแจงต่อเนื่องบางรูปแบบคือ การแจกแจงแบบปกติ การแจกแจงแบบที และการแจกแจงแบบไคสแควร์ จากผลการศึกษาพบว่า การประมาณค่าความแปรปรวนของ Trimmed mean จากตัวอย่างโดยวิธีบูตสแตรพและวิธีแจ็คไนฟ์ เป็นวิธีการประมาณค่าที่มีประสิทธิภาพไม่แตกต่างกัน ยกเว้นในกรณีการใช้ตัวอย่างขนาดเล็กที่มาจากประชากรที่มีลักษณะเบ้ (การแจกแจงไคสแควร์) พบว่า การประมาณค่าความแปรปรวนของ Trimmed mean จากตัวอย่างทั้ง 2 วิธี ให้ผลของการประมาณความแปรปรวนของ Trimmed mean จากตัวอย่างไม่ค่อยดี

4. Parr (1983) ได้เปรียบเทียบการประมาณความแปรปรวนของตัวประมาณสัมประสิทธิ์สหสัมพันธ์จากตัวอย่างโดยวิธีแจ็คไนฟ์ กับวิธีบูตสแตรพ วิธีแบบเดลต้า (Delta method) โดยใช้วิธีการจำลองแบบด้วยวิธี Monte Carlo โดยใช้ตัวอย่างสุ่มขนาด 20, 100 จากประชากรแบบปกติของ 2 ตัวแปร โดยมีค่าสหสัมพันธ์ของประชากร 0, 0.5, 0.9 ตามลำดับ จากผลการศึกษาพบว่า ในกรณีที่ใช้ตัวอย่างขนาดเล็ก (ตัวอย่างสุ่มขนาด 20) การประมาณความแปรปรวนของตัวประมาณสัมประสิทธิ์สหสัมพันธ์จากตัวอย่าง ใกล้เคียงกับความแปรปรวนของสัมประสิทธิ์สหสัมพันธ์จากประชากรมากกว่าวิธีแจ็คไนฟ์ มีค่าประมาณใกล้เคียงกับความแปรปรวนของสัมประสิทธิ์สหสัมพันธ์จากประชากรไม่แตกต่างกัน จึงสรุปว่า การประมาณความแปรปรวนของตัวประมาณสัมประสิทธิ์สหสัมพันธ์จากตัวอย่างโดยวิธีบูตสแตรพมีประสิทธิภาพมากกว่าการประมาณความแปรปรวนของตัวประมาณสัมประสิทธิ์สหสัมพันธ์จากตัวอย่างโดยวิธีแจ็คไนฟ์ ในกรณีที่ตัวอย่างมีขนาดเล็ก

5. Beran (1984) ได้เปรียบเทียบการประมาณความแปรปรวนของตัวประมาณความแปรปรวนจากตัวอย่างประมาณความเอนเอียงของตัวประมาณความแปรปรวนจากตัวอย่าง และการประมาณความเบ้ของตัวประมาณความแปรปรวนจากตัวอย่างโดยวิธีบูตสแตรพและวิธีแจ็คไนฟ์ สำหรับประชากรที่มีการแจกแจงแบบปกติมาตรฐาน โดยใช้ตัวอย่างสุ่มขนาด 10 , 20 , 40 และ 80 ตามลำดับ จากผลการศึกษาพบว่า ในทุกขนาดตัวอย่าง การประมาณความแปรปรวนของตัวประมาณความแปรปรวนจากตัวอย่างโดยวิธีบูตสแตรพและวิธีแจ็คไนฟ์ จะให้ค่าประมาณใกล้เคียงกับความแปรปรวนของตัวประมาณความแปรปรวนจากตัวอย่างที่แท้จริงได้ดีพอกัน สำหรับการเปรียบเทียบการประมาณความเอนเอียงของตัวประมาณความแปรปรวนจากตัวอย่างและการประมาณความเบ้ของตัวประมาณความแปรปรวนจากตัวอย่าง โดยวิธีแจ็คไนฟ์และวิธีบูตสแตรพ ก็ให้ผลคล้ายคลึงกับที่กล่าวมาคือ มีค่าประมาณความเอนเอียงและค่าประมาณความเบ้ของตัวประมาณความแปรปรวนจากตัวอย่างโดยวิธีบูตสแตรพและวิธีแจ็คไนฟ์ไม่แตกต่างกัน

จึงได้ข้อสรุปว่า การใช้วิธีแจ็คไนฟ์และวิธีบูตสแตรพในการประมาณความแปรปรวนของตัวประมาณความแปรปรวนจากตัวอย่าง ประมาณความเอนเอียงของตัวประมาณความแปรปรวนจากตัวอย่าง และประมาณความเบ้ของตัวประมาณความแปรปรวนจากตัวอย่างเป็นวิธีที่มีประสิทธิภาพดีพอๆ กัน

3. วิธีประมาณความแปรปรวนแบบมีเงื่อนไข

วิธีประมาณความแปรปรวนแบบมีเงื่อนไข เป็นอีกวิธีหนึ่งสำหรับการประมาณความแปรปรวนของตัวประมาณ

ให้ X_i , i = 1 , 2 , … , n มีการแจกแจงเหมือนกัน เป็นอิสระกันและสมมาตรรอบ θ โดยมีฟังก์ชันการแจกแจง $F(x)$ ให้ $\Delta_i = |X_i - \theta|$, $i = 1,2,...,n$ คือ ขนาดระยะห่างของ X_i จาก θ แล้วจะได้ว่า สำหรับ Δ_i , i = 1 , 2 , … , n ทุกๆเซตของค่าสังเกตที่เป็นไปได้ในรูป $\theta \pm \Delta_1$, $\theta \pm \Delta_2$, … , $\theta \pm \Delta_n$ มีความน่าจะเป็นของการเกิดขึ้น คือ $\frac{1}{2^n}$

ตัวอย่างเช่น n = 3 ค่าสังเกตแบบมีเงื่อนไข จำนวน 2^3 ที่เป็นไปได้บนเงื่อนไข $\Delta_i = |x_i - \theta|$, i = 1 , 2 , 3 คือ

ชุดที่ 1	$\theta + \Delta_1$,	$\theta + \Delta_2$,	$\theta + \Delta_3$
ชุดที่ 2	$\theta - \Delta_1$,	$\theta + \Delta_2$,	$\theta + \Delta_3$
ชุดที่ 3	$\theta + \Delta_1$,	$\theta - \Delta_2$,	$\theta + \Delta_3$
ชุดที่ 4	$\theta - \Delta_1$,	$\theta - \Delta_2$,	$\theta + \Delta_3$
ชุดที่ 5	$\theta + \Delta_1$,	$\theta + \Delta_2$,	$\theta - \Delta_3$
ชุดที่ 6	$\theta - \Delta_1$,	$\theta + \Delta_2$,	$\theta - \Delta_3$
ชุดที่ 7	$\theta + \Delta_1$,	$\theta - \Delta_2$,	$\theta - \Delta_3$
ชุดที่ 8	$\theta - \Delta_1$,	$\theta - \Delta_2$,	$\theta - \Delta_3$

สมมติว่าประมาณ θ ด้วย $\hat{\theta}_\phi$ โดย $\hat{\theta}_\phi$ เป็นคำตอบของสมการ

$$T_\varphi(X, \hat{\theta}) = \sum_{i=1}^{n} \text{sign}\left(x_i - \hat{\theta}\right)\varphi\left(\left|x_i - \hat{\theta}\right|\right) \tag{3.1}$$

ดังนั้น ความแปรปรวนแบบมีเงื่อนไขของ $\hat{\theta}_{\phi}$ จะถูกนิยามดังนี้

บทนิยาม : ความแปรปรวนแบบมีเงื่อนไขของ $\hat{\theta}_{\phi}$ คือ ความแปรปรวนของเซตของคำตอบจากสมการ 3.1 จำนวน 2^n ชุด ที่ถูกสร้างขึ้นจากเซตของค่าสังเกตที่เป็นไปได้ทั้งหมดในรูป $\theta \pm \Delta_1$, $\theta \pm \Delta_2$, … , $\theta \pm \Delta_n$ จำนวน 2^n เซต ที่มีเงื่อนไขบน Δ_1 , Δ_2 , … , Δ_n

ตัวอย่างต่อไปนี้เป็นวิธีการหาความแปรปรวนแบบมีเงื่อนไขของ $\hat{\theta}_{\phi}$

ตัวอย่างที่ 1 ให้ $\phi(u) = u$ ดังนั้น

$$T_{\varphi}\ (X\,,\ \hat{\theta}) = \sum_{i=1}^{n} \operatorname{sign}\left(x_i - \hat{\theta}\right)\varphi\left(\left|x_i - \hat{\theta}\right|\right)$$

ซึ่งคำตอบที่สอดคล้องกับสมการคือค่าเฉลี่ยตัวอย่าง $\bar{x}$

เราสามารถเขียน $\bar{x} = \theta + \dfrac{\sum_{i=1}^{n} u_i}{n}$

เมื่อ $P(u_i = +\Delta_i) = P(u_i = -\Delta_i) = \dfrac{1}{2}$

และ $u_1, u_2, ..., u_n$ เป็นอิสระกันอย่างมีเงื่อนไข

ดังนั้น ความแปรปรวนแบบมีเงื่อนไขของ $\bar{x}$ ที่เขียนอยู่ในรูปของ $\hat{V}ar_{con}\left(\bar{x}\right)$ จึงเท่ากับ

$$\hat{V}ar_{con}\left(\bar{x}\right) = \frac{\sum_{i=1}^{n} (\bar{x} - \theta)^2}{n}$$

$$= \frac{\sum_{i=1}^{n} \left(\theta + \frac{\sum_{i=1}^{n} u_i}{n} - \theta\right)^2}{n}$$

$$= \frac{\sum_{i=1}^{n} \frac{\left(\sum_{i=1}^{n} u_i\right)^2}{n^2}}{n}$$

$$= \frac{n\dfrac{\left(\sum_{i=1}^{n} u_i\right)^2}{n^2}}{n} \quad = \frac{\sum_{i=1}^{n} u_i^2}{n^2} \quad = \frac{\sum_{i=1}^{n} \Delta_i^2}{n^2} \qquad = \frac{\sum_{i=1}^{n} (|x_i - \theta|)^2}{n^2}$$

ดังนั้น ในกรณีที่ $\phi(u) = u$ ความแปรปรวนแบบมีเงื่อนไขคือความแปรปรวนแบบไม่มีเงื่อนไขที่เราใช้กันตามปกติ

ตัวอย่างที่ 2 ถ้า $\phi(u) = 1$ สมการ (3.1) จะเปลี่ยนเป็น

$$\mathrm{T}_{\varphi}(\mathbf{X},\hat{\theta}) \;=\; \mathrm{S}_{\varphi}(\mathbf{X},\hat{\theta}) \;=\; \sum_{i=1}^{n} sign(x_i - \hat{\theta})$$

ซึ่งคือตัวสถิติเครื่องหมาย และค่ากลางที่สอดคล้องกับฟังก์ชันคือมัธยฐาน ซึ่งสำหรับ

$n = 3$

เซตของค่าสังเกตที่เป็นไปได้ ดังนี้

$\theta - \Delta_{(1)}$, $\theta - \Delta_{(2)}$, $\theta - \Delta_{(3)}$, median = $\theta - \Delta_{(2)}$

$\theta - \Delta_{(1)}$, $\theta - \Delta_{(2)}$, $\theta + \Delta_{(3)}$, median = $\theta - \Delta_{(1)}$

$\theta - \Delta_{(1)}$, $\theta + \Delta_{(2)}$, $\theta - \Delta_{(3)}$, median = $\theta - \Delta_{(1)}$

$\theta - \Delta_{(1)}$, $\theta + \Delta_{(2)}$, $\theta + \Delta_{(3)}$, median = $\theta + \Delta_{(2)}$

$\theta + \Delta_{(1)}$, $\theta - \Delta_{(2)}$, $\theta - \Delta_{(3)}$, median = $\theta - \Delta_{(2)}$

$\theta + \Delta_{(1)}$, $\theta - \Delta_{(2)}$, $\theta + \Delta_{(3)}$, median = $\theta + \Delta_{(1)}$

$\theta + \Delta_{(1)}$, $\theta + \Delta_{(2)}$, $\theta - \Delta_{(3)}$, median = $\theta + \Delta_{(1)}$

$\theta + \Delta_{(1)}$, $\theta + \Delta_{(2)}$, $\theta + \Delta_{(3)}$, median = $\theta + \Delta_{(2)}$

เมื่อ $\Delta_{(1)} < \Delta_{(2)} < \Delta_{(3)}$ คือ สถิติอันดับของ Δ_i , $i = 1,2,\ldots,n$

ดังนั้น ความแปรปรวนแบบมีเงื่อนไขของมัธยฐานตัวอย่างเท่ากับ $\frac{1}{2}\left(\Delta_{(1)}^2 + \Delta_{(2)}^2\right)$

เมื่อ $\Delta_{(1)} < \Delta_{(2)}$

ในทำนองเดียวกัน สำหรับ n = 5 ความแปรปรวนแบบมีเงื่อนไขของมัธยฐานคือ $\frac{6}{16}\Delta_{(1)}^2+\frac{6}{16}\Delta_{(2)}^2+\frac{4}{16}\Delta_{(3)}^2$ เมื่อ $\Delta_{(1)}<\Delta_{(2)}<\Delta_{(3)}$

ความแปรปรวนแบบมีเงื่อนไขสามารถใช้เหมือนกับความแปรปรวนแบบไม่มีเงื่อนไข ตัวอย่างเช่น ในการถ่วงน้ำหนักของตัวประมาณ 2 ตัว สมมติว่า ความแปรปรวนแบบมีเงื่อนไข ของ $\hat{\theta}_1$ และ $\hat{\theta}_2$ โดยที่ $\hat{\theta}_1$ และ $\hat{\theta}_2$ เป็นตัวประมาณของ θ ตัวเดียวกันและเป็นค่าประมาณร่วมคือ

$$\frac{n_1\hat{\theta}_1+n_2\hat{\theta}_2}{n_1+n_2}$$

ความคลาดเคลื่อนมาตรฐานแบบมีเงื่อนไขถ่วงน้ำหนัก คือ

$$\sqrt{\frac{n_1^2\ \mathrm{var}\left(\hat{\theta}_1\right)+n_2^2\ \mathrm{var}\left(\hat{\theta}_2\right)}{\left(n_1+n_2\right)^2}}$$

ช่วงความเชื่อมั่น 95% ในการประมาณ θ คือ

$$\frac{n_1\hat{\theta}_1+n_2\hat{\theta}_2}{n_1+n_2}\pm 1.65\sqrt{\frac{n_1^2\ \mathrm{var}\left(\hat{\theta}_1\right)+n_2^2\ \mathrm{var}\left(\hat{\theta}_2\right)}{\left(n_1+n_2\right)^2}}$$

แม้ว่าสัมประสิทธิ์ความเชื่อมั่น (95%) ได้มาจากความแปรปรวนแบบมีเงื่อนไข สัมประสิทธิ์ความเชื่อมั่นแบบไม่มีเงื่อนไขก็มีค่าเท่ากัน

ความแปรปรวนแบบมีเงื่อนไขมีความได้เปรียบกว่าความแปรปรวนแบบไม่มีเงื่อนไข เพราะความแปรปรวนแบบมีเงื่อนไขสามารถหาค่าได้เสมอ แต่ความแปรปรวนแบบไม่มีเงื่อนไขของ $\hat{\theta}_\phi$ อาจหาค่าไม่ได้ ตัวอย่างเช่น ค่าเฉลี่ยตัวอย่างของความแปรปรวนแบบไม่มีเงื่อนไขในการแจกแจงแบบ Cauchy จะไม่สามารถหาค่าได้ ในกรณีอื่น ค่าเฉลี่ยอาจจะหาค่าได้แต่ความแปรปรวนจะเท่ากับ ∞ ตัวอย่างเช่น ให้ $F(x)=1-\frac{1}{1+x^2}$, $x\geq 0$ ในกรณีนี้ $E\left(X\right)=-2$ และความแปรปรวนของ X เท่ากับ ∞ ดังนั้น ความแปรปรวนแบบไม่มีเงื่อนไขของค่าเฉลี่ยตัวอย่างจึงเท่ากับ ∞

ลักษณะเฉพาะของความแปรปรวนแบบมีเงื่อนไข คือ ความแปรปรวนแบบมีเงื่อนไขเป็นตัวประมาณค่าของความแปรปรวนแบบไม่มีเงื่อนไข เมื่อเป็นจำนวนจำกัด จากทฤษฎีเบื้องต้นเราทราบว่า

$$\operatorname{var}\left(\hat{\theta}_\phi\right) = \operatorname{var}\left(E\left(\hat{\theta}_\phi \mid \Delta\right)\right) + E\left(\operatorname{var}\left(\hat{\theta}_\phi \mid \Delta\right)\right) \tag{3.2}$$

เมื่อ $\Delta = (\Delta_1 , \Delta_2 , \ldots , \Delta_n)$

ถ้า $\hat{\theta}_\phi$ มีการแจกแจงแบบสมมาตรรอบ θ ดังนั้น $E\left(\hat{\theta}_\phi \mid \Delta\right) = \theta$ และ $\operatorname{var}(E\left(\hat{\theta}_\phi \mid \Delta\right)) = 0$ สมการ 3.2 ลดรูปเป็น

$$\operatorname{var}\left(\hat{\theta}_\phi\right) = E\left(\operatorname{var}\left(\hat{\theta}_\phi \mid \Delta\right)\right)$$

ดังนั้น ความแปรปรวนแบบไม่มีเงื่อนไขของ $\hat{\theta}_\phi$ เป็นค่าคาดหวังของความแปรปรวนแบบมีเงื่อนไขของ $\hat{\theta}_\phi$ เมื่อ $\hat{\theta}_\phi$ เป็นการแจกแจงแบบสมมาตรรอบ θ

การประมาณความแปรปรวนแบบมีเงื่อนไข

ในทางปฏิบัติความแปรปรวนแบบมีเงื่อนไขขึ้นอยู่กับ θ ซึ่ง θ เป็นพารามิเตอร์ที่ไม่ทราบค่า จึงต้องประมาณค่า θ ด้วย $\hat{\theta}_\phi$ ที่เป็นค่าตัวอย่างจากข้อมูลที่ได้จากการสังเกต ถ้า $\hat{\theta}_\phi$ เป็นตัวประมาณที่คงเส้นคงวาของ θ นั่นคือ $\left|\hat{\theta}_\phi - \theta\right| \xrightarrow{P} 0$ เมื่อ $n \to \infty$

ดังนั้น

$$\lim_{n\to\infty}\left|\hat{\Delta}_i - \Delta_i\right| = \lim_{n\to\infty}\left|x_i - \hat{\theta}_\phi - x_i + \theta\right| = \lim_{n\to\infty}\left|\hat{\theta}_\phi - \theta\right| = 0$$

ดังนั้น $\hat{\Delta}_i$, i = 1 , 2 , … , n เป็นตัวประมาณที่คงเส้นคงวาของ Δ_i , i = 1 , 2 , … , n แทน θ ด้วย $\hat{\theta}$ และ Δ_i ด้วย $\hat{\Delta}_i$, i = 1 , 2 , … , n ในนิยามของความแปรปรวนแบบมีเงื่อนไขได้แสดงตัวประมาณที่คงเส้นคงวาของความแปรปรวนแบบมีเงื่อนไขไว้ ดังนั้น ค่าประมาณความแปรปรวนแบบมีเงื่อนไขของ $\hat{\theta}_\phi$ คือ การหาความแปรปรวนของคำตอบจากสมการ (3.1) ที่ถูกสร้างโดยเซตของค่าสังเกตที่เป็นไปได้ทั้งหมดจำนวน 2^n ชุด ในเทอมของ $\theta \pm \Delta_1$, $\theta \pm \Delta_2$, … , $\theta \pm \Delta_n$ แทนด้วย $\hat{\theta}_\varphi \pm \Delta_1$, $\hat{\theta}_\varphi \pm \Delta_2$, … , $\hat{\theta}_\varphi \pm \Delta_n$ ซึ่งขั้นตอนในการคำนวณค่าประมาณความแปรปรวนแบบมีเงื่อนไขของ $\hat{\theta}_\phi$ หรือ $\left(\hat{\operatorname{var}}_{con1}\left(\hat{\theta}\right)\right)$ คือ

1) หาคำตอบของ $\hat{\theta}_{\phi}$ ในสมการ 3.1 โดยใช้ค่าสังเกตจากตัวอย่าง

2) กำหนดตัวอย่างที่เป็นไปได้ในเทอมของ $\theta \pm \Delta_1$, $\theta \pm \Delta_2$, … , $\theta \pm \Delta_n$

3) ดำเนินการขั้นตอนที่ 2 จำนวน 2^n ครั้ง

4) คำนวณความแปรปรวนที่ได้ในขั้นตอนที่ 3

การสร้างเซตจำนวน 2^n เซตของค่าสังเกตที่เป็นไปได้ อาจเป็นปัญหาเมื่อตัวอย่างมีขนาดใหญ่ ซึ่งจำเป็นต้องหาวิธีการเฉพาะสำหรับกรณีที่ตัวอย่างมีขนาดใหญ่ แล้วแก้สมการหาค่าของ $\hat{\theta}_{\phi}$ อย่างไรก็ตามในกรณีของมัธยฐานตัวอย่าง เราสามารถหาสูตรทั่วไปสำหรับประมาณความแปรปรวนแบบมีเงื่อนไข ซึ่งจะได้กล่าวต่อไปในหัวข้อวิธีการประมาณค่าความแปรปรวนแบบมีเงื่อนไขโดยใช้มัธยฐานตัวอย่าง การหาสูตรดังกล่าวได้ทำให้ง่ายที่จะหาค่าประมาณความแปรปรวนแบบมีเงื่อนไขของมัธยฐานในทางปฏิบัติ

ลักษณะเฉพาะที่น่าสนใจของการประมาณความแปรปรวนแบบมีเงื่อนไข คือ การประมาณความแปรปรวนแบบมีเงื่อนไขเป็นการหาลิมิตความเชื่อมั่นที่เป็นไปได้จริงๆ และนำไปสู่ค่าจริงของความแปรปรวนแบบมีเงื่อนไขของ $\hat{\theta}_{\phi}$ ซึ่งตัวอย่างต่อไปนี้จะใช้อธิบายวิธีการหาลิมิตความเชื่อมั่น สำหรับความแปรปรวนแบบมีเงื่อนไขของ $\hat{\theta}_{\phi}$:

สมมติค่าสังเกตจากตัวอย่างขนาด 5 คือ

3.89 4.23 4.47 5.15 5.72

กำหนดเงื่อนไขบน $\Delta_i = |x_i - \theta|$, $i = 1,2,\ldots,n$ เมื่อ θ ถูกแทนด้วย $\hat{\theta}_{\phi}$ = 4.47 ซึ่งเป็นมัธยฐานตัวอย่างที่เป็นไปได้ของมัธยฐานตัวอย่างจำนวน 32 ค่าคือ

$$\overbrace{3.89,\ldots,3.89}^{4},\overbrace{4.23,\ldots,4.23}^{6},\overbrace{4.47,\ldots,4.47}^{12},\overbrace{4.71,\ldots,4.71}^{6},\overbrace{5.05,\ldots,5.05}^{4}$$

ลิมิตความเชื่อมั่น $\left(1-\frac{2}{32}\right)100\%$ สำหรับ θ คือ $(3.89,5.05)$ ความแปรปรวนแบบมีเงื่อนไขของ $\hat{\theta}$ เท่ากับ $\frac{6}{16}\left(\Delta_{(1)}^2\right)+\frac{6}{16}\left(\Delta_{(2)}^2\right)+\frac{4}{16}\left(\Delta_{(3)}^2\right)$ เมื่อ $\Delta_{(1)} < \Delta_{(2)} < \Delta_{(3)}$

สมมติ

$\theta = 3.89$, $\Delta_{(1)} = 0$, $\Delta_{(2)} = .34$, $\Delta_{(3)} = .58$, $con.\mathrm{var} = 0.1274$

$\theta = 4.23$, $\Delta_{(1)} = 0$, $\Delta_{(2)} = .24$, $\Delta_{(3)} = .34$, $con.\mathrm{var} = 0.0505$

$\theta = 4.235$, $\Delta_{(1)} = 0.005$, $\Delta_{(2)} = .235$, $\Delta_{(3)} = .345$, $con.\mathrm{var} = 0.05047$

$\theta = 4.47$, $\Delta_{(1)} = 0$, $\Delta_{(2)} = .24$, $\Delta_{(3)} = .58$, $con.\mathrm{var} = 0.1057$

$\theta = 4.71$, $\Delta_{(1)} = .24$, $\Delta_{(2)} = .44$, $\Delta_{(3)} = .48$, $con.\mathrm{var} = 0.1518$

$\theta = 5.05$, $\Delta_{(1)} = .1$, $\Delta_{(2)} = .58$, $\Delta_{(3)} = .67$, $con.\mathrm{var} = 0.2420$

จากตัวอย่าง ช่วงความเชื่อมั่น $\left(1-\frac{2}{32}\right)100\%$ ของความแปรปรวนแบบมีเงื่อนไขจริง คือ $(0.05047, 0.242)$ การประมาณแบบจุดของช่วงความเชื่อมั่นของ $\hat{\theta}_{\phi} =$ 0.1057 ซึ่งหาได้ที่ $\hat{\theta}_{\phi} =$ 4.47

การประมาณความแปรปรวนของค่ากลางสำหรับตัวอย่าง 1 กลุ่ม

สมมติมีค่าสังเกตที่เป็นอิสระกันของตัวแปรสุ่ม X จำนวน n ค่า โดยที่ X มีฟังก์ชันการแจกแจงแบบต่อเนื่องและสมมาตรรอบ θ จากนิยาม ความแปรปรวนแบบมีเงื่อนไขของมัธยฐานตัวอย่าง คือ ความแปรปรวนของมัธยฐานตัวอย่างที่เป็นไปได้ทั้งหมดจำนวน 2^n เซตบนเงื่อนไข $\Delta_i = |x_i - \theta|$ เซตของค่าสังเกตที่เป็นไปได้ 2^n เซต สามารถเขียนในเทอมของ $\hat{x}_1$, $\hat{x}_2$, … , $\hat{x}_n$ เมื่อ $\hat{x}_i = \theta \pm \Delta_i$, i = 1 , 2 , … , n

ให้ $\Delta_{(1)}$, $\Delta_{(2)}$, ..., $\Delta_{(n)}$ เป็นสถิติอันดับของ Δ_i , i = 1 , 2 , … , n เมื่อ $n = 2m+1$ จะมีวิธีการของมัธยฐานที่เป็นไปได้ 3 วิธี คือ

1) ให้ $\theta - \Delta_{(i)}$, i = 1 , 2 , … , m+1 เป็นมัธยฐานของตัวอย่างที่เป็นไปได้ ถ้าเครื่องหมายสัมประสิทธิ์ของ $\Delta_{(i+1)}$, $\Delta_{(i+2)}$, … , $\Delta_{(n)}$, 1 , 2 , … , m+1 เป็นลบจำนวน m ครั้งจากจำนวนทั้งหมด $n-i$ ครั้ง โดยขณะที่สัมประสิทธิ์ของ $\Delta_{(1)}$, $\Delta_{(2)}$, … , $\Delta_{(i-1)}$, i = 1 , 2 , … , m+1 อาจจะมีเครื่องหมายอะไรก็ได้ ซึ่งเหตุการณ์ดังกล่าวเกิดขึ้นได้ $\sum_{i=2}^{m+1} 2^{i-1} C_m^{n-i}$ วิธี

2) ให้ $\theta + \Delta_{(i)}$, i = 1 , 2 , … , m+1 เป็นมัธยฐานตัวอย่างที่เป็นไปได้ ถ้าเครื่องหมายสัมประสิทธิ์ของ $\Delta_{(i+1)}$, $\Delta_{(i+2)}$, … , $\Delta_{(n)}$, i = 1 , 2 , … , m+1 เป็นลบจำนวน $m-i+1$ ครั้ง จากจำนวนทั้งหมด $n-i$ ครั้ง โดยที่สัมประสิทธิ์ของ $\Delta_{(1)}$, $\Delta_{(2)}$, … , $\Delta_{(i-1)}$, i = 1 , 2 , … , m+1 อาจจะ มีเครื่องหมายอะไรก็ได้ ซึ่งเหตุการณ์ดังกล่าวเกิดขึ้นได้ $\sum_{i=2}^{m+1} 2^{i-1} C_{m-i+1}^{n-i}$ วิธี

3) ให้ $\theta - \Delta_{(1)}$ หรือ $\theta + \Delta_{(1)}$, i = 1 , 2 , … , m+1 เป็นมัธฐานตัวอย่างที่เป็นไปได้ ถ้าเครื่องหมายสัมประสิทธิ์ของ $\Delta_{(2)}$, $\Delta_{(3)}$, … , $\Delta_{(n)}$ เป็นลบจำนวน m ครั้ง จากจำนวนทั้งหมด $n-i$ ครั้ง ซึ่งสามารถเกิดขึ้นได้ C_m^{n-i} วิธี

ค่าคาดหวังของมัธยฐานตัวอย่างแบบมีเงื่อนไข คือ

$$E_{con2}\left(\hat{\theta}\right) = \frac{1}{2^n}\left[\sum_{i=2}^{m+1} 2^{i-1} C_m^{n-i}\left(\theta - \Delta_{(i)}\right) + 2^{i-1} C_{m-i+1}^{n-i}\left(\theta + \Delta_{(i)}\right) + C_m^{n-1}\left(\theta - \Delta_{(i)}\right) + C_m^{n-1}\left(\theta + \Delta_{(1)}\right)\right]$$

$$= \theta$$

และความแปรปรวนแบบมีเงื่อนไขของมัธยฐานตัวอย่าง คือ

$$Var_{con2}\left(\hat{\theta}\right) = \frac{1}{2^n}\left[\sum_{i=2}^{m+1} 2^i C_m^{n-i}\left(\theta - \Delta_{(i)} - \theta\right)^2 + 2C_m^{n-1}\Delta_{(1)}^2\right]$$

$$= \frac{1}{2^n}\left[\sum_{i=2}^{m+1} 2^i C_m^{n-i}\Delta_{(i)}^2 + 2C_m^{n-1}\Delta_{(1)}^2\right]$$

เมื่อ $\Delta_{(1)} < \Delta_{(2)} < \ldots < \Delta_{(n)}$ เป็นสถิติอันดับของ Δ_i , i = 1 ,2 , … , n

ตัวอย่างเช่น $n = 3$

$$Var_{con2}\left(\hat{\theta}\right) = \frac{1}{8}\left(4\Delta_{(2)}^2 + 4\Delta_{(1)}^2\right)$$

$$= \frac{1}{4}\Delta_{(1)}^2 + \frac{1}{2}\Delta_{(2)}^2$$

หรือถ้า $n = 5$ จะได้ว่า

$$Var_{con2}\left(\hat{\theta}\right) = \frac{1}{2^5}\left(12\Delta_{(2)}^2 + 8\Delta_{(3)}^2 + 12\Delta_{(1)}^2\right)$$

$$= \frac{1}{8}\left(3\Delta_{(1)}^2 + 3\Delta_{(2)}^2 + 2\Delta_{(3)}^2\right)$$

เมื่อ $n = 2m$ ให้มัธยฐานของเซตตัวอย่างที่เป็นไปได้ถูกนิยามว่าคือค่าเฉลี่ยของสถิติอันดับตัวที่ m และ $m+1$ ของ $\theta \pm \Delta_1$, $\theta \pm \Delta_2$, … , $\theta \pm \Delta_n$ วิธีที่เป็นไปได้ในการหามัธยฐาน คือ

1) $\dfrac{\theta - \Delta_{(i)} + \theta - \Delta_{(j)}}{2}, i < j$ โดยที่ i = 2 , 3 , … , m ; j = 3 , 4 , … , m+1 คือ มัธยฐานของตัวอย่างที่เป็นไปได้ ถ้าเครื่องหมายของสัมประสิทธิ์ $\Delta_{(i+1)}$ ถึง $\Delta_{(j-1)}$ เป็นลบทั้งหมด ขณะที่สัมประสิทธิ์ของ $\Delta_{(j)}$ ถึง $\Delta_{(n)}$ เป็นลบจำนวน $m - j + 1$ ครั้ง จากทั้งหมด $n - j$ ครั้ง และสัมประสิทธิ์ของ $\Delta_{(1)}$ ถึง $\Delta_{(j-1)}$ มีเครื่องหมายอะไรก็ได้ซึ่งเกิดขึ้นได้ $\sum_{i=2}^{m}\sum_{j=3}^{m+1} 2^{i-1} C^{n-j}_{\frac{n}{2}-j+1}$ วิธี

2) $\dfrac{\theta - \Delta_{(i)} + \theta - \Delta_{(j)}}{2}, i < j$ โดยที่ i = 2 , 3 , … , m ; j = 3 , 4 , … , m+1 คือ มัธยฐานของตัวอย่างที่เป็นไปได้ ถ้าเครื่องหมายของสัมประสิทธิ์ $\Delta_{(i+1)}$ ถึง $\Delta_{(j-1)}$ i = 2 , 3 , … , m ; j = 3 , 4 , … , m+1 เป็นบวกทั้งหมด ขณะที่สัมประสิทธิ์ของ $\Delta_{(j)}$ ถึง $\Delta_{(n)}$ เป็นลบ $\left(\dfrac{n}{2} - 1\right)$ ครั้ง จากทั้งหมด $n - j$ ครั้ง และสัมประสิทธิ์ของ $\Delta_{(1)}$ ถึง $\Delta_{(i-1)}$ มีเครื่องหมายอะไรก็ได้ซึ่งเกิดขึ้นได้ $\sum_{i=2}^{m}\sum_{j=3}^{m+1} 2^{i-1} C^{n-j}_{\frac{n}{2}-1}$ วิธี

3) $\dfrac{\theta + \Delta_{(1)} + \theta - \Delta_{(j)}}{2}$ หรือ $\dfrac{\theta - \Delta_{(1)} + \theta + \Delta_{(j)}}{2}$ โดยที่ j = 2 , 3 , … , m+1 คือ มัธยฐานตัวอย่างที่เป็นไปได้ ถ้าเครื่องหมายของสัมประสิทธิ์ $\Delta_{(2)}$ ถึง $\Delta_{(j)}$ โดยที่ j = 2 , 3 , … , m+1 เป็นลบทั้งหมด ขณะที่สัมประสิทธิ์ของ $\Delta_{(j+1)}$ ถึง $\Delta_{(n)}$ เป็นลบ $\left(\dfrac{n}{2} - j + 1\right)$ ครั้ง จากทั้งหมด $n - j$ ครั้ง โดยที่ j = 2 , 3 , … , m+1 ซึ่งแต่ละกรณีเกิดขึ้นได้ $\sum_{j=2}^{m-1} C^{n-j}_{\frac{n}{2}-j+1}$ วิธี

4) $\dfrac{\theta + \Delta_{(1)} + \theta - \Delta_{(j)}}{2}$ หรือ $\dfrac{\theta - \Delta_{(1)} + \theta + \Delta_{(j)}}{2}$ คือ มัธยฐานตัวอย่างที่เป็นไปได้ ถ้าเครื่องหมายของสัมประสิทธิ์ $\Delta_{(2)}$ ถึง $\Delta_{(j-1)}$ โดยที่ j = 2 , 3 , … , m+1 เป็นบวกทั้งหมด และสัมประสิทธิ์ของ $\Delta_{(j+1)}$ ถึง $\Delta_{(n)}$ เป็นลบ $\left(\dfrac{n}{2} - 1\right)$ ครั้ง จากทั้งหมด $n - j$ ครั้ง ซึ่งแต่ละกรณีเกิดขึ้นได้ $\sum_{j=2}^{m+1} C^{n-j}_{\frac{n}{2}-1}$ วิธี

เราสามารถแสดงให้เห็นว่า $E_{con2}(\hat{\theta}) = \theta$ และความแปรปรวนแบบมีเงื่อนไขของมัธยฐานตัวอย่าง คือ

$$\hat{V}ar_{con}\left(\hat{\theta}\right)$$

$$= \frac{1}{2^n}\left[\sum_{i=2}^{m}\sum_{j=3}^{m+1}\left\{\begin{array}{l} 2^{i-1}C^{n-j}_{\frac{n}{2}-j+1}\left(\dfrac{\theta+\Delta_{(i)}+\theta+\Delta_{(j)}}{2}-\theta\right)^2 \\ +\,2^{i-1}C^{n-j}_{\frac{n}{2}-1}\left(\dfrac{\theta-\Delta_{(i)}+\theta-\Delta_{(j)}}{2}-\theta\right)^2 \end{array}\right\} + \sum_{j=2}^{m+1}\left[\begin{array}{l} C^{n-j}_{\frac{n}{2}-j+1}\left\{\left(\dfrac{\theta+\Delta_{(i)}+\theta-\Delta_{(j)}}{2}-\theta\right)^2+\left(\dfrac{\theta-\Delta_{(1)}+\theta-\Delta_{(j)}}{2}-\theta\right)^2\right\} \\ +\,C^{n-j}_{\frac{n}{2}-1}\left\{\left(\dfrac{\theta+\Delta_{(1)}+\theta-\Delta_{(j)}}{2}-\theta\right)^2+\left(\dfrac{\theta-\Delta_{(1)}+\theta-\Delta_{(j)}}{2}-\theta\right)^2\right\} \end{array}\right]\right]$$

ดังนั้น

$$\hat{V}ar_{con}\left(\hat{\theta}\right)=\frac{1}{2^n}\left[\sum_{i=2}^{m}\sum_{j=3}^{m+1}2^{i}\,C^{n-j}_{\frac{n}{2}-j+1}\left(\frac{\Delta_{(i)}+\Delta_{(j)}}{4}\right)^2+\sum_{j=2}^{m+1}C^{n-j}_{\frac{n}{2}-j+1}\left(\Delta^2_{(1)}+\Delta^2_{(j)}\right)^2\right]$$

เมื่อ $\Delta_{(1)}<\Delta_{(2)}<\Delta_{(3)}$ เป็นสถิติอันดับของ Δ_i , $i=1,2,...,n$

ตัวอย่างเช่น $n=4$

$$\hat{V}ar_{con}\left(\hat{\theta}\right)=\frac{1}{2^4}\left[\left(\Delta_{(2)}+\Delta_{(3)}\right)^2+2\left(\Delta^2_{(1)}+\Delta^2_{(2)}\right)+\Delta^2_{(1)}+\Delta^2_{(3)}\right]$$

$$=\frac{1}{2^4}\left[\left(\Delta_{(2)}+\Delta_{(3)}\right)^2+3\Delta^2_{(1)}+2\Delta^2_{(2)}+\Delta^2_{(3)}\right]$$

4. งานวิจัยที่เกี่ยวข้องกับการประมาณความแปรปรวน

P. Ratanaprasert (1987) ได้ศึกษาการประมาณความแปรปรวนของ intercept ในการวิเคราะห์การถดถอยเชิงเส้นตรงในกรณีที่ไม่ทราบการแจกแจง โดยใช้ความแปรปรวนแบบมีเงื่อนไขและวิธีการอื่น พบว่าความเอนเอียง (bias) ในการประมาณค่าความแปรปรวนโดยใช้ความแปรปรวนแบบมีเงื่อนไขจะลดลงเมื่อ n มีขนาดใหญ่ เนื่องจากมัธยฐานตัวอย่างมีการแจกแจงแบบสมมาตร ค่าคาดหวังที่ได้จากความแปรปรวนแบบมีเงื่อนไขเท่ากับความแปรปรวนแบบไม่มีเงื่อนไข ดังนั้นความเอนเอียง (bias) ในการประมาณความแปรปรวนแบบมีเงื่อนไขซึ่งไปประมาณความแปรปรวนแบบไม่มีเงื่อนไขจึงลดลง

Jun Shao and Wu (1989) ได้เสนอวิธี Delete-d jackknife variance estimation ไปประมาณความแปรปรวนของค่ามัธยฐานของประชากร เมื่อขนาดตัวอย่าง n เพิ่มขึ้น ซึ่งแสดงว่าวิธี Delete-d jackknife variance estimation จะได้ตัวประมาณที่มีความคงเส้นคงวา และเป็นวิธีการประมาณความแปรปรวนของค่ามัธยฐานจากตัวอย่างอีกวิธีหนึ่ง ซึ่ง Jun Shao (1989) ได้พัฒนาวิธีแจ็คไนฟ์ที่ยังคงรูปแบบวิธี Delete-d jackknife variance estimation ไว้ แต่ได้ปรับ

วิธีการสุ่มตัวอย่าง เมื่อใช้สำหรับกรณีที่ตัวอย่างสุ่มขนาด n เพียงชุดเดียวมีขนาดใหญ่มากๆ วิธีการประมาณความแปรปรวนของตัวประมาณก็คือ ทำการสุ่มตัวอย่าง 2 ขั้นตอน คือ ขั้นตอนแรกจะกระทำเช่นเดียวกับวิธี Delete-d jackknife variance estimation คือ สร้างตัวอย่างใหม่จากตัวอย่างสุ่มขนาด n เพียงชุดเดียว โดยตัดหน่วยของตัวอย่างครั้งละ d หน่วย ซึ่งจะได้ตัวอย่างใหม่ที่เป็นไปได้ คือ $\binom{n}{d}$ เมื่อ $d < n$ และขั้นที่สอง ทำการสุ่มตัวอย่างแบบไม่แทนที่จำนวน m ชุด จากตัวอย่างใหม่ซึ่งมีจำนวน $\binom{n}{d}$ ชุดดังกล่าว สำหรับการประมาณความแปรปรวนของตัวประมาณก็ทำเช่นเดียวกับวิธี Delete-d jackknife variance estimation แต่เป็นการประมาณความแปรปรวนของตัวอย่างใหม่ จำนวน m ชุด และเรียกวิธีการประมาณความแปรปรวนของตัวประมาณโดยวิธีแจ็คไนฟ์แบบนี้ว่า Jackknife-sampling variance estimation (JSVE) ซึ่งพบว่าการใช้ Jackknife-sampling variance estimation และวิธี Delete-d jackknife variance estimation มีประสิทธิภาพไม่แตกต่างกัน

บทที่ 3
วิธีดำเนินการวิจัย

ในการดำเนินการวิจัยครั้งนี้ต้องการเปรียบเทียบประสิทธิภาพการประมาณความแปรปรวนของมัธยฐานโดยวิธีความแปรปรวนแบบมีเงื่อนไขที่ได้นำเสนอขึ้นใหม่ กับวิธีแจ็คไนฟ์ และวิธีบูตสแตรพ ซึ่งเป็นวิธีที่นิยมใช้กันในปัจจุบันในกรณีที่ต้องการประมาณความแปรปรวนของตัวสถิติที่มีฟังก์ชันที่ซับซ้อนแต่ในการวิจัยครั้งนี้จะศึกษาเฉพาะการประมาณความแปรปรวนของมัธยฐานตัวอย่าง สำหรับการศึกษาจะใช้การจำลองแบบโดยใช้โปรแกรม FORTRAN 6 มีรายละเอียดและวิธีการดำเนินการวิจัยดังนี้

1. ประชากรที่นำมาศึกษา ผู้วิจัยได้เลือกประชากรที่มีลักษณะดังนี้

แบบที่ 1 ประชากรที่มีการแจกแจงแบบปกติ ผู้วิจัยได้เลือกประชากรที่มีการแจกแจงแบบปกติมาตรฐานเพื่อใช้ในการจำลองแบบ

แบบที่ 2 ประชากรที่มีการแจกแจงแบบสมมาตรแต่ไม่ใช่การแจกแจงแบบปกติ ผู้วิจัยได้ใช้การแจกแจงแบบเอกรูปในช่วง (0 , 1) เพื่อเป็นตัวแทนของประชากรชนิดนี้

แบบที่ 3 การแจกแจงของประชากรที่ไม่ใช่ทั้งการแจกแจงแบบปกติและการแจกแจงแบบสมมาตร ผู้วิจัยได้ใช้การแจกแจงผสมระหว่างการแจกแจงแบบปกติ 2 การแจกแจงที่มีค่าเฉลี่ยและความแปรปรวนต่างกัน โดยในที่นี้กำหนดให้การแจกแจงหลักเป็นการแจกแจงแบบปกติมาตรฐานผสมกับการแจกแจงแบบปกติที่มีค่าเฉลี่ย 10 และความแปรปรวน 4 โดยให้ความน่าจะเป็นของการปลอมปนเท่ากับ 5 และ 10 ซึ่งเราเรียกการแจกแจงในลักษณะนี้ว่าการแจกแจงปกติปลอมปน (Contaminated Normal Distribution)

2. วิธีการจำลองแบบ มีขั้นตอนในการดำเนินการจำลองแบบดังนี้

2.1 สุ่มตัวอย่างขนาด n จากประชากรที่มีการแจกแจงในลักษณะต่างๆ ที่กำหนด โดยกำหนดให้ขนาดตัวอย่าง $n = 6 , 10$ และ 15 แล้วคำนวณค่ามัธยฐานของตัวอย่างแต่ละชุด

2.2 จากตัวอย่างที่สุ่มได้ในข้อ 2.1 คำนวณค่าความแปรปรวนของมัธยฐานโดยใช้วิธีการแจ็คไนฟ์ วิธีบูตสแตรพ และวิธีความแปรปรวนแบบมีเงื่อนไข เนื่องจากวิธีบูตสแตรพการประมาณความแปรปรวนของตัวสถิติขึ้นกับจำนวนครั้งของการสุ่มตัวอย่างซ้ำ ในการวิจัยครั้งนี้จะกำหนดจำนวนครั้งของการสุ่มตัวอย่างซ้ำสำหรับวิธีบูตสแตรพให้เท่ากับ 500 ครั้ง

2.3 ดำเนินการตามข้อ 2.1 – 2.2 แล้วคำนวณค่าเฉลี่ยของตัวประมาณความแปรปรวนของค่ามัธยฐาน และความแปรปรวนของตัวประมาณความแปรปรวนของค่ามัธยฐานตัวอย่างด้วยวิธีการแจ็คไนฟ์ วิธีบูตสแตรพ และวิธีความแปรปรวนแบบมีเงื่อนไข ตามลำดับ

สำหรับการคำนวณค่าเฉลี่ยของค่าประมาณความแปรปรวนของค่ามัธยฐานและความแปรปรวนของตัวประมาณความแปรปรวนของค่าเฉลี่ยตัวอย่างแต่ละวิธีดังนี้

2.3.1 กรณีของแจ็คไนฟ์

ให้ $\hat{V}ar_{Ji}(\overline{X})$ เป็นค่าประมาณความแปรปรวนของค่ามัธยฐานตัวอย่างด้วยวิธีแจ็คไนฟ์ในการสุ่มตัวอย่างครั้งที่ i , $\mathrm{i}=1, 2, ..., 1000$ ค่าเฉลี่ยและความแปรปรวนของค่าประมาณความแปรปรวนของค่ามัธยฐานตัวอย่างด้วยวิธีแจ็คไนฟ์หาได้ดังนี้

2.3.1.1 ค่าเฉลี่ยของค่าประมาณความแปรปรวนของค่ามัธยฐานตัวอย่าง คือ

$$\overline{\hat{V}ar}_J(\tilde{X}) = \frac{\sum_{i=1}^{1000} \hat{V}ar_{Ji}\left(\tilde{X}\right)}{1000}$$

2.3.1.2 ค่าความแปรปรวนของค่าประมาณความแปรปรวนของค่ามัธยฐานตัวอย่าง คือ

$$Var(\hat{V}ar_J(\tilde{X})) = \frac{\sum_{i=1}^{1000}\left(\hat{V}ar_{Ji}\left(\tilde{X}\right) - \overline{\hat{V}ar}_J\left(\tilde{X}\right)\right)^2}{1000}$$

2.3.2 กรณีของบูตสแตรพ

ให้ $\hat{V}ar_{Bi}(\overline{X}^*)$ เป็นค่าประมาณความแปรปรวนของค่ามัธยฐานตัวอย่างด้วยวิธีบูตสแตรพในการสุ่มตัวอย่างครั้งที่ i , $i = 1,2,...,1000$ ค่าเฉลี่ยและค่าความแปรปรวนของค่าประมาณความแปรปรวนของค่ามัธยฐานตัวอย่างด้วยวิธีบูตสแตรพหาได้ดังนี้

2.3.2.1 ค่าเฉลี่ยของค่าประมาณความแปรปรวนของค่ามัธยฐานตัวอย่าง คือ

$$\overline{\hat{V}ar}_B(\tilde{X}) = \frac{\sum_{i=1}^{1000} \hat{V}ar_B\left(\tilde{X}\right)}{1000}$$

2.3.2.2 ค่าความแปรปรวนของค่าประมาณความแปรปรวนของค่ามัธยฐานตัวอย่าง คือ

$$Var(\hat{V}ar_B(\tilde{X})) = \frac{\sum_{i=1}^{1000}\left(\hat{V}ar_{Bi}\left(\tilde{X}\right) - \overline{\hat{V}ar}_B\left(\tilde{X}\right)\right)^2}{1000}$$

2.3.3 กรณีของความแปรปรวนแบบมีเงื่อนไข

ให้ $\hat{V}ar_{con}\left(\tilde{X}\right)$ เป็นค่าประมาณความแปรปรวนของมัธยฐานตัวอย่างด้วยวิธีความแปรปรวนแบบมีเงื่อนไข ในการสุ่มตัวอย่างครั้งที่ i , $i = 1,2,...,100$ ซึ่งมีวิธีดำเนินการดังนี้

2.3.3.1 สุ่มตัวอย่างขนาด $n =$ 6 , 10 และ 15 จากประชากรที่กำหนด

2.3.3.2 สำหรับแต่ละตัวอย่างสุ่มตามขนาดต่างๆ คำนวณค่ามัธยฐานจากตัวอย่างคือ $\tilde{x}$

2.3.3.3 หาค่า $\Delta_i = \left|x_i - \tilde{x}\right|$, $i = 1,2,...,n$

2.3.3.4 หาค่า $\Delta_{(1)} < \Delta_{(2)} < ... < \Delta_{(n)}$ ซึ่งเป็นสถิติอันดับของ Δ_i , $i = 1,2,...,n$ คำนวณค่า $\hat{V}ar_{con}\left(\tilde{X}\right) = \frac{1}{2^n}\left[\sum_{i=2}^{m+1} 2^i C_m^{n-i} \Delta_{(i)}^2 + 2C_m^{n-1}\Delta_{(1)}^2\right]$ กรณีที่ n เป็นจำนวนคี่ หรือสมการ

$$\hat{V}ar_{con}\left(\tilde{X}\right) = \frac{1}{2^n}\left[\sum_{i=2}^{m}\sum_{j=3}^{m+1} 2^i C_{\frac{n}{2}-j+1}^{n-j}\left(\frac{\Delta_{(i)} + \Delta_{(j)}}{4}\right)^2 + \sum_{j=2}^{m+1} C_{\frac{n}{2}-j+1}^{n-j}\left(\Delta_{(1)}^2 + \Delta_{(j)}^2\right)^2\right]$$

กรณีที่ n เป็นจำนวนคู่ เพื่อหาค่าความแปรปรวนของ $\tilde{x}$ ที่แทนด้วยสัญลักษณ์ $\hat{V}ar_{con}\left(\tilde{X}\right)$

จากนั้นจึงคำนวนค่าเฉลี่ยและค่าความแปรปรวนของค่าประมาณความแปรปรวนของค่ามัธยฐานตัวอย่างดังนี้

2.3.3.5 ค่าเฉลี่ยของค่าประมาณความแปรปรวนของค่ามัธยฐานตัวอย่าง คือ

$$\overline{\hat{V}ar}_{con}(\tilde{X}) = \frac{\sum_{i=1}^{1000} \hat{V}ar_{con}\left(\tilde{X}\right)}{1000}$$

2.3.3.6 ค่าความแปรปรวนของค่าประมาณความแปรปรวนของค่ามัธยฐานตัวอย่าง คือ

$$Var(\hat{V}ar_{con}(\tilde{X})) = \frac{\sum_{i=1}^{1000} \left(\hat{V}ar_{con\,i}\left(\tilde{X}\right) - \overline{\hat{V}ar}_{con}\left(\tilde{X}\right)\right)^2}{1000}$$

บทที่ 4
ผลการวิจัย

จากการศึกษา วิธีการประมาณความแปรปรวนของตัวประมาณด้วยความแปรปรวนแบบมีเงื่อนไข เพื่อเปรียบเทียบผลกับ วิธีการประมาณความแปรปรวนของตัวประมาณด้วยวิธีแจ็คไนฟ์ และวิธีบูตสแตรพ : กรณีศึกษาการประมาณความแปรปรวนของมัธยฐาน เกณฑ์ที่ใช้ในการเปรียบเทียบ คือ พิจารณาว่าวิธีการประมาณความแปรปรวนของมัธยฐานตัวอย่างในแต่ละวิธี มีค่าประมาณแตกต่างจากค่าความแปรปรวนของมัธยฐานตัวอย่างที่แท้จริง มากน้อยเพียงใด โดยถือว่าวิธีประมาณความแปรปรวนของมัธยฐานตัวอย่างวิธีใดให้ค่าประมาณใกล้เคียงกับค่าความแปรปรวนแบบ Empirical ของมัธยฐานตัวอย่าง และมีความแปรปรวนของค่าประมาณต่ำกว่าย่อมจะมีประสิทธิภาพมากกว่า ซึ่งผลการจำลองแบบสำหรับการแจกแจงแบบต่างๆ เมื่อตัวอย่างสุ่มมีขนาด 6 , 10 และ 15 ได้แสดงไว้ในตาราง ดังต่อไปนี้

ตารางที่ 1 ความแปรปรวนแบบ Empirical ของมัธยฐานตัวอย่างเทียบกับค่าประมาณความแปรปรวนของมัธยฐานด้วยวิธีความแปรปรวนแบบมีเงื่อนไข วิธีแจ็คไนฟ์ และวิธีบูตสแตรพ เมื่อกำหนดให้ประชากรที่ตัวอย่างสุ่มมามีการแจกแจงแบบปกติมาตรฐาน สำหรับตัวอย่างสุ่มขนาด 6 , 10 และ 15

ขนาดตัวอย่าง	ความแปรปรวนแบบ Empirical ของมัธยฐานตัวอย่าง (1)	วิธีการประมาณค่าความแปรปรวน	ค่าเฉลี่ยของค่าประมาณความแปรปรวนของมัธยฐานตัวอย่าง (2)	ผลต่างระหว่าง |(1)-(2)|	ความแปรปรวนของค่าประมาณความแปรปรวนของมัธยฐานตัวอย่าง
n = 6	0.2076334506	แจ็คไนฟ์	0.0714602694	0.136173182	0.0143062174
		บูตสแตรพ	0.2563662231	0.048732772*	0.0353542939
		ความแปรปรวนแบบมีเงื่อนไข	0.0761446804	0.131488771	0.0141683090

ตารางที่ 1 (ต่อ)

ขนาดตัวอย่าง	ความแปรปรวนแบบ Empirical ของมัธยฐานตัวอย่าง (1)	วิธีการประมาณค่าความแปรปรวน	ค่าเฉลี่ยของค่าประมาณความแปรปรวนของมัธยฐานตัวอย่าง (2)	ผลต่างระหว่าง |(1)-(2)|	ความแปรปรวนของค่าประมาณความแปรปรวนของมัธยฐานตัวอย่าง
n = 10	0.1383202225	แจ็คไนฟ์	0.0293514263	0.108968797	0.0031724249
		บูตสแตรพ	0.1710952371	0.032775014	0.0143725043
		ความแปรปรวนแบบมีเงื่อนไข	0.1562902927	0.01797007*	0.0225083679
n = 15	0.1056057662	แจ็คไนฟ์	0.0098411590	0.095764607	0.0001702658
		บูตสแตรพ	0.1301026791	0.024496913*	0.0074130977
		ความแปรปรวนแบบมีเงื่อนไข	0.0717778951	0.033827871	0.0033557469

เครื่องหมาย * คือ ค่าเฉลี่ยของค่าประมาณความแปรปรวนของมัธยฐานตัวอย่างที่ใกล้เคียงกับความแปรปรวนแบบ Empirical ของมัธยฐานตัวอย่างมากที่สุดสำหรับแต่ละขนาดตัวอย่าง

จากตารางที่ 1 เมื่อกำหนดให้ตัวอย่างถูกสุ่มมาจากประชากรที่มีการแจกแจงแบบปกติมาตรฐาน พบว่า เมื่อใช้ตัวอย่างขนาด 6 ค่าเฉลี่ยของค่าประมาณความแปรปรวนของมัธยฐานตัวอย่างแต่ละวิธีไม่แตกต่างจากความแปรปรวนแบบ Empirical ของมัธยฐานตัวอย่างมากนัก โดยพบว่าวิธีการประมาณความแปรปรวนของมัธยฐานตัวอย่างด้วยวิธีบูตสแตรพ มีค่าประมาณความแปรปรวนของมัธยฐานตัวอย่างโดยเฉลี่ยใกล้เคียงกับความแปรปรวนแบบ Empirical ของมัธยฐานตัวอย่างมากที่สุด แต่ก็มีความแปรปรวนของค่าประมาณความแปรปรวนของมัธยฐานตัวอย่างสูงกว่าวิธีการประมาณความแปรปรวนของมัธยฐานตัวอย่างด้วยวิธีแจ็คไนฟ์และวิธีความแปรปรวนแบบมีเงื่อนไข

เมื่อสุ่มตัวอย่างจากประชากรที่มีขนาดตัวอย่างเท่ากับ 10 พบว่า ค่าเฉลี่ยของตัวประมาณความแปรปรวนของมัธยฐานตัวอย่างโดยวิธีความแปรปรวนแบบมีเงื่อนไขมีค่าใกล้เคียงกับความแปรปรวนแบบ Empirical ของมัธยฐานตัวอย่างมากที่สุด แต่ก็มีความแปรปรวน

ของค่าประมาณความแปรปรวนของมัธยฐานตัวอย่างสูงกว่าวิธีการประมาณความแปรปรวนของมัธยฐานตัวอย่างด้วยวิธีแจ็คไนฟ์และวิธีบูตสแตรพ

เมื่อเพิ่มขนาดของตัวอย่างสุ่มจากประชากรเป็น 15 พบว่า วิธีการประมาณความแปรปรวนของมัธยฐานตัวอย่างด้วยวิธีบูตสแตรพ มีค่าประมาณความแปรปรวนของมัธยฐานตัวอย่างโดยเฉลี่ยใกล้เคียงกับความแปรปรวนแบบ Empirical ของมัธยฐานตัวอย่างมากที่สุด สำหรับวิธีประมาณความแปรปรวนแบบมีเงื่อนไขซึ่งมีค่าใกล้เคียงกับความแปรปรวนแบบ Empirical ของมัธยฐานเป็นอันดับถัดมา พบว่ามีความแปรปรวนของค่าประมาณความแปรปรวนของมัธยฐานตัวอย่างต่ำกว่าวิธีบูตสแตรพ แต่มีความแปรปรวนของค่าประมาณความแปรปรวนของมัธยฐานตัวอย่างสูงกว่าวิธีการประมาณความแปรปรวนของมัธยฐานตัวอย่างด้วยวิธีแจ็คไนฟ์และวิธีความแปรปรวนแบบมีเงื่อนไข

ซึ่งแสดงว่า วิธีการประมาณความแปรปรวนของมัธยฐานตัวอย่างโดยวิธีความแปรปรวนแบบมีเงื่อนไข สามารถใช้ได้ดีไม่แตกต่างกับวิธีบูตสแตรพมากนัก และดีกว่าวิธีการประมาณความแปรปรวนโดยวิธีแจ็คไนฟ์

ตารางที่ 2 ความแปรปรวนแบบ Empirical ของมัธยฐานตัวอย่างเทียบกับค่าประมาณความแปรปรวนของมัธยฐานด้วยวิธีความแปรปรวนแบบมีเงื่อนไข วิธีแจ็คไนฟ์ และวิธีบูตสแตรพ เมื่อกำหนดให้ประชากรที่ตัวอย่างสุ่มมามีการแจกแจงแบบเอกรูปในช่วง (0,1) สำหรับตัวอย่างสุ่มขนาด 6, 10 และ 15

ขนาดตัวอย่าง	ความแปรปรวนแบบ Empirical ของมัธยฐานตัวอย่าง (1)	วิธีการประมาณค่าความแปรปรวน	ค่าเฉลี่ยของค่าประมาณความแปรปรวนของมัธยฐานตัวอย่าง (2)	ผลต่างระหว่าง |(1)-(2)|	ความแปรปรวนของค่าประมาณความแปรปรวนของมัธยฐานตัวอย่าง
n = 6	0.0262472443	แจ็คไนฟ์	0.0092199249	0.017027319	0.0002411406
		บูตสแตรพ	0.0259435009	0.000303743*	0.0002868861
		ความแปรปรวนแบบมีเงื่อนไข	0.0052299602	0.00522996	0.0000300247

ตารางที่ 2 (ต่อ)

ขนาดตัวอย่าง	ความแปรปรวนแบบ Empirical ของมัธยฐานตัวอย่าง (1)	วิธีการประมาณค่าความแปรปรวน	ค่าเฉลี่ยของค่าประมาณความแปรปรวนของมัธยฐานตัวอย่าง (2)	ผลต่างระหว่าง |(1)-(2)|	ความแปรปรวนของค่าประมาณความแปรปรวนของมัธยฐานตัวอย่าง
n = 10	0.0181109272	แจ็คไนฟ์	0.0038350313	0.014275896	0.0000468357
		บูตสแตรพ	0.0202840213	0.002173094*	0.0001516681
		ความแปรปรวนแบบมีเงื่อนไข	0.0185716040	0.018571604	0.0002231612
n = 15	0.0148498723	แจ็คไนฟ์	0.0013576042	0.013492268	0.0000036409
		บูตสแตรพ	0.0165823027	0.001732431*	0.0001153397
		ความแปรปรวนแบบมีเงื่อนไข	0.0091631664	0.009163166	0.0000597283

เครื่องหมาย * คือ ค่าเฉลี่ยของค่าประมาณความแปรปรวนของมัธยฐานตัวอย่างที่ใกล้เคียงกับความแปรปรวนแบบ Empirical ของมัธยฐานตัวอย่างมากที่สุดสำหรับแต่ละขนาดตัวอย่าง

จากตารางที่ 2 เมื่อกำหนดให้ตัวอย่างถูกสุ่มมาจากประชากรที่มีการแจกแจงแบบเอกรูป ในช่วง (0,1) พบว่า เมื่อใช้ตัวอย่างขนาด 6 , 10 และ 15 สำหรับทุกขนาดตัวอย่างสุ่ม ค่าเฉลี่ยของตัวประมาณความแปรปรวนของมัธยฐานตัวอย่างในแต่ละวิธีไม่แตกต่างจากความแปรปรวนแบบ Empirical ของมัธยฐานตัวอย่างมากนัก โดยพบว่าวิธีการประมาณความแปรปรวนของมัธยฐานตัวอย่างด้วยวิธีบูตสแตรพ มีค่าประมาณความแปรปรวนของมัธยฐานตัวอย่างโดยเฉลี่ยใกล้เคียงกับความแปรปรวนแบบ Empirical ของมัธยฐานตัวอย่างมากที่สุดสำหรับทุกขนาดตัวอย่าง แต่ก็มีความแปรปรวนของค่าประมาณความแปรปรวนของมัธยฐานตัวอย่างสูงกว่าวิธีการประมาณความแปรปรวนของมัธยฐานตัวอย่างด้วยวิธีแจ็คไนฟ์ สำหรับทุกขนาดตัวอย่างด้วยเช่นกัน

ซึ่งแสดงว่า วิธีการประมาณความแปรปรวนของมัธยฐานตัวอย่างด้วยวิธีบูตสแตรพ สามารถใช้ได้ดีที่สุด เมื่อเปรียบเทียบกับวิธีการประมาณความแปรปรวนของมัธยฐานตัวอย่างด้วยวิธีความแปรปรวนแบบมีเงื่อนไขและวิธีแจ็คไนฟ์

ตารางที่ 3 ความแปรปรวนแบบ Empirical ของมัธยฐานตัวอย่างเทียบกับค่าประมาณความแปรปรวนของมัธยฐานด้วยวิธีความแปรปรวนแบบมีเงื่อนไข วิธีแจ็คไนฟ์ และวิธีบูตสแตรพ เมื่อกำหนดให้ประชากรที่ตัวอย่างสุ่มมามีการแจกแจงแบบปกติปลอมปน (Contaminated Normal Distribution) ซึ่งในที่นี้กำหนดการแจกแจงแบบ $CN(0,1,p,10,4)$ เมื่อ $p = 0.05$ สำหรับตัวอย่างสุ่มขนาด 6 , 10 และ 15

ขนาดตัวอย่าง	ความแปรปรวนแบบ Empirical ของมัธยฐานตัวอย่าง (1)	วิธีการประมาณค่าความแปรปรวน	ค่าเฉลี่ยของค่าประมาณความแปรปรวนของมัธยฐานตัวอย่าง (2)	ผลต่างระหว่าง |(1)-(2)|	ความแปรปรวนของค่าประมาณความแปรปรวนของมัธยฐานตัวอย่าง
n = 6	0.2545497715	แจ็คไนฟ์	0.0832298547	0.171319917	0.0444880351
		บูตสแตรพ	1.0844622850	0.829912513*	4.3128194809
		ความแปรปรวนแบบมีเงื่อนไข	0.1387928873	0.115756885	1.0598276854
n = 10	0.1530686915	แจ็คไนฟ์	0.0328266770	0.120242015	0.0032324761
		บูตสแตรพ	0.3968350291	0.243766337	0.8363263607
		ความแปรปรวนแบบมีเงื่อนไข	0.1871861517*	0.03411746*	0.0318775252
n = 15	0.1198006598	แจ็คไนฟ์	0.0106343785	0.109166282	0.0002531526
		บูตสแตรพ	0.1650266796	0.04522602	0.0627482608
		ความแปรปรวนแบบมีเงื่อนไข	0.0787628442*	0.041037816*	0.0045396439

เครื่องหมาย * คือ ค่าเฉลี่ยของค่าประมาณความแปรปรวนของมัธยฐานตัวอย่างที่ใกล้เคียงกับความแปรปรวนแบบ Empirical ของมัธยฐานตัวอย่างมากที่สุดสำหรับแต่ละขนาดตัวอย่าง

จากตารางที่ 3 เมื่อกำหนดให้ตัวอย่างถูกสุ่มมาจากประชากรที่มีการแจกแจงแบบปกติปลอมปน (Contaminated Normal Distribution) ซึ่งในที่นี้กำหนดการแจกแจงแบบ $CN(0,1,p,10,4)$ เมื่อ $p = 0.05$ พบว่า เมื่อใช้ตัวอย่างขนาด 6 ค่าเฉลี่ยของค่าประมาณความแปรปรวนของมัธยฐานตัวอย่างแต่ละวิธีไม่แตกต่างจากความแปรปรวนแบบ Empirical ของมัธย

ฐานตัวอย่างมากนัก โดยพบว่าวิธีการประมาณความแปรปรวนของมัธยฐานตัวอย่างด้วยวิธีบูตสแตรพ มีค่าประมาณความแปรปรวนของมัธยฐานตัวอย่างโดยเฉลี่ยใกล้เคียงกับความแปรปรวนแบบ Empirical ของมัธยฐานตัวอย่างมากที่สุด แต่ก็มีความแปรปรวนของค่าประมาณความแปรปรวนของมัธยฐานตัวอย่างสูงกว่าวิธีการประมาณความแปรปรวนของมัธยฐานตัวอย่างด้วยวิธีแจ็คไนฟ์ และ วิธีความแปรปรวนแบบมีเงื่อนไข

เมื่อเพิ่มขนาดของตัวอย่างที่สุ่มจากประชากรเป็น 10 และ 15 พบว่าความแปรปรวนของค่าประมาณความแปรปรวนของมัธยฐานตัวอย่างโดยวิธีความแปรปรวนแบบมีเงื่อนไข วิธีมีค่าใกล้เคียงกับความแปรปรวนแบบ Empirical ของมัธยฐานตัวอย่างมากขึ้น ในระดับที่ดีกว่าหรือใกล้เคียงกับวิธีการประมาณความแปรปรวนโดยวิธีของบูตสแตรพ นอกจากนี้ยังพบว่า ความแปรปรวนของค่าประมาณความแปรปรวนของมัธยฐานตัวอย่างด้วยวิธีความแปรปรวนแบบมีเงื่อนไข มีค่าไม่แตกต่างจากวิธีการประมาณความแปรปรวนโดยวิธีของบูตสแตรพมากนักในทุกกรณี

ซึ่งแสดงว่าในกรณีที่ข้อมูลมีค่าผิดปกติ วิธีการประมาณความแปรปรวนของมัธยฐานตัวอย่างด้วยวิธีความแปรปรวนแบบมีเงื่อนไขสามารถใช้ได้ดีเมื่อเปรียบเทียบกับวิธีการประมาณความแปรปรวนโดยวิธีบูตสแตรพ และดีกว่าวิธีการประมาณความแปรปรวนโดยวิธีแจ็คไนฟ์

ตารางที่ 4 ความแปรปรวนแบบ Empirical ของมัธยฐานตัวอย่างเทียบกับค่าประมาณความแปรปรวนของมัธยฐานด้วยวิธีความแปรปรวนแบบมีเงื่อนไข วิธีแจ็คไนฟ์ และวิธีบูตสแตรพ เมื่อกำหนดให้ประชากรที่ตัวอย่างสุ่มมามีการแจกแจงแบบปกติปนเปื้อนปกติปลอมปน (Contaminated Normal Distribution) ซึ่งในที่นี้กำหนดการแจกแจงแบบ $CN(0,1,p,10,4)$ เมื่อ $p = 0.10$ สำหรับตัวอย่างสุ่มขนาด 6 , 10 และ 15

<table>
<tr><th>ขนาดตัวอย่าง</th><th>ความแปรปรวนแบบ Empirical ของมัธยฐานตัวอย่าง (1)</th><th>วิธีการประมาณค่าความแปรปรวน</th><th>ค่าเฉลี่ยของค่าประมาณความแปรปรวนของมัธยฐานตัวอย่าง (2)</th><th>ผลต่างระหว่าง |(1)-(2)|</th><th>ความแปรปรวนของค่าประมาณความแปรปรวนของมัธยฐานตัวอย่าง</th></tr>
<tr><td rowspan="3">n = 6</td><td rowspan="3">0.6881325841</td><td>แจ็คไนฟ์</td><td>0.2711063325*</td><td>0.417026252*</td><td>2.6105844975</td></tr>
<tr><td>บูตสแตรพ</td><td>2.0847656727</td><td>1.396633089</td><td>10.2591609955</td></tr>
<tr><td>ความแปรปรวนแบบมีเงื่อนไข</td><td>2.1493597031</td><td>1.461227119</td><td>416.9306030273</td></tr>
<tr><td rowspan="3">n = 10</td><td rowspan="3">0.3909905553</td><td>แจ็คไนฟ์</td><td>0.0632538721</td><td>0.327736683</td><td>0.3280021548</td></tr>
<tr><td>บูตสแตรพ</td><td>0.8998481035</td><td>0.508857549</td><td>3.3161761761</td></tr>
<tr><td>ความแปรปรวนแบบมีเงื่อนไข</td><td>0.4516572058</td><td>0.060666651*</td><td>26.9912471771</td></tr>
<tr><td rowspan="3">n = 15</td><td rowspan="3">0.1291187406</td><td>แจ็คไนฟ์</td><td>0.0126189552</td><td>0.116499786</td><td>0.0002905747</td></tr>
<tr><td>บูตสแตรพ</td><td>0.3694368303</td><td>0.240318089</td><td>0.8638269305</td></tr>
<tr><td>ความแปรปรวนแบบมีเงื่อนไข</td><td>0.0959184170*</td><td>0.033200324*</td><td>0.0079547213</td></tr>
</table>

เครื่องหมาย * คือ ค่าเฉลี่ยของค่าประมาณความแปรปรวนของมัธยฐานตัวอย่างที่ใกล้เคียงกับความแปรปรวนแบบ Empirical ของมัธยฐานตัวอย่างมากที่สุดสำหรับแต่ละขนาดตัวอย่าง

จากตารางที่ 4 เมื่อกำหนดให้ตัวอย่างสุ่มมาจากประชากรที่มีการแจกแจงแบบปกติปลอมปน (Contaminated Normal Distribution) ซึ่งในที่นี้กำหนดการแจกแจงแบบ

$CN(0,1,p,10,4)$ เมื่อ $p = 0.10$ พบว่า เมื่อใช้ตัวอย่างขนาด 6 ค่าเฉลี่ยของตัวประมาณความแปรปรวนของมัธยฐานตัวอย่างโดยวิธีแจ็คไนฟ์มีค่าใกล้เคียงกับความแปรปรวนแบบ Empirical ของมัธยฐานตัวอย่างมากที่สุด และมีความแปรปรวนของค่าประมาณความแปรปรวนของมัธยฐานตัวอย่างต่ำที่สุดด้วยเช่นกัน

เมื่อเพิ่มขนาดของตัวอย่างที่สุ่มจากประชากรเป็น 10 และ 15 พบว่า วิธีการประมาณความแปรปรวนของค่ามัธยฐานด้วยวิธีความแปรปรวนแบบมีเงื่อนไข มีค่าประมาณความแปรปรวนของมัธยฐานตัวอย่างโดยเฉลี่ยใกล้เคียงกับความแปรปรวนแบบ Empirical ของมัธยฐานตัวอย่างมากที่สุด แต่ก็มีความแปรปรวนของค่าประมาณความแปรปรวนของมัธยฐานตัวอย่างสูงกว่าวิธีการประมาณความแปรปรวนของมัธยฐานตัวอย่างด้วยวิธีบูตสแตรพและวิธีแจ็คไนฟ์

ซึ่งแสดงว่า เมื่อมีค่าผิดปกติเพิ่มมากขึ้น วิธีความแปรปรวนแบบมีเงื่อนไขสามารถใช้ได้ดีเมื่อเปรียบเทียบกับวิธีการประมาณความแปรปรวนโดยวิธีแจ็คไนฟ์ และดีกว่าวิธีการประมาณความแปรปรวนโดยวิธีบูตสแตรพ

บทที่ 5
สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

ผลการวิจัยในเรื่อง วิธีประมาณความแปรปรวนของตัวประมาณด้วยความแปรปรวนแบบมีเงื่อนไข : กรณีศึกษาความแปรปรวนของมัธยฐาน มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพในการประมาณความแปรปรวนของมัธยฐานตัวอย่าง ด้วยวิธีความแปรปรวนแบบมีเงื่อนไขที่เสนอใหม่ เทียบกับวิธีการประมาณความแปรปรวนที่มีใช้กันอยู่คือวิธีแจ็คไนฟ์ และวิธีบูตสแตรพ โดยการศึกษาใช้วิธีจำลองแบบ ด้วยการกำหนดให้ตัวอย่างถูกสุ่มจากประชากรที่มีการแจกแจงแบบปกติมาตรฐาน การแจกแจงแบบเอกรูป และการแจกแจงแบบปกติปลอมปนในรูป CN(0 , 1 , p , 10 , 4) เมื่อ $p = 0.05$ และ $p = 0.10$ ตามลำดับ จากผลการวิจัยสามารถสรุปผลการวิจัยได้ดังนี้

1. ประสิทธิภาพของการประมาณความแปรปรวนของค่ามัธยฐานตัวอย่างด้วยวิธีความแปรปรวนแบบมีเงื่อนไข วิธีแจ็คไนฟ์ และวิธีบูตสแตรพ สำหรับกรณีที่ประชากรมีการแจกแจงแบบปกติมาตรฐาน สรุปได้ว่า

เมื่อพิจารณาแต่ละขนาดตัวอย่าง สำหรับตัวอย่างสุ่มขนาด 6 , 10 และ 15 วิธีการประมาณความแปรปรวนของค่ามัธยฐานตัวอย่าง 3 วิธี คือ วิธีความแปรปรวนแบบมีเงื่อนไข วิธีแจ็คไนฟ์ และวิธีบูตสแตรพ ส่วนใหญ่มีค่าเฉลี่ยของตัวประมาณความแปรปรวนแต่ละวิธีไม่แตกต่างกันมากนัก โดยพบว่าเมื่อใช้ขนาดตัวอย่าง 6 และ 15 วิธีบูตสแตรพให้ค่าประมาณความแปรปรวนของค่ามัธยฐานตัวอย่างใกล้เคียงกับความแปรปรวนแบบ Empirical ของมัธยฐานตัวอย่างมากที่สุดรองลงมาคือวิธีความแปรปรวนแบบมีเงื่อนไข และเมื่อใช้ขนาดตัวอย่างเท่ากับ 10 วิธีความแปรปรวนแบบมีเงื่อนไข ให้ค่าประมาณความแปรปรวนของค่ามัธยฐานตัวอย่างใกล้เคียงกับความแปรปรวนแบบ Empirical ของมัธยฐานตัวอย่างมากที่สุด แต่ก็มีความแปรปรวนของค่าประมาณความแปรปรวนของค่ามัธฐานตัวอย่างสูงกว่าวิธีแจ็คไนฟ์ และวิธีบูตสแตรพ ซึ่งแสดงว่าวิธีบูตสแตรพสามารถใช้ได้ดี เมื่อเทียบกับวิธีความแปรปรวนแบบมีเงื่อนไข และวิธีแจ็คไนฟ์ โดยที่วิธีความแปรปรวนแบบมีเงื่อนไข สามารถใช้ในการประมาณความแปรปรวนของค่ามัธยฐานตัวอย่างได้ดีไม่แตกต่างจากวิธีบูตสแตรพมากนัก และดีกว่าวิธีแจ็คไนฟ์

2. ประสิทธิภาพของการประมาณความแปรปรวนของค่ามัธยฐานตัวอย่างด้วยวิธีความแปรปรวนแบบมีเงื่อนไข วิธีแจ็คไนฟ์ และวิธีบูตสแตรพ สำหรับกรณีที่ประชากรมีการแจกแจงแบบเอกรูป ซึ่งในที่นี้กำหนดการแจกแจงแบบเอกรูปที่ศึกษาให้มีค่าเฉลี่ย(μ) และความแปรปรวน (σ^2) ในช่วง (0 , 1) สรุปได้ว่า

เมื่อพิจารณาแต่ละขนาดตัวอย่าง สำหรับตัวอย่างสุ่มขนาด 6 , 10 และ 15 วิธีการประมาณความแปรปรวนของค่ามัธยฐานตัวอย่างทั้ง 3 วิธี คือ วิธีความแปรปรวนแบบมีเงื่อนไข วิธีแจ็คไนฟ์ และวิธีบูตสแตรพ ซึ่งวิธีบูตสแตรพ ให้ค่าประมาณความแปรปรวนของค่ามัธยฐานตัวอย่างใกล้เคียงกับความแปรปรวนแบบ Empirical ของค่ามัธยฐานตัวอย่างมากที่สุดสำหรับทุกขนาดตัวอย่าง แต่ก็มีความแปรปรวนของค่าประมาณความแปรปรวนของค่ามัธฐานตัวอย่างสูงกว่าวิธีแจ็คไนฟ์ สำหรับทุกขนาดตัวอย่างด้วยเช่นกัน แสดงว่า วิธีบูตสแตรพสามารถใช้ในการประมาณความแปรปรวนของค่ามัธยฐานตัวอย่างได้ดีที่สุด เมื่อเปรียบเทียบกับวิธีประมาณความแปรปรวนของค่าประมาณความแปรปรวนของมัธยฐานตัวอย่างโดยวิธีความแปรปรวนแบบมีเงื่อนไขและวิธีแจ็คไนฟ์ ในกรณีที่ข้อมูลมีการแจกแจงเอกรูป

3. ประสิทธิภาพของการประมาณความแปรปรวนของค่ามัธยฐานตัวอย่างด้วยวิธีความแปรปรวนแบบมีเงื่อนไข วิธีแจ็คไนฟ์ และวิธีบูตสแตรพ สำหรับกรณีที่ประชากรมีการแจกแจงแบบปกติปลอมปน CN(0 , 1 , p , 10 , 4) **เมื่อ** $p = 0.05$ สรุปได้ว่า

เมื่อพิจารณาแต่ละขนาดตัวอย่าง สำหรับตัวอย่างสุ่มขนาด 6 , 10 และ 15 วิธีการประมาณความแปรปรวนของค่ามัธยฐานตัวอย่าง 3 วิธี คือ วิธีความแปรปรวนแบบมีเงื่อนไข วิธีแจ็คไนฟ์ และวิธีบูตสแตรพ ส่วนใหญ่มีค่าเฉลี่ยของตัวประมาณความแปรปรวนของค่ามัธยฐานตัวอย่างในแต่ละวิธีใกล้เคียงกับความแปรปรวนแบบ Empirical ของมัธยฐานตัวอย่าง เมื่อใช้ขนาดตัวอย่างเท่ากับ 6 วิธีบูตสแตรพให้ค่าประมาณความแปรปรวนของค่ามัธยฐานตัวอย่างใกล้เคียงกับความแปรปรวนแบบ Empirical ของมัธยฐานตัวอย่างมากที่สุด แต่ก็มีความแปรปรวนของค่าประมาณความแปรปรวนของค่ามัธฐานตัวอย่างสูงกว่าวิธีแจ็คไนฟ์ และเมื่อเพิ่มขนาดตัวอย่างเป็น 10 และ 15 พบว่าวิธีความแปรปรวนแบบมีเงื่อนไขให้ค่าประมาณความแปรปรวนของมัธยฐานตัวอย่าโดยเฉลี่ยใกล้เคียงกับค่าความแปรปรวนแบบ Empirical ของมัธยฐานตัวอย่างมากที่สุด แต่ก็มีความแปรปรวนของค่าประมาณความแปรปรวนของค่ามัธฐานตัวอย่างสูงกว่าวิธีแจ็คไนฟ์ แสดงว่า วิธีความแปรปรวนแบบมีเงื่อนไข สามารถใช้ได้ดีเมื่อเปรียบเทียบกับวิธีบูตสแตรพและวิธีแจ็คไนฟ์ ในกรณีที่ข้อมูลมีการแจกแจงแบบปกติปลอมปน CN(0 , 1 , p , 10 , 4) เมื่อ $p = 0.05$

4. ประสิทธิภาพของการประมาณความแปรปรวนของค่ามัธยฐานตัวอย่างด้วยวิธีความแปรปรวนแบบมีเงื่อนไข วิธีแจ็คไนฟ์ และวิธีบูตสแตรพ สำหรับกรณีที่ประชากรมีการแจกแจงแบบปกติปนเปื้อน CN(0 , 1 , p , 10 , 4) **เมื่อ** $p = 0.10$ สรุปได้ว่า

เมื่อพิจารณาแต่ละขนาดตัวอย่าง สำหรับตัวอย่างสุ่มขนาด 6 , 10 และ 15 วิธีการประมาณความแปรปรวนของค่ามัธยฐานตัวอย่าง 3 วิธี คือ วิธีความแปรปรวนแบบมีเงื่อนไข วิธีแจ็คไนฟ์ และวิธีบูตสแตรพ โดยที่วิธีแจ็คไนฟ์ ให้ค่าเฉลี่ยของตัวประมาณความแปรปรวนของค่ามัธยฐานตัวอย่างใกล้เคียงกับความแปรปรวนแบบ Empirical ของมัธยฐานตัวอย่างมากที่สุด และมีความแปรปรวนของค่าประมาณความแปรปรวนของค่ามัธยฐานตัวอย่างต่ำที่ด้วยเช่นกัน สำหรับตัวอย่างสุ่มที่มีขนาดเท่ากับ 6 เมื่อเพิ่มขนาดตัวอย่างเป็น 10 และ 15 ค่าเฉลี่ยของตัวประมาณความแปรปรวนของค่ามัธยฐานตัวอย่างโดยวิธีความแปรปรวนแบบมีเงื่อนไข จะมีค่าเฉลี่ยใกล้เคียงกับความแปรปรวนแบบ Empirical ของมัธยฐานตัวอย่างมากที่สุด แต่ก็มีความแปรปรวนของค่าประมาณความแปรปรวนของค่ามัธฐานตัวอย่างสูงกว่าวิธีแจ็คไนฟ์ แสดงว่า วิธีความแปรปรวนแบบมีเงื่อนไข สามารถใช้ในการประมาณความแปรปรวนของค่ามัธยฐานตัวอย่างได้ดีไม่แตกต่างจากวิธีแจ็คไนฟ์ และดีกว่าวิธีบูตสแตรพ ในกรณีที่ข้อมูลมีการแจกแจงแบบปกติปนเปื้อน CN(0 , 1 , p , 10 , 4) เมื่อ $p = 0.10$

อภิปรายผล

จากผลการศึกษาครั้งนี้ พบว่า กรณีที่ข้อมูลมีการแจกแจงแบบปกติ วิธีการประมาณความแปรปรวนของค่ามัธยฐานตัวอย่างโดยวิธีความแปรปรวนแบบมีเงื่อนไข สามารถใช้ประมาณค่าความแปรปรวนของค่ามัธยฐานตัวอย่างได้ดีไม่แตกต่างกับวิธีบูตสแตรพ และดีกว่าวิธีแจ็คไนฟ์ กล่าวคือ วิธีประมาณความแปรปรวนของค่ามัธยฐานแต่ละวิธีมีค่าเฉลี่ยใกล้เคียงกัน แต่วิธีความแปรปรวนแบบมีเงื่อนไข จะให้ค่าประมาณใกล้เคียงค่าความแปรปรวนแบบ Empirical ของมัธยฐานตัวอย่างมากกว่าวิธีบูตสแตรพเพียงเล็กน้อย และกรณีที่ข้อมูลมีการแจกแจงแบบเอกรูป พบว่า วิธีประมาณความแปรปรวนของค่ามัธยฐานตัวอย่างโดยบูตสแตรพ สามารถใช้ประมาณความแปรปรวนของค่ามัธยฐานตัวอย่างได้ดีที่สุด เมื่อเปรียบเทียบกับวิธีความแปรปรวนแบบมีเงื่อนไขและวิธีแจ็คไนฟ์ เนื่องจากวิธีการประมาณความแปรปรวนของมัธยฐานตัวอย่างด้วยวิธีบูตสแตรพ มีค่าประมาณความแปรปรวนของมัธยฐานตัวอย่างโดยเฉลี่ยใกล้เคียงกับความแปรปรวนแบบ Empirical ของมัธยฐานตัวอย่างมากที่สุด สำหรับทุกขนาดตัวอย่าง ภายใต้ข้อมูลที่มีการแจกแจงแบบเอกรูป สำหรับการแจกแจงแบบปกติปนเปื้อนอีก 2 แบบ คือ การแจกแจงแบบปกติปลอมปน ซึ่งกำหนดการแจกแจงแบบ CN(0 , 1 , p , 10 , 4) เมื่อ $p = 0.05$ และการ

แจกแจงแบบปกติปลอมปน ซึ่งกำหนดการแจกแจงแบบ CN(0 , 1 , p , 10 , 4) เมื่อ $p = 0.10$ พบว่า วิธีความแปรปรวนแบบมีเงื่อนไข สามารถใช้ในการประมาณความแปรปรวนของค่ามัธยฐานตัวอย่างได้ดีสำหรับขนาดตัวอย่าง 10 และ 15 ซึ่งให้ค่าประมาณความแปรปรวนของค่ามัธยฐานตัวอย่างใกล้เคียงกับความแปรปรวนแบบ Empirical มากกว่าที่สุด เมื่อเปรียบเทียบกับวิธีบูตสแตรพและวิธีแจ็คไนฟ์ ยกเว้นตัวอย่างสุ่มขนาด 6 ที่วิธีบูตสแตรพ ให้ค่าเฉลี่ยของตัวประมาณความแปรปรวนของค่ามัธยฐานตัวอย่างใกล้เคียงกับความแปรปรวนแบบ Empirical ของมัธยฐานตัวอย่างมากที่สุด และมีความแปรปรวนของค่าประมาณความแปรปรวนของค่ามัธยฐานตัวอย่างต่ำที่สุดด้วยเช่นกัน สำหรับการแจกแจงแบบ CN(0 , 1 , p , 10 , 4) เมื่อ $p = 0.05$

ข้อเสนอแนะของงานวิจัย

ผู้วิจัยมีข้อเสนอแนะ สำหรับการประมาณความแปรปรวนของค่ามัธยฐานตัวอย่างด้วยวิธีความแปรปรวนแบบมีเงื่อนไข ดังนี้

1. จากผลการศึกษา พบว่า ในกรณีที่ข้อมูลมีการแจกแจงแบบปกติ การแจกแจงแบบปกติปนเปื้อน ซึ่งกำหนดการแจกแจงแบบปกติปลอมปน CN(0 , 1 , p , 10 , 4) เมื่อ $p = 0.05$ และการแจกแจงแบบปกติปลอมปน CN(0 , 1 , p , 10 , 4) เมื่อ $p = 0.10$ วิธีความแปรปรวนแบบมีเงื่อนไข วิธีแจ็คไนฟ์ และวิธีบูตสแตรพ มีประสิทธิภาพในการประมาณความแปรปรวนของค่ามัธยฐานตัวอย่างไม่แตกต่างกัน แต่ในกรณีที่ข้อมูลมีการแจกแจงแบบเอกรูป การประมาณความแปรปรวนของค่ามัธยฐานตัวอย่างด้วยวิธีบูตสแตรพจะมีประสิทธิภาพในการประมาณความแปรปรวนของค่ามัธยฐานตัวอย่างดีกว่าวิธีความแปรปรวนแบบมีเงื่อนไขเพียงเล็กน้อย และดีกว่าวิธีแจ็คไนฟ์ ดังนั้น จึงควรลองพิจารณาการแจกแจงแบบอื่นเช่น การแจกแจงแบบล็อกนอร์มอล

2. งานวิจัยนี้ผู้วิจัยได้นำ วิธีความแปรปรวนแบบมีเงื่อนไข ไปหาความแปรปรวนของตัวประมาณเพียงตัวเดียว คือ ค่ามัธยฐานตัวอย่างเมื่อตัวอย่างมีขนาดเล็ก จึงควรศึกษาต่อไปว่า การใช้วิธีความแปรปรวนแบบมีเงื่อนไข ไปหาความแปรปรวนของตัวประมาณแบบอื่นๆ เช่น ค่ามัธยฐานตัวอย่างเมื่อตัวอย่างมีขนาดใหญ่ จะให้ผลแตกต่างจากงานวิจัยนี้หรือไม่

บรรณานุกรม

ภาษาไทย

ปราณี นิลกรณ์. ทฤษฎีความน่าจะเป็น. นครปฐม : ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยศิลปากร, 2547.

__________. การจำลองแบบ SIMULATION. นครปฐม : ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยศิลปากร, 2531.

สุชาดา กีระนันทน์. ทฤษฎีและการสำรวจตัวอย่าง. กรุงเทพมหานคร : โรงพิมพ์จุฬาลงกรณ์มหาวิทยาลัย, 2538

ภาษาต่างประเทศ

Beran, R.J. “Jackknife Approximations to Bootstrap Estimates.”The Annals of Statistics 12(1984) : 101-118.

Bickel, P.J., and D.A. Freedman. “Some Asymptotic Theory for the Bootstrap.”The Annals of Statistics 9(1981) : 1196-1217.

Davison, A.C, D.V. Hinkley, and E. Schechtman. “Efficient Bootstrap Simulation.” Biometrika 73(1986) : 555-556.

Efron, B., and R. Tibshirani. “Bootstrap Methods for Standard Errors : Confidence Intervals, and Other Measures of Statistical Accuracy.” Statistical Science 1(1986) : 54-77.

__________. An Introduction to the Bootstrap. London : Chapman & Hall, 1993.

Efron, B. “Bootstrap Methods : Another Look at the Jackknife.” The Annals of Statistics 7(1979) : 1-26.

__________. “Nonparametric Estimates of Standard Error : The Jackknife, the Bootstrap and Other Methods.” Biometrika 68(1981) : 589-599.

__________. The Jackknife, the Bootstrap and Other Resampling Plans. Philadelphia : SIAM , 1982

__________. “More Efficient Bootstrap Computations.” Journal of the American Statistical Association 85(1990) : 79-92.

Hinkley, D.V. “Improving the Jackknife with Special Reference to Correlation Estimation.” Biometrika 65(1978) : 13-22.

Miller, R.G. “The Jackknife – a Review.” Biometrika 61(1974) : 1-15.

Chernick, M. R. Bootstrap Methods : A Practitioner’s Guide. New York : Wiley & sons, 1999.

Parr, W.C. “A Notes on the Jackknife, the Bootstrap and the Delta Method Estimators of Bias and Variance.” Biometrika 70(1983) : 719-722.

__________. “Jackknifing Differentiable Statistical Functionals.” Journal of the Royal Statistical Society Series B 47(1985) : 56-66.

Quenonuille, M.H. “Problems in Plane Sampling.” The Annals of Mathematical Statistics 20(1949) : 136-137.

Ratanaprasert, P. Estimating the Variance of the Estimated Intercept in Distribution-free Straight Line Regression by the Conditional Variance and Other Methods. Melbourne : La Trobe University, 1987.

Shao, J., and C.F.J. Wu. “A General Theory for the Jackknife Variance Estimations.” The Annals of Statistics 17(1989) : 1176-1197.

Shao, J. “The Efficiency and Consistency of Approximations to the Jackknife Variance Estimators.” Journal of the American Statistical Association 84(1989) : 114-119.

ภาคผนวก

ขั้นตอนการทำงานของโปรแกรม

ในขั้นตอนการทำงานของโปรแกรมประกอบด้วย

1. กำหนดค่าคงที่และตัวแปร
2. กำหนดการแจกแจงของประชากร
3. ทำการเลือกตัวอย่างสุ่มมาจากประชากรในข้อ 2 โดยกำหนดการสุ่มตัวอย่างขนาด 6 , 10 และ 15 ตามลำดับ
4. ทำการเลือกตัวอย่างสุ่มซ้ำ ๆ ตัวอย่างสุ่ม ในข้อ 2 ด้วยวิธีการต่างๆ คือ วิธีแจ็คไนฟ์, วิธีบูตสแตรพ โดยทำการสุ่มตัวอย่างซ้ำเท่ากับ 500 ครั้ง และวิธีความแปรปรวนแบบมีเงื่อนไข
5. คำนวณค่าสถิติต่างๆ ดังนี้

   การประมาณค่ามัธยฐานของตัวอย่าง

   การประมาณค่าเฉลี่ยด้วยวิธีแจ็คไนฟ์

   การประมาณค่าความแปรปรวนของค่ามัธยฐานตัวอย่างด้วยวิธีแจ็คไนฟ์

   การประมาณค่าเฉลี่ยด้วยวิธีบูตสแตรพ

   การประมาณค่าความแปรปรวนของค่ามัธยฐานตัวอย่างด้วยวิธีบูตสแตรพ

   การประมาณค่าเฉลี่ยด้วยวิธีความแปรปรวนแบบมีเงื่อนไข

   การประมาณค่าความแปรปรวนของค่ามัธยฐานตัวอย่างด้วยวิธีความแปรปรวนแบบมีเงื่อนไข
6. เก็บค่าสะสมผลการคำนวณข้อ 5 ทั้งหมด 1,000 ครั้ง
7. คำนวณหาค่าต่าง ๆ สำหรับวิธีการประมาณค่าความแปรปรวนของค่ามัธยฐานตัวอย่าง ดังนี้

   ค่าเฉลี่ยของการประมาณค่าความแปรปรวนของค่ามัธยฐานตัวอย่างด้วยวิธีแจ็คไนฟ์

   ค่าประมาณความแปรปรวนของการประมาณค่าความแปรปรวนของค่ามัธยฐานตัวอย่างด้วยวิธีแจ็คไนฟ์

   ค่าเฉลี่ยของการประมาณค่าความแปรปรวนของค่ามัธยฐานตัวอย่างด้วยวิธีบูตสแตรพ

   ค่าประมาณความแปรปรวนของการประมาณค่าความแปรปรวนของค่ามัธยฐานตัวอย่างด้วยวิธีบูตสแตรพ

ค่าเฉลี่ยของการประมาณค่าความแปรปรวนของค่ามัธยฐานตัวอย่างด้วยวิธีความแปรปรวนแบบมีเงื่อนไข

ค่าประมาณความแปรปรวนของการประมาณค่าความแปรปรวนของค่ามัธยฐานตัวอย่างด้วยวิธีความแปรปรวนแบบมีเงื่อนไข

8. จบการทำงาน

ความหมายของตัวแปรที่ใช้ในโปรแกรม

ในโปรแกรมมีการกำหนดตัวแปร และความหมายดังต่อไปนี้

<u>ตัวแปรที่ใช้ในการสุ่มตัวอย่าง</u>

NR	ขนาดของประชากร
MAXLOOP	จำนวนครั้งของการทำซ้ำ 1000 ครั้ง
MAXB	จำนวนตัวอย่างบูตสแตรพ 500 ชุด
MVXMedS	ค่าเฉลี่ยของค่าประมาณความแปรปรวนของค่ามัธยฐานตัวอย่าง
VVXMedS	ค่าความแปรปรวนของค่าประมาณความแปรปรวนของค่ามัธยฐานตัวอย่าง

<u>ตัวแปรที่ใช้ในการหาค่าต่าง ๆ ด้วยวิธีแจ็คไนฟ์</u>

MVXMedJ	ค่าเฉลี่ยของค่าประมาณความแปรปรวนของค่ามัธยฐานตัวอย่างโดยวิธีแจ็คไนฟ์
VVXMedJ	ค่าความแปรปรวนของค่าประมาณความแปรปรวนของค่ามัธยฐานตัวอย่าง โดยวิธีแจ็คไนฟ์

<u>ตัวแปรที่ใช้ในการหาค่าต่าง ๆ ด้วยวิธีบูตสแตรพ</u>

MVXMedB	ค่าเฉลี่ยของค่าประมาณความแปรปรวนของค่ามัธยฐานตัวอย่างโดยวิธีบูตสแตรพ
VVXMedB	ค่าความแปรปรวนของค่าประมาณความแปรปรวนของค่ามัธยฐานตัวอย่าง โดยวิธีบูตสแตรพ

<u>ตัวแปรที่ใช้ในการหาค่าต่าง ๆ ด้วยวิธีความแปรปรวนแบบมีเงื่อนไข</u>

MVXMedC	ค่าเฉลี่ยของค่าประมาณความแปรปรวนของค่ามัธยฐานตัวอย่างโดยวิธีความแปรปรวนแบบมีเงื่อนไข
VVXMedC	ค่าความแปรปรวนของค่าประมาณความแปรปรวนของค่ามัธยฐานตัวอย่าง โดยวิธีความแปรปรวนแบบมีเงื่อนไข

โปรแกรมวิเคราะห์ประสิทธิภาพของการประมาณความแปรปรวนของมัธยฐานตัวอย่างด้วยวิธีแจ็คไนฟ์ , วิธีบูตสแตรพ และวิธีความแปรปรวนแบบมีเงื่อนไข

```
Use MSIMSL
        Integer, parameter :: MAXLOOP = 1000
        Integer, parameter :: MAXB = 500
        Real   MVXMedS
Real   VVXMedS
        Real   MVXMedJ
Real   VVXMedJ
Real   MVXMedB
Real   VVXMedB
        Real   MVXMedC
Real   VVXMedC
        Real, dimension(1:MAXLOOP) :: TablMed
        Real, dimension(1:MAXLOOP,1:15) :: XMed
        Real      alph
        Integer :: NR
        Open(1, FILE = 'Conclusion.txt')
        Open(2, FILE = 'DataAnal1.txt')
      Open(19, FILE = 'NoProcess.txt')
        NR = 6
        Call StoreRNUni(NR,XMed)
        CALL WriteUniXMed(NR,XMed)
        Call
ProcUni(NR,XMed,MVXMedS,VVXMedS,MVXMedJ,VVXMedJ,MVXMedB,VVXMedB,MV
XMedC,VVXMedC)
        NR = 10
        Call StoreRNUni(NR,XMed)
        CALL WriteUniXMed(NR,XMed)
```

```
Call
ProcUni(NR,XMed,MVXMedS,VVXMedS,MVXMedJ,VVXMedJ,MVXMedB,VVXMedB,MV
XMedC,VVXMedC)
NR = 15
Call StoreRNUni(NR,XMed)
CALL WriteUniXMed(NR,XMed)
Call
ProcUni(NR,XMed,MVXMedS,VVXMedS,MVXMedJ,VVXMedJ,MVXMedB,VVXMedB,MV
XMedC,VVXMedC)
close(2)
Open(2, FILE = 'DataAnal2.txt')
NR = 6
Call StoreRNorm(NR,XMed)
CALL WriteNormXMed(NR,XMed)
Call
ProcNorm(NR,XMed,MVXMedS,VVXMedS,MVXMedJ,VVXMedJ,MVXMedB,VVXMedB,M
VXMedC,VVXMedC)
NR = 10
Call StoreRNorm(NR,XMed)
CALL WriteNormXMed(NR,XMed)
Call
ProcNorm(NR,XMed,MVXMedS,VVXMedS,MVXMedJ,VVXMedJ,MVXMedB,VVXMedB,M
VXMedC,VVXMedC)
NR = 15
Call StoreRNorm(NR,XMed)
CALL WriteNormXMed(NR,XMed)
Call
ProcNorm(NR,XMed,MVXMedS,VVXMedS,MVXMedJ,VVXMedJ,MVXMedB,VVXMedB,M
VXMedC,VVXMedC)
close(2)
```

```
Open(2, FILE = 'DataAnal3.txt')
write(2,*) "CONTAMINATE 5 percent"
alph = 0.05
NR = 6
Call StoreRNCon(alph,NR,XMed)
CALL WriteContamXMed(NR,XMed)
Call
ProcContam5(NR,XMed,MVXMedS,VVXMedS,MVXMedJ,VVXMedJ,MVXMedB,VVXMedB,
MVXMedC,VVXMedC)
NR = 10
Call StoreRNCon(alph,NR,XMed)
CALL WriteContamXMed(NR,XMed)
Call
ProcContam5(NR,XMed,MVXMedS,VVXMedS,MVXMedJ,VVXMedJ,MVXMedB,VVXMedB,
MVXMedC,VVXMedC)
NR = 15
Call StoreRNCon(alph,NR,XMed)
CALL WriteContamXMed(NR,XMed)
Call
ProcContam5(NR,XMed,MVXMedS,VVXMedS,MVXMedJ,VVXMedJ,MVXMedB,VVXMedB,
MVXMedC,VVXMedC)
close(2)
Open(2, FILE = 'DataAnal4.txt')
write(2,*) "CONTAMINATE 10 percent"
alph = 0.1
NR = 6
Call StoreRNCon(alph,NR,XMed)
CALL WriteContamXMed(NR,XMed)
```

```
            Call
ProcContam10(NR,XMed,MVXMedS,VVXMedS,MVXMedJ,VVXMedJ,MVXMedB,VVXMed
B,MVXMedC,VVXMedC)
            NR = 10
            Call StoreRNCon(alph,NR,XMed)
            CALL WriteContamXMed(NR,XMed)
            Call
ProcContam10(NR,XMed,MVXMedS,VVXMedS,MVXMedJ,VVXMedJ,MVXMedB,VVXMed
B,MVXMedC,VVXMedC)
            NR = 15
            Call StoreRNCon(alph,NR,XMed)
            CALL WriteContamXMed(NR,XMed)
            Call
ProcContam10(NR,XMed,MVXMedS,VVXMedS,MVXMedJ,VVXMedJ,MVXMedB,VVXMed
B,MVXMedC,VVXMedC)
            close(1)
            close(2)
            contains
            Subroutine StoreRNUni(N,XMed)
                implicit None
                integer, intent(in) :: N
                Real, intent(out) :: XMed(1:MAXLOOP,1:15)
                Real,dimension(1:15) :: X
                Integer   i
                do i = 1, MAXLOOP
                    Call UniRn(N,X)
                    XMed(i,:) = X
                end do
            end subroutine
```

```
subroutine UniRn(N,X)
    implicit None
    Integer, intent(in) :: N
    Real, intent(out) :: X(1:15)
    integer   Iseed

    Call RNGET(Iseed)
    Call RNSET(Iseed)
    Call RNUN(N,X)
end subroutine

Subroutine   WriteUniXMed(NR,XMed)
    Implicit None
    integer, intent(in) :: NR
    Real, intent(in) :: XMed(1:MAXLOOP,1:15)
    Real, dimension(1:15) :: temp
    Integer i

    write(2,'("UNIFORM DIST. sample size =  ",I4)') NR
    do i = 1,MAXLOOP
        temp = XMed(i,:)
        CALL WriteDetail(NR,temp)
    end do
end subroutine

Subroutine WriteNormXMed(N,XMed)
    Implicit None
    integer, intent(in) :: N
    Real, intent(in) :: XMed(1:MAXLOOP,1:15)
    Real, dimension(1:15) :: temp
```

```
    Integer i

    write(2,'("NORMAL DIST. sample size =  ",I4)') NR
    do i = 1,MAXLOOP
        temp = XMed(i,:)
        CALL WriteDetail(NR,temp)
    end do
end subroutine

Subroutine WriteContamXMed(NR,XMed)
    Implicit None
    integer, intent(in) :: NR
    Real, intent(in) :: XMed(1:MAXLOOP,1:15)
    Real, dimension(1:15) :: temp
    Integer i
    write(2,'("CONTAMINATE DIST. sample size =  ",I4)') NR
    do i = 1,MAXLOOP
        temp = XMed(i,:)
        CALL WriteDetail(NR,temp)
    end do
end subroutine

Subroutine WriteDetail(NR,temp)
    Implicit None
    Integer, intent(in) :: NR
    Real, intent(in) :: temp(1:15)
        if (NR == 6) then
            CALL Detail6(temp)
        else if (NR == 10) then
            CALL Detail10(temp)
```

```
            else
            CALL Detail15(temp)
        end if
end subroutine

Subroutine   Detail6(temp)
    Implicit None
    Real, intent(in) :: temp(:)
    write(2,'(6F15.10)') temp(1:6)
end subroutine

Subroutine   Detail10(temp)
    Implicit None
    Real, intent(in) :: temp(:)
    write(2,'(10F15.10)') temp(1:10)
end subroutine

Subroutine   Detail15(temp)
    Implicit None
    Real, intent(in) :: temp(:)
    write(2,'(15F15.10)') temp(1:15)
end subroutine

Subroutine StoreRNorm(N,XMed)
    implicit None
    integer, intent(in) :: N
    Real, intent(out) :: XMed(1:MAXLOOP,1:15)
    Real,dimension(1:15) :: X
    Integer   i
    do i = 1, MAXLOOP
```

```
        Call NormRn(N,X)
        XMed(i,:) = X
    end do
end subroutine

subroutine NormRn(N,X)
    implicit None
    Integer, intent(in) :: N
    Real, intent(out) :: X(1:15)
    integer   Iseed
    Call RNGET(Iseed)
    Call RNSET(Iseed)
    Call RNNOA(N,X)
end subroutine

Subroutine StoreRNCon(lev,N,XMed)
    implicit None
    Real, intent(in) :: lev
    integer, intent(in) :: N
    Real, intent(out) :: XMed(1:MAXLOOP,1:15)
    Real,dimension(1:15) :: X
    Integer   i
    do i = 1, MAXLOOP
        Call NormContam(lev,N,X)
        XMed(i,:) = X
    end do
end subroutine

subroutine NormContam(lev,N,X)
    implicit None
```

```
Real, intent(in) :: lev
Integer, intent(in) :: N
Real, intent(out) :: X(1:15)
Real, dimension(1:15) :: temp
integer   i,Iseed
Call RNGET(Iseed)
Call RNSET(Iseed)
Call RNUN(N,temp)
Call RNNOA(N,X)
do i = 1,N
    if (temp(i)<= lev) then
        X(i) = 10+2*X(i)
    end if
end do
end subroutine
Subroutine
ProcUni(N,XMed,MVXMedS,VVXMedS,MVXMedJ,VVXMedJ,MVXMedB,VVXMedB,MVX
MedC,VVXMedC)
implicit None
integer, intent(in) :: N
Real, intent(out) :: MVXMedS
Real, intent(out) :: VVXMedS
Real, intent(out) :: MVXMedJ
Real, intent(out) :: VVXMedJ
Real, intent(out) :: MVXMedB
Real, intent(out) :: VVXMedB
Real, intent(out) :: MVXMedC
Real, intent(out) :: VVXMedC
Real, intent(in) :: XMed(1:MAXLOOP,1:15)
call ProcSimple(N,XMed,MVXMedS,VVXMedS)
```

```
call ProcJackknife(N,XMed,MVXMedJ,VVXMedJ)
call ProcBootTrap(N,XMed,MVXMedB,VVXMedB)
call ProcContam(N,XMed,MVXMedC,VVXMedC)
Write (19,'(" Uniform sample size : ",I4)') N
Write (19,'(" Number of samplings : ",I4)') MAXLOOP
Write (6,'(" Uniform sample size : ",I4)') N
Write (6,'(" Number of samplings : ",I4)') MAXLOOP
Write (1,'(" Uniform sample size : ",I4)') N
Write (1,'(" Number of samplings : ",I4)') MAXLOOP
Call
WriteResult(MVXMedS,VVXMedS,MVXMedJ,VVXMedJ,MVXMedB,VVXMedB,MVXMedC
,VVXMedC)
end subroutine

Subroutine
ProcNorm(N,XMed,MVXMedS,VVXMedS,MVXMedJ,VVXMedJ,MVXMedB,VVXMedB,MV
XMedC,VVXMedC)
implicit None
integer, intent(in) :: N
Real, intent(out) :: MVXMedS
Real, intent(out) :: VVXMedS
Real, intent(out) :: MVXMedJ
Real, intent(out) :: VVXMedJ
Real, intent(out) :: MVXMedB
Real, intent(out) :: VVXMedB
Real, intent(out) :: MVXMedC
Real, intent(out) :: VVXMedC
Real, intent(in) :: XMed(1:MAXLOOP,1:15)
call ProcSimple(N,XMed,MVXMedS,VVXMedS)
call ProcJackknife(N,XMed,MVXMedJ,VVXMedJ)
```

```
        call ProcBootTrap(N,XMed,MVXMedB,VVXMedB)
        call ProcContam(N,XMed,MVXMedC,VVXMedC)
    Write (19,'(" Normal sample size : ",I4)') N
    Write (19,'(" Number of samplings : ",I4)') MAXLOOP
    Write (6,'(" Normal sample size : ",I4)') N
    Write (6,'(" Number of samplings : ",I4)') MAXLOOP
    Write (1,'(" Normal sample size : ",I4)') N
    Write (1,'(" Number of samplings : ",I4)') MAXLOOP
        Call
WriteResult(MVXMedS,VVXMedS,MVXMedJ,VVXMedJ,MVXMedB,VVXMedB,MVXMedC,
VVXMedC)
    end subroutine

    Subroutine
ProcContam5(N,XMed,MVXMedS,VVXMedS,MVXMedJ,VVXMedJ,MVXMedB,VVXMedB,
MVXMedC,VVXMedC)
        implicit None
        integer, intent(in) :: N
        Real, intent(out) :: MVXMedS
        Real, intent(out) :: VVXMedS
        Real, intent(out) :: MVXMedJ
        Real, intent(out) :: VVXMedJ
        Real, intent(out) :: MVXMedB
        Real, intent(out) :: VVXMedB
        Real, intent(out) :: MVXMedC
        Real, intent(out) :: VVXMedC
        Real, intent(in) :: XMed(1:MAXLOOP,1:15)
        call ProcSimple(N,XMed,MVXMedS,VVXMedS)
        call ProcJackknife(N,XMed,MVXMedJ,VVXMedJ)
        call ProcBootTrap(N,XMed,MVXMedB,VVXMedB)
```

```
      call ProcContam(N,XMed,MVXMedC,VVXMedC)
   Write (19,'(" Contamination 5  sample size : ",I4)') N
   Write (19,'(" Number of samplings : ",I4)') MAXLOOP
   Write (6,'(" Contamination 5  sample size : ",I4)') N
   Write (6,'(" Number of samplings : ",I4)') MAXLOOP
   Write (1,'(" Contamination 5  sample size : ",I4)') N
   Write (1,'(" Number of samplings : ",I4)') MAXLOOP
      Call
WriteResult(MVXMedS,VVXMedS,MVXMedJ,VVXMedJ,MVXMedB,VVXMedB,MVXMedC,
VVXMedC)
   end subroutine

   Subroutine
ProcContam10(N,XMed,MVXMedS,VVXMedS,MVXMedJ,VVXMedJ,MVXMedB,VVXMedB,
MVXMedC,VVXMedC)
      implicit None
      integer, intent(in) :: N
      Real, intent(out) :: MVXMedS
      Real, intent(out) :: VVXMedS
      Real, intent(out) :: MVXMedJ
      Real, intent(out) :: VVXMedJ
      Real, intent(out) :: MVXMedB
      Real, intent(out) :: VVXMedB
      Real, intent(out) :: MVXMedC
      Real, intent(out) :: VVXMedC
      Real, intent(in) :: XMed(1:MAXLOOP,1:15)
      call ProcSimple(N,XMed,MVXMedS,VVXMedS)
      call ProcJackknife(N,XMed,MVXMedJ,VVXMedJ)
      call ProcBootTrap(N,XMed,MVXMedB,VVXMedB)
      call ProcContam(N,XMed,MVXMedC,VVXMedC)
```

```
Write (19,'(" Contamination 10  sample size : ",I4)') N
Write (19,'(" Number of samplings : ",I4)') MAXLOOP
Write (1,'(" Contamination 10  sample size : ",I4)') N
Write (1,'(" Number of samplings : ",I4)') MAXLOOP
Write (6,'(" Number of samplings : ",I4)') MAXLOOP
Write (6,'(" Contamination 10  sample size : ",I4)') N
    Call
WriteResult(MVXMedS,VVXMedS,MVXMedJ,VVXMedJ,MVXMedB,VVXMedB,MVXMedC,
VVXMedC)
end subroutine

Subroutine WriteResult(MVS,VVS,MVJ,VVJ,MVB,VVB,MVC,VVC)
    Implicit None
    Real, intent(in) :: MVS,VVS,MVJ,VVJ,MVB,VVB,MVC,VVC
    Write (6,'(" MeanVarXMed do Nothing : ",1F15.10)') MVS
    Write (19,'(" MeanVarXMed do Nothing : ",1F15.10)') MVS
    Write (6,'(" VarVarXMed do Nothing  : ",1F15.10)') VVS
    Write (19,'(" VarVarXMed do Nothing  : ",1F15.10)') VVS
    Write (6,'(" MeanVarXMed Jackknife : ",1F15.10)') MVJ
    Write (1,'(" MeanVarXMed Jackknife : ",1F15.10)') MVJ
    Write (6,'(" VarVarXMed Jackknife  : ",1F15.10)') VVJ
    Write (1,'(" VarVarXMed Jackknife  : ",1F15.10)') VVJ
    Write (6,'(" MeanVarXMed Bootstrap   : ",1F15.10)') MVB
    Write (1,'(" MeanVarXMed Bootstrap   : ",1F15.10)') MVB
    Write (6,'(" VarVarXMed Bootstrap    : ",1F15.10)') VVB
    Write (1,'(" VarVarXMed Bootstrap    : ",1F15.10)') VVB
        Write (6,'(" MeanVarXMed Con Variance   : ",1F15.10)') MVC
    Write (1,'(" MeanVarXMed Con Variance   : ",1F15.10)') MVC
    Write (6,'(" VarVarXMed Con Variance    : ",1F15.10)') VVC
    Write (1,'(" VarVarXMed Con Variance    : ",1F15.10)') VVC
```

```
end subroutine

Subroutine ProcSimple(N,XMed,MVXMedS,VVXMedS)
    implicit None
    integer, intent(in) :: N
    Real, intent(in) :: XMed(1:MAXLOOP,1:15)
    Real, intent(out) :: MVXMedS
    Real, intent(out) :: VVXMedS
    Real, dimension(1:15) :: TempS
    Real, dimension(1:MAXLOOP) :: MXMedS
    Real, dimension(1:MAXLOOP) :: VXMedS
    Integer  i
    do i = 1, MAXLOOP
        TempS = XMed(i,:)
        Call SortData(N,TempS)   ! เริ่มด้วยการเรียงลำดับ ข้อมูลที่สุ่มมา
        MXMedS(i) = FindMedian(N,TempS)   ! แล้วก็หา Median เก็บไว้ 1000 ตัว

    end do
    MVXMedS = FindMean(MAXLOOP,MXMedS)   ! หาค่าเฉลี่ยของ Median นั้น
    VVXMedS = FindVariance(MAXLOOP,MXMedS,MVXMedS)  ! หา Var ของ
Median นั้นๆ

end subroutine

Subroutine ProcJackknife(N,XMed,MVXMedJ,VVXMedJ)
    implicit None
    integer, intent(in) :: N
    Real, intent(out) :: MVXMedJ
    Real, intent(out) :: VVXMedJ
    Real, intent(in) :: XMed(1:MAXLOOP,1:15)
```

```
    Real, dimension(1:MAXLOOP,1:15) :: XMedJ
    Real, dimension(1:15) :: TempJ
    Real, dimension(1:MAXLOOP) :: MXMedJ
    Real, dimension(1:MAXLOOP) :: VXMedJ
    Integer  i
    do i = 1,MAXLOOP
        TempJ = XMed(i,:)
        Call DoJackknife(N,TempJ)
        XMedJ(i,:) = TempJ
    end do

    write(2,*) 'JackKnife  (Value in XMedJ )'
    do i = 1,MAXLOOP
        TempJ = XMedJ(i,:)
        Call WriteDetail(N,TempJ)
    end do
    write(2,*) 'JackKnife MEAN    VARIANCE (Value in MXMedJ and VXMedJ)'

    do i = 1, MAXLOOP
        TempJ = XMedJ(i,:)
        MXMedJ(i) = FindMean(N,TempJ)
        VXMedJ(i) = FindVariance(N,TempJ,MXMedJ(i))

        write(2,'(2F15.10)') MXMedJ(i),VXMedJ(i)
    end do
    MVXMedJ = FindMean(MAXLOOP,VXMedJ)
    VVXMedJ = FindVariance(MAXLOOP,VXMedJ,MVXMedJ)
end subroutine

Subroutine DoJackknife(N,TempJ)
```

```
    implicit None
    integer, intent(in) :: N
    Real, intent(out) :: TempJ(1:15)
    Real, dimension(1:15) :: temp1,temp2
    integer :: pos,i,j
    do pos = 1, N
    j = 1
        do i = 1, N
            if (i /= pos) then
             temp1(j) = TempJ(i)
             j = j+1
            end if
        end do
        Call SortData(N-1,temp1)
        temp2(pos) = FindMedian(N-1,temp1)
    end do
    TempJ = temp2
end subroutine

Subroutine ProcBootTrap(N,XMed,MVXMedB,VVXMedB)
    implicit None
    integer, intent(in) :: N
    Real, intent(out) :: MVXMedB
    Real, intent(out) :: VVXMedB
    Real, intent(in) :: XMed(1:MAXLOOP,1:15)
    Real, dimension(1:15) :: TempB
    Real, dimension(1:MAXB,1:15) :: XMedB
    Real, dimension(1:MAXLOOP) :: VXMedB
    Integer   i
    do i = 1, MAXLOOP
```

```
            TempB = XMed(i,:)
            Call doBootTrap(N,TempB,XMedB)    !construct XMedB
            VXMedB(i) = FindVarianceBoot(N,XMedB)
        end do
        MVXMedB = FindMean(MAXLOOP,VXMedB)
        VVXMedB = FindVariance(MAXLOOP,VXMedB,MVXMedB)
!---------- Write Data VXMedB
        write(2,*) 'Value in VXMedB of BOOTTRAP'
        do i = 1,MAXLOOP
            write(2,'(1F15.10)') VXMedB(i)
        end do
    end subroutine

    subroutine doBootTrap(N,TempB,XMedB)
        implicit None
        integer, intent(in) :: N
        Real, intent(in) :: TempB(1:15)
        Real, intent(out) :: XMedB(1:MAXB,1:15)
        Integer, dimension(1:N) :: temp1,temp2
        Integer   i,j,k,Iseed
        do i = 1, MAXB
            do j = 1,N
                Call RNGET(Iseed)
                Call RNSET(Iseed)
                Call RNPER(N,temp1)
                temp2(j) = temp1(j)
            end do
            do j = 1,N
                k = temp2(j)
                XMedB(i,j) = TempB(k)
```

```
        end do
    end do
end subroutine

Real Function FindVarianceBoot(N,XMedB)
    implicit None
    integer, intent(in) :: N
    Real, intent(in) :: XMedB(1:MAXB,1:15)
    Real, dimension(1:MAXB) :: MedTab
    Real, dimension(1:15) :: temp
    Real   MedBar
    Integer   i
    do i = 1, MAXB
        temp = XMedB(i,:)
        Call SortData(N,temp)
        MedTab(i) = FindMedian(N,temp)
    end do
    MedBar = FindMean(MAXB,MedTab)
    FindVarianceBoot = FindVariance(MAXB,MedTab,MedBar)
end function

Subroutine ProcContam(N,XMed,MVXMedC,VVXMedC)
    implicit None
    integer, intent(in) :: N
    Real, intent(out) :: MVXMedC
    Real, intent(out) :: VVXMedC
    Real, intent(in) :: XMed(1:MAXLOOP,1:15)
    Real, dimension(1:MAXLOOP,1:15) :: XMedC
    Real, dimension(1:15) :: TempC
    Real, dimension(1:MAXLOOP) :: MXMedC
```

```
    Real, dimension(1:MAXLOOP) :: VXMedC
    REal   MedVal
    Integer   i,j
    Write(2,*) 'DATA in DELTA of Condition'
    do i = 1, MAXLOOP
        TempC = XMed(i,:)
        Call SortData(N,TempC)
        MedVal = FindMedian(N,TempC)
        do j = 1,N
            TempC(j) = abs(TempC(j)-MedVal)
        end do
        CALL WriteDetail(N,TempC)
        Call SortData(N,TempC)
        VXMedC(i) = FindVarianceCon(N,TempC)
    end do
    MVXMedC = FindMean(MAXLOOP,VXMedC)
    VVXMedC = FindVariance(MAXLOOP,VXMedC,MVXMedC)
    Write(2,*) 'DATA in VXMedC Condition'
    do i = 1,MAXLOOP
        write(2,'(1F15.10)') VXMedC(i)
    end do
end subroutine

Real Function FindVarianceCon(N,Delta)
    implicit None
    integer, intent(in) :: N
    Real, intent(in) :: Delta(1:15)
    Real        temp
    integer   r
    r = mod(N,2)
```

```
        if (r == 0) then
            temp = EvenVariance(N,Delta)
        else
            temp = OddVariance(N,Delta)
        end if
        FindVarianceCon = temp
end Function

Real Function EvenVariance(N,Delta)
        implicit None
        integer, intent(in) :: N
        Real, intent(in) :: Delta(1:15)
        Integer  i,j,m
        Real       temp,SumX
        SumX= 0
        m = N/2
        do i = 2,m
            do j = 3,m+1
                SumX = SumX+2**i*nCr(N-j,m-j+1)*((Delta(i)+Delta(j))/4)**2
            end do
        end do
        temp = 0
        do j = 2,m+1
            temp = temp + nCr(N-j,m-j+1)*(Delta(1)**2+Delta(j)**2)**2
        end do
        EvenVariance = (sumX+temp)/(2**N)
end function

Real Function OddVariance(N,Delta)
        implicit None
```

```
    integer, intent(in) :: N
    Real, intent(in) :: Delta(1:15)
    integer   i,m
    Real        temp,SumX
    SumX= 0
    m = (N-1)/2
    do i = 2, m+1
        SumX = SumX+2**i*nCr(N-i,m)*Delta(i)**2
    end do
    temp = 2*nCr(N-1,m)*Delta(1)**2
    OddVariance = (sumX+temp)/(2**N)
end function

Real Function FindMean(N,X)
    implicit None
    integer, intent(in) :: N
    Real, intent(in) :: X(:)
    Real   SumX
    Integer   i
    SumX = 0
    do i = 1, N
        SumX = SumX+X(i)
    end do
    FindMean = SumX/N
end function


Real Function FindVarianceS(N,X,Xbar)
    implicit None
    integer, intent(in) :: N
```

```
    Real, intent(in) :: X(1:N)
    Real, intent(in) :: Xbar
    Real   SumX2
    Integer   i
    SumX2 = 0
    do i = 1, N
        sumX2 = SumX2+(X(i)-Xbar)**2
    end do
    FindVarianceS = SumX2/(N-1)
end function

Real Function FindVariance(N,X,Xbar)
    implicit None
    integer, intent(in) :: N
    Real, intent(in) :: X(1:N)
    Real, intent(in) :: Xbar
    Real   SumX2
    Integer   i
    SumX2 = 0
    do i = 1, N
        sumX2 = SumX2+(X(i)-Xbar)**2
    end do
    FindVariance = SumX2/N
end function

Real Function Factorial(N)
    Implicit None
    Integer, intent(in) :: N
    Integer   i
    Real   prod
```

```
    prod = 1.0
    if (N == 0) then
        prod =1.0
    else
        do i = 1, N
            prod = prod*i
        end do
    end if
    Factorial = prod
end function

Real Function nCr(N,R)
    Implicit None
    Integer, intent(in) :: N
    Integer, intent(in) :: R

    nCr = Factorial(N)/(Factorial(N-R)*Factorial(R))
end function

subroutine SortData(N,X)
    implicit None
    Integer, intent(in) :: N
    Real, intent(out) :: X(:)
    integer   i,j
    Real temp
    do i = 1, N
        do j = i+1, N
            if (X(i) > X(j)) then
             temp = X(i)
             X(i) = X(j)
```

```
                X(j) = temp
            end if
        end do
    end do
end subroutine



Real Function FindMedian(N,X)
    implicit None
    integer, intent(in) :: N
    Real, intent(in) :: X(:)
    Real :: medval
    integer :: r
    r = mod(N,2)
    if (r == 0) then
        medVal = (X(N/2)+X((N/2)+1))/2.0
    else
        medVal = X((N+1)/2)
    end if
    FindMedian = medVal
end function
END program
```

ประวัติผู้วิจัย

ชื่อ	นางสาวกัญญารัตน์ หัสมา
ที่อยู่	122 หมู่ 6 ถ. ละงู – ปากบารา ต. ละงู อ. ละงู จ. สตูล 91110 โทรศัพท์ 074-781751
ที่ทำงาน	ห้างหุ้นส่วนจำกัด พี.เค. สปีดโบ้ท ทราเวล ต. ละงู อ. ละงู จ. สตูล โทรศัพท์ 074-721621 , 082-4327315
ประวัติการศึกษา	
พ.ศ. 2547	สำเร็จการศึกษาบัณฑิต วิชาเอกคณิตศาสตร์ มหาวิทยาลัยทักษิณ วิทยาเขตสงขลา
พ.ศ. 2547	ศึกษาต่อระดับปริญญามหาบัณฑิต สาขาวิชาสถิติประยุกต์ บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร
ประวัติการทำงาน	
พ.ศ. 2547	อาจารย์(อัตราจ้าง) วิทยาลัยการอาชีพนวมินทราชูทิศ กรุงเทพมหานคร
พ.ศ. 2548 - 2549	อาจารย์ สถาบันไอ แอม จีเนียส สาขาวัชรพล กรุงเทพมหานคร
พ.ศ. 2550 – 2551	พนักงานสัมภาษณ์ บริษัทวิจัยการตลาด ไทเลอร์ แนลสัน ซอฟเฟรส บางรัก กรุงเทพมหานคร
พ.ศ. 2551 – ปัจจุบัน	ผู้จัดการ ห้างหุ้นส่วนจำกัด พี.เค. สปีดโบ้ท ทราเวล ต. ละงู อ. ละงู จ. สตูล