การประมาณพารามิเตอร์ของการแจกแจงแบบทวินาม โดยวิธีการประมาณแบบเบส์ภายใต้ฟังก์ชันสูญเสียคลาดเคลื่อนกำลังสอง

อาจารย์กาญจนา ใจบุญ

มหาวิทยาลัยราชภัฏเพชรบูรณ์ พ.ศ. ๒๕๕๕

บทคัดย่อ

หัวข้อวิจัย การประมาณพารามิเตอร์ของการแจกแจงแบบทวินามโดยวิธีการประมาณแบบเบส์ภายใต้ฟังก์ชันสูญเสียคลาดเคลื่อนกำลังสอง

ผู้วิจัย กาญจนา ใจบุญ

พ.ศ. 2555

การวิจัยครั้งนี้มีจุดมุ่งหมายเพื่อเปรียบเทียบค่าประมาณของพารามิเตอร์ของการแจกแจงแบบทวินามโดยวิธีการประมาณแบบเบส์ภายใต้ฟังก์ชันสูญเสียคลาดเคลื่อนกำลังสอง (Squared error loss function) $L[(\theta,T)]=(\theta-T)^2$ โดยที่ θ คือ ความน่าจะเป็นของการเกิดผลสำเร็จในการแจกแจงทวินาม T =T(X_1, X_2, ...,X_n) เป็นตัวประมาณของ θ โดยมีการแจกแจงเบื้องต้นของ θ เป็นการแจกแจงต่างๆ ได้แก่ ตามวิธีของ Hardd Jeffreys ตามวิธีของ Laplace prior และเป็นการแจกแจงแบบ Beta (a,b) และเพื่อหาข้อสรุปในการเลือกใช้ตัวประมาณพารามิเตอร์ของการแจกแจงแบบทวินามโดยวิธีการประมาณแบบเบส์ภายใต้ฟังก์ชันสูญเสียคลาดเคลื่อนกำลังสองโดยมีการแจกแจงเบื้องต้นต่างๆ กัน

สรุปผลการวิจัย พบว่า เมื่อใช้ขนาดของพารามิเตอร์ในการพิจารณาจะพบว่าตัวประมาณที่เหมาะสมสำหรับพารามิเตอร์ที่มีขนาดเล็กคือตัวประมาณแบบเบส์ ที่มีการแจกแจงเบื้องต้นเป็นแบบ Beta (1,3) และเมื่อขนาดพารามิเตอร์เพิ่มขึ้นพบว่า ตัวประมาณที่เหมาะสม คือตัวประมาณแบบเบส์ที่มีการแจกแจงเบื้องต้นเป็นแบบ Beta (2,3) ยกเว้นเมื่อตัวอย่างขนาดเล็ก ตัวประมาณแบบเบส์ที่มีการแจกแจงเบื้องต้นเป็นแบบ Beta (3,2) จะเป็นตัวประมาณที่เหมาะสม และจะสังเกตได้ว่าตัวประมาณค่าทุกตัวมีค่า MSE ลดลง เมื่อขนาดตัวอย่างมีขนาดเพิ่มขึ้น และถ้าพิจารณาค่าความเสี่ยงของเบส์เมื่อจำแนกตามขนาดตัวอย่าง จากตารางที่ 6. สามารถสรุปผลได้ว่า ตัวประมาณที่เหมาะสมสำหรับตัวอย่างขนาดเล็กคือตัวประมาณแบบเบส์ที่มีการแจกแจงเบื้องต้นเป็นแบบ Beta (1,3) (เมื่อ n=5) และ Beta (0.3,0.3) (เมื่อ n=25) สำหรับตัวอย่างขนาดกลาง และขนาดใหญ่ตัวประมาณแบบเบส์ ที่มีการแจกแจงเบื้องต้นเป็นแบบ Beta (0.3,0.3) จะเป็นตัวประมาณที่ดีที่สุด

คำสำคัญ : การประมาณพารามิเตอร์, การแจกแจงแบบทวินาม, การประมาณแบบเบส์ และฟังก์ชันสูญเสียคลาดเคลื่อนกำลังสอง

Abstract

Research Title	Eatimation parameters of Binomail distribution Squared error loss function Methods
The Researcher	Kanchana Jaiboon
Year	2012

This research's purpose was comparing estimates the Bayes estimation of binomial distribution using parameters in squared error loss function or $L[(\theta,T)]=(\theta-T)^2$, θ = Binomial distribution success probability, T = T(X_1, X_2, …,X_n) or the estimator of θ .According the prior distribution of θ to Hardd Jeffreys method ,Laplace method and Beta distribution (a,b).For selecting the estimation of binomial distribution using parameters in different distribution squared error loss function.

The research's results show a reasonable estimator for small parameters are the initial distribution Beta (1,3) and add more parameters are the distribution Beta (2,3) but the small sample size data reasonable estimator is Beta (3,2) of Bayes estimation. So all estimators, the MSE decreases with increasing data sample size. About the risk, Bayes classified in the sample size table 6 that reasonable estimators for small sample are Beta (1,3) (If n = 5) and Beta (0.3. , 0.3) (If n = 25). But the best estimators for medium and large sample are Beta (0.3,0.3) .

Keywords : Parameters' estimation, Binomial distribution , Bayes estimation and Squared error loss function.

กิตติกรรมประกาศ

งานวิจัยฉบับนี้ได้รับอนุเคราะห์ทุนอุดหนุนการวิจัยประเภททั่วไป ประจำปีงบประมาณ พ.ศ. ๒๕๕๕ จากมหาวิทยาลัยราชภัฏเพชรบูรณ์ ความสำเร็จดังกล่าวเกิดขึ้นได้ ก็ด้วยความกรุณาของคณาจารย์สาขาวิชาคณิตศาสตร์ คณะวิทยาศาสตร์และเทคโนโลยีทุกท่าน ผู้วิจัยต้องขอกราบขอบพระคุณคณาจารย์ทุกท่านเป็นอย่างสูงไว้ ณ ที่นี้

กราบขอบพระคุณกรรมการผู้ทรงคุณวุฒิภายนอก ที่กรุณาให้คำแนะนำและชี้แนะ ตลอดจนคณาจารย์ทุกท่านที่ประสิทธิ์ประสาทวิชา ให้คำแนะนำ ชี้แนะ อบรมสั่งสอน ให้กำลังใจในการจัดทำวิจัย

ขอขอบพระคุณ บุคลากร และเจ้าหน้าที่ทุกท่านที่ให้ความร่วมมือช่วยเหลือเป็นอย่างดี ทำให้วิจัยนี้สำเร็จสมบูรณ์ลงได้ ผู้บังคับบัญชาและเพื่อนร่วมงานทุกท่านที่ให้กำลังใจเสมอมา ท้ายสุดต้อง ขอขอบพระคุณผู้ที่อยู่เบื้องหลังความสำเร็จในครั้งนี้ คือ บิดา มารดา สามีและลูก ที่แสนดี ที่ให้ความช่วยเหลือ ให้กำลังใจ ทำให้วิจัยนี้สำเร็จลงได้

หากวิจัยฉบับนี้เกิดประโยชน์ต่อผู้เกี่ยวข้องทั้งทางตรงและทางอ้อม ผู้วิจัยขอมอบความดีในครั้งนี้ให้แก่มหาวิทยาลัยราชภัฏเพชรบูรณ์ มหาวิทยาลัยที่ทรงคุณค่าในการพัฒนาบุคลากร และท้องถิ่น

กาญจนา ใจบุญ

หัวหน้าโครงการวิจัย

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหาการวิจัย

ในการศึกษาเกี่ยวกับประชากรใดๆนั้นเป็นไปได้ยากที่จะศึกษาจากข้อมูลของประชากรทั้งหมด จึงใช้หลักทฤษฎีของสถิติอนุมาน ซึ่งเป็นศาสตร์ที่อาศัยทฤษฎีความน่าจะเป็นมาช่วยในการวิเคราะห์ โดยการนำข้อมูลที่รวบรวมจากตัวอย่างสุ่มไปอนุมานเพื่อนำไปสู่ข้อสรุปของประชากร การใช้ข้อมูลจากตัวอย่างสุ่มเพื่อประมาณพารามิเตอร์ให้มีความถูกต้องใกล้เคียงกับพารามิเตอร์นั้น ทั้งนี้ขึ้นอยู่กับผลของข้อมูลที่รวบรวมได้จากตัวอย่างที่สุ่มมาจากประชากรนั้นๆ การอนุมานสามารถทำได้สองลักษณะคือ การประมาณค่า (Estimation) และการทดสอบสมมติฐาน (Test of hypothesis) ซึ่งในที่นี้ผู้วิจัยพิจารณาเกี่ยวกับการประมาณค่าเท่านั้น โดยการประมาณค่าเป็นวิธีการวิเคราะห์ทางสถิติที่มีความสำคัญมาก และจะพบว่าในปัจจุบันนี้องค์กรต่างๆได้นำการประมาณค่ามาใช้กันอย่างแพร่หลาย

การประมาณค่าเป็นวิธีการใช้ค่าสถิติที่ได้จากตัวอย่างสุ่มไปประมาณค่าหรือสรุปผลเกี่ยวกับค่าพารามิเตอร์ที่สนใจว่ามีค่าเป็นเท่าไร โดยจะแบ่งออกเป็นสองแบบคือ การประมาณค่าแบบจุด (Point estimation) และ การประมาณค่าแบบช่วง (Interval estimation) ซึ่งการประมาณค่าแบบจุด คือ การประมาณค่าพารามิเตอร์ โดยประมาณเป็นค่าใดค่าหนึ่งเพียงค่าเดียว การประมาณค่าแบบช่วง คือ การประมาณค่าพารามิเตอร์ของประชากรว่าจะอยู่ในช่วงใดช่วงหนึ่งที่ระดับความเชื่อมั่นหนึ่ง

ในการประมาณค่า แนวคิดแบบดั้งเดิม (Classical approach) การประมาณค่าพารามิเตอร์ (θ) เริ่มจากการสุ่มตัวอย่างจากประชากรที่มีความหนาแน่น $f(x;\theta)$ และถือว่าพารามิเตอร์ (θ) เป็นค่าคงที่แต่ไม่ทราบค่า แต่ในแนวคิดแบบเบส์มีการนำเอาข้อมูลที่เกี่ยวกับ (θ) มาใช้ให้เกิดประโยชน์ต่อการประมาณค่า θ ให้ดียิ่งขึ้น การวิจัยครั้งนี้จะมุ่งความสนใจไปที่การประมาณค่าพารามิเตอร์ (θ) ซึ่งแทนความน่าจะเป็นของการเกิดผลสำเร็จของประชากรในการแจกแจงทวินาม โดยวิธีการประมาณแบบเบส์

1.2 วัตถุประสงค์ของการวิจัย

1. เพื่อเปรียบเทียบค่าประมาณของพารามิเตอร์ของการแจกแจงแบบทวินามโดยวิธีการประมาณแบบเบส์ภายใต้ฟังก์ชันสูญเสียคลาดเคลื่อนกำลังสอง (Squared error loss function) $L[(\theta, T)] = (\theta - T)^2$

โดยที่ θ คือ ความน่าจะเป็นของการเกิดผลสำเร็จในการแจกแจงทวินาม

T =T($X_1, X_2, \ldots, X_n$) เป็นตัวประมาณของ θ

โดยมีการแจกแจงเบื้องต้นของ θ เป็นการแจกแจงต่างๆดังนี้

1.1 ตามวิธีของ Hardd Jeffreys

1.2 ตามวิธีของ Laplace prior (เป็นการแจกแจงแบบ Uniform แบบหนึ่ง)

1.3 เป็นการแจกแจงแบบ Beta (a,b)

2. เพื่อหาข้อสรุปในการเลือกใช้ตัวประมาณพารามิเตอร์ของการแจกแจงแบบทวินามโดยวิธีการประมาณแบบเบส์ภายใต้ฟังก์ชันสูญเสียคลาดเคลื่อนกำลังสองโดยมีการแจกแจงเบื้องต้นต่าง ๆ กัน

1.3 ขอบเขตของการวิจัย

การศึกษาในงานวิจัยนี้ ทำการศึกษาภายใต้ขอบเขต ดังนี้

1.ข้อมูลที่ทำการศึกษาได้จากการจำลองข้อมูลให้มีการแจกแจงแบบทวินาม โดยใช้เทคนิคมอนติคาร์โลกระทำซ้ำ 500 ครั้งในแต่ละสถานการณ์ โดยกำหนดขนาดตัวอย่าง (n) ที่ใช้ในการศึกษา มีดังนี้

1.1 ตัวอย่างมีขนาดเล็ก คือ 5, 25

1.2 ตัวอย่างมีขนาดกลาง คือ 50, 80

1.3 ตัวอย่างมีขนาดใหญ่ คือ 100, 150

2. กำหนดค่าความน่าจะเป็นของการเกิดผลสำเร็จ(θ)ในการแจกแจงทวินามมีค่าเท่ากับ 0.1, 0.2, 0.3, 0.4, 0.5

3. การหาตัวประมาณจะเป็นตัวประมาณแบบวิธีของเบส์ภายใต้ฟังก์ชันสูญเสียคลาดเคลื่อนกำลังสอง

4. เกณฑ์ที่ใช้ทำการเปรียบเทียบในการประมาณมี 2 วิธี คือ

4.1 ตัวประมาณใดที่ให้ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean squared error: MSE) ต่ำสุดจะเป็นตัวประมาณที่เหมาะสมที่สุดสำหรับสถานการณ์นั้นๆ กล่าวคือ มีค่า $\sum_{i=1}^{500} (\hat{\theta}_i - \theta)^2 / 500$ ต่ำที่สุด โดยจะทำการเปรียบเทียบตามกลุ่มของแต่ละขนาดของตัวอย่าง

โดยที่ θ คือ ความน่าจะเป็นของการเกิดผลสำเร็จในการแจกแจงทวินามที่กำหนดไว้ข้างต้น

$\hat{\theta}$ คือ ค่าประมาณที่ได้จากตัวประมาณแบบเบส์ที่มีการแจกแจงเบื้องต้นตามที่กำหนดไว้ $(\hat{\theta}_{Bayes'Jeffrey}, \hat{\theta}_{Bayes'Laplace}, \hat{\theta}_{Bayes'Beta})$

4.2 ตัวประมาณใดที่ให้ค่าความเสี่ยงของเบส์ (Bayes' Risk) $r(T) = E_{\Theta}[R(\theta,T)]$ ต่ำสุดจะเป็นตัวประมาณที่เหมาะสมที่สุดสำหรับสถานการณ์นั้นๆ โดยจะทำการเปรียบเทียบตามกลุ่มของแต่ละขนาดของตัวอย่าง

1.4 สมมติฐานของการวิจัย

1. ตัวประมาณแบบเบส์ภายใต้ฟังก์ชันสูญเสียคลาดเคลื่อนกำลังสอง โดยมีการแจกแจงเบื้องต้นตามวิธีของ Hardd Jeffreys เป็นตัวประมาณ พารามิเตอร์ของการแจกแจงแบบทวินาม ที่ให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำกว่าวิธีอื่นๆที่พิจารณา

2. ตัวประมาณแบบเบส์ภายใต้ฟังก์ชันสูญเสียคลาดเคลื่อนกำลังสอง โดยมีการแจกแจงเบื้องต้นตามวิธีของ Hardd Jeffreys เป็นตัวประมาณ พารามิเตอร์ของการแจกแจงแบบทวินาม ที่ให้ค่าความเสี่ยงของเบส์ ต่ำกว่าวิธีอื่นๆที่พิจารณา

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1. ผลที่ได้จากการวิจัยจะเป็นแนวทางในการเลือกค่าประมาณของพารามิเตอร์ที่เหมาะสมในแต่ละสถานการณ์ สำหรับการแจกแจงแบบทวินาม

2. เป็นแนวทางในการศึกษาเปรียบเทียบวิธีการประมาณค่าแบบจุดสำหรับการแจกแจงแบบอื่นต่อไป

1.6 วิธีดำเนินการวิจัย

1. สร้างข้อมูลที่ใช้ในการวิจัยโดยใช้เทคนิคมอนติคาร์โล ด้วยโปรแกรมสำเร็จรูปทางสถิติ กระทำซ้ำ 500 ครั้งในแต่ละสถานการณ์ตามขนาดตัวอย่างและค่าความน่าจะเป็นของการเกิดผลสำเร็จในการแจกแจงทวินามที่กำหนดไว้

2. คำนวณค่าประมาณของพารามิเตอร์ในแต่ละสถานการณ์ตามขนาดตัวอย่างและค่าความน่าจะเป็นของการเกิดผลสำเร็จในการแจกแจงทวินามที่กำหนดไว้ โดยวิธีการประมาณแบบเบส์ภายใต้ฟังก์ชันสูญเสียคลาดเคลื่อนกำลังสอง ตามการแจกแจงเบื้องต้นที่กำหนด

3. คำนวณค่าความคลาดเคลื่อนกำลังสองเฉลี่ย และค่าความเสี่ยงของเบส์ ในแต่ละสถานการณ์ตามขนาดตัวอย่างและค่าความน่าจะเป็นของการเกิดผลสำเร็จในการแจกแจงทวินามโดยวิธีการประมาณแบบเบส์ภายใต้ฟังก์ชันสูญเสียคลาดเคลื่อนกำลังสอง ตามการแจกแจงเบื้องต้นที่กำหนด

4. ทำการเปรียบเทียบค่าความคลาดเคลื่อนกำลังสองเฉลี่ย และค่าความเสี่ยงของเบส์ ของตัวประมาณแบบเบส์ภายใต้ฟังก์ชันสูญเสียคลาดเคลื่อนกำลังสองตามการแจกแจงเบื้องต้นที่กำหนด

5. สรุปผลการวิจัยว่าวิธีใดเป็นวิธีที่ดีที่สุดในการประมาณค่าพารามิเตอร์ในสถานการณ์ต่าง ๆ ที่ได้กำหนดไว้

บทที่ 2

ทบทวนวรรณกรรม

2.1 ทฤษฎีและบทความที่เกี่ยวข้องกับการวิจัยในครั้งนี้มีดังต่อไปนี้

การแจกแจงแบบเบอร์นูลลี (Bernoulli Distribution)

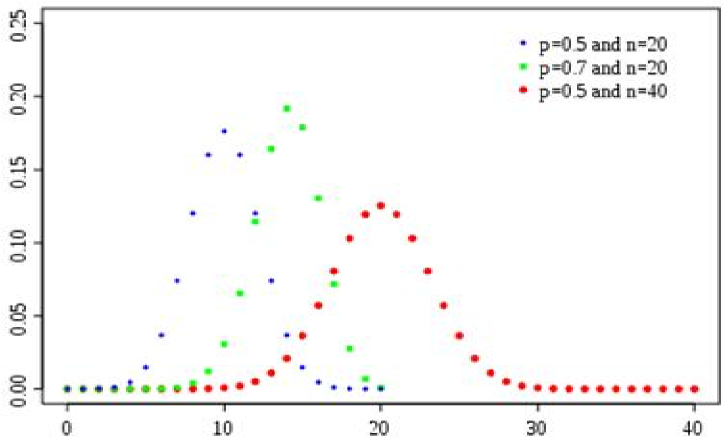
การแจกแจงแบบเบอร์นูลลี (Bernoulli Distribution) เป็นการแจกแจงที่อธิบายถึงสถานการณ์ของการทดลองใด ๆ ที่มีผลการทดลองเป็น 2 ลักษณะ คือ “ลักษณะที่สนใจ” หรือ “เกิดผลสำเร็จ” (Success) และ “ลักษณะที่ไม่สนใจ” หรือ “ไม่เกิดผลสำเร็จ” (Failure) โดยมีค่าความน่าจะเป็นของการ “เกิดผลสำเร็จ” เท่ากับ θ และความน่าจะเป็นของการ “ไม่เกิดผลสำเร็จ” เท่ากับ 1 - θ เมื่อทำการทดลองแบบเบอร์นูลลี(Bernoulli trial) จำนวน n ครั้ง โดยที่การทดลองในแต่ละครั้งเป็นอิสระต่อกันและมีค่าความน่าจะเป็นของการ “เกิดผลสำเร็จ” เท่ากับ θ คงที่ตลอดการทดลอง จะได้ว่าจำนวนครั้งของการ “เกิดผลสำเร็จ” ที่เกิดขึ้นจากการทดลองทั้ง n ครั้ง มีการแจกแจงแบบทวินาม (Binomial Distribution) ที่มีพารามิเตอร์ n และ θ นั่นคือ

ถ้ากำหนดให้ $X_1, X_2, \ldots, X_n$ เป็นตัวแปรสุ่มแต่ละค่าที่เป็นอิสระกัน และมีการแจกแจงแบบเบอร์นูลลี โดยมีพารามิเตอร์คือ θ จะได้ว่า $Y = \sum_{i=1}^{n} X_i$ เป็นตัวแปรสุ่มที่มีการแจกแจงแบบทวินาม ซึ่งรูปแบบการแจกแจงความน่าจะเป็นแบบทวินาม (Binomial probability distribution) สามารถแสดงได้ดังนี้

$$f(y;\theta) = P(Y = y)$$

$$= \begin{cases} \binom{n}{y}\theta^{y}\left(1-\theta\right)^{n-y} & ;\quad y = 0, 1, 2, \ldots, n \;;\; 0 \le \theta \le 1 \\ 0 & ;\quad \text{เมื่อ } y \text{ มีค่าอื่น ๆ} \end{cases}$$

และ $E(Y) = n\theta$, $V(Y) = n\theta(1-\theta)$

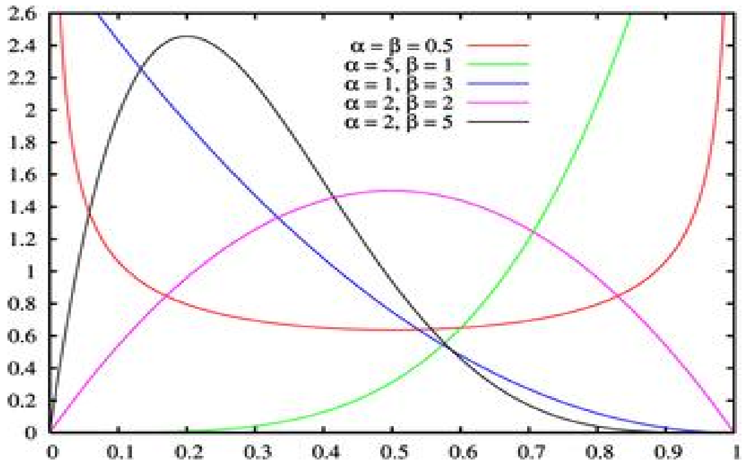


การแจกแจงแบบเบต้า (Beta Distribution)

ให้ตัวแปรสุ่ม X มีการแจกแจงแบบเบต้า ด้วยพารามิเตอร์ a และ b ฟังก์ชันความน่าจะเป็นหนาแน่นอยู่ในรูปของ

$$f(x;a,b) = \begin{cases} \dfrac{1}{B(a,b)} x^{a-1}(1-x)^{b-1} & ;\ x > 0 \text{ และ } a > 0,\ b > 0 \\ 0 & ;\ \text{เมื่อ } x \text{ มีค่าอื่น ๆ} \end{cases}$$

และ $E(X) = \dfrac{a}{a+b}$, $V(X) = \dfrac{ab}{(a+b)^2(a+b+1)}$



การหาตัวประมาณแบบเบส์โดยมีการแจกแจงเบื้องต้นต่าง ๆ ดังนี้

1. **การหาตัวประมาณค่าแบบเบส์โดยมีการแจกแจงเบื้องต้นตามวิธี Hardd Jeffreys** มีขั้นตอนดังนี้

1.1 หาฟังก์ชันภาวะน่าจะเป็น (Likelihood Function) : $L(\theta, y)$

1.2 หาอนุพันธ์อันดับหนึ่งของฟังก์ชันภาวะน่าจะเป็นดังกล่าวเทียบกับพารามิเตอร์ที่สนใจ

1.3 หาอนุพันธ์อันดับสองของฟังก์ชันภาวะน่าจะเป็นดังกล่าวเทียบกับพารามิเตอร์ที่สนใจ

1.4 หา $J^{\frac{1}{2}}(\theta)$ โดยที่

$$J^{\frac{1}{2}}(\theta) = \left[-\frac{1}{n}\frac{\partial^2}{\partial\theta^2}L(\theta|Y)\right]^{\frac{1}{2}}$$

จะได้ฟังก์ชันการแจกแจงเบื้องต้นของ θ คือ $p(\theta) \propto j^{\frac{1}{2}}(\theta)$

1.5 หาการแจกแจงโพสทีเรียจาก $h(\theta|y) = constant \bullet p(\theta)L(y|\theta)$

(http://en.wikipedia.org/wiki/Bayes'_theorem)

หาการแจกแจงเบื้องต้น (Prior Distribution)

ให้ $X_1, X_2, \ldots, X_n$ เป็นตัวแปรสุ่มจากประชากรที่มีการแจกแจงแบบเบอร์นูลลี $\text{Ber}(\theta)$ เมื่อ n แทนขนาดของตัวอย่าง และ θ แทนพารามิเตอร์ที่ไม่ทราบค่า โดยมีฟังก์ชันความน่าจะเป็นอยู่ในรูปของ

$$f(x;\theta) = \theta^x(1-\theta)^{1-x} \qquad ; x = 0, 1 \;,\; 0 \le \theta \le 1$$

ดังนั้นตัวสถิติ $Y = \sum_{i=1}^{n} X_i$ มีการแจกแจงแบบทวินาม Bin (n, θ)

1. ฟังก์ชันภาวะน่าจะเป็นของตัวอย่างสุ่มคือ

$$L(\theta|y) = \binom{n}{y}\theta^y(1-\theta)^{n-y} \qquad ; y = 0, 1, 2, \ldots, n \;,\; 0 \le \theta \le 1$$

$$\ln L(\theta|y) = \ln\binom{n}{y} + y\ln\theta + (n-y)(1-\theta)$$

2. $$\frac{\partial \ln L(\theta|y)}{\partial\theta} = \frac{y}{\theta} - \frac{(n-y)}{(1-\theta)}$$

3. $\dfrac{\partial^2 \ln L(\theta \mid y)}{\partial \theta^2} = -\dfrac{y}{\theta^2} - \dfrac{(n-y)}{(1-\theta)^2} \quad = \quad -\dfrac{n\bar{x}}{\theta^2} - \dfrac{n(1-\bar{x})}{(1-\theta)^2}$

4. หา $J^{\frac{1}{2}}(\theta) = \left[-\dfrac{1}{n}\dfrac{\partial^2}{\partial \theta^2} L(\theta \mid Y)\right]^{\frac{1}{2}} \quad = \quad \left[-\dfrac{n}{n}\left(-\dfrac{\bar{x}}{\theta^2} - \dfrac{(1-\bar{x})}{(1-\theta)^2}\right)\right]^{\frac{1}{2}}$

ประมาณค่าแบบภาวะน่าจะเป็นสูงสุด(Maximum likelihood estimation)

ของพารามิเตอร์ θ ได้จากการแก้สมการ $\dfrac{\partial \ln L(\theta \mid y)}{\partial \theta} = 0$ และได้ $\bar{x}$ เป็นตัวประมาณแบบภาวะน่าจะเป็นสูงสุด (Maximum likelihood estimator) ของพารามิเตอร์ θ

ดังนั้น $\quad J^{\frac{1}{2}}(\theta) = [\theta(1-\theta)]^{-\frac{1}{2}}$

5. $p(\theta) \propto J^{\frac{1}{2}}(\theta)$ จะได้ว่า $p(\theta) \propto [\theta(1-\theta)]^{-\frac{1}{2}}$

ขั้นตอนการหาการแจกแจงโพสทีเรีย (Posterior Distribution)

จาก $\quad h(\theta \mid y) = constant \bullet p(\theta) L(y \mid \theta)$

$$h(\theta \mid y) = c\binom{n}{y}\theta^{y-\frac{1}{2}}(1-\theta)^{n-y-\frac{1}{2}}$$

จะสามารถหาค่า c ได้โดยใช้คุณสมบัติความน่าจะเป็น คือ

$$\int_0^1 h(\theta \mid y)\, d\theta = 1$$

$$\int_0^1 c\binom{n}{y}\theta^{y-\frac{1}{2}}(1-\theta)^{n-y-\frac{1}{2}}\, d\theta = 1$$

$$\int_0^1 c\binom{n}{y}\theta^{(y+\frac{1}{2})-1}(1-\theta)^{(n-y+\frac{1}{2})-1}\, d\theta = 1$$

$$c\binom{n}{y}B\left(y+\frac{1}{2}, n-y+\frac{1}{2}\right)\int_0^1 \frac{\theta^{(y+\frac{1}{2})-1}(1-\theta)^{(n-y+\frac{1}{2})-1}}{B\left(y+\frac{1}{2}, n-y+\frac{1}{2}\right)} d\theta = 1$$

เนื่องจาก $$\int_0^1 \frac{\theta^{(y+\frac{1}{2})-1}(1-\theta)^{(n-y+\frac{1}{2})-1}}{B\left(y+\frac{1}{2},n-y+\frac{1}{2}\right)}d\theta = 1$$

จะได้ $$c\binom{n}{y}B\left(y+\frac{1}{2},n-y+\frac{1}{2}\right) = 1$$

$$c = \frac{1}{\binom{n}{y}B\left(y+\frac{1}{2},n-y+\frac{1}{2}\right)}$$

เมื่อ $B(a,b)$ แทนฟังก์ชันเบต้า (Beta function) = $\frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$

ดังนั้น $$h(\theta\mid y)=\frac{\theta^{(y+\frac{1}{2})-1}(1-\theta)^{(n-y+\frac{1}{2})-1}}{B\left(y+\frac{1}{2},n-y+\frac{1}{2}\right)}$$

นั่นคือ $$\theta\mid y \sim Beta\left(y+\frac{1}{2},n-y+\frac{1}{2}\right)$$

$E\left(\theta-\hat{\theta}\right)^2$ จะมีค่าต่ำสุดเมื่อ $\hat{\theta}=E(\theta\mid y)$ (Rao-Blackwell Theorem) ดังนั้น

$$\hat{\theta}_{Bayes'\,jeffrey}=E(\theta\mid y)=\int_0^1\theta h(\theta\mid y)d\theta$$

$$=\frac{1}{B\left(y+\frac{1}{2},n-y+\frac{1}{2}\right)}\int_0^1\theta^{\left(y+1\frac{1}{2}\right)-1}(1-\theta)^{\left(n-y+\frac{1}{2}\right)-1}d\theta$$

$$=\frac{B\left(y+1\frac{1}{2},n-y+\frac{1}{2}\right)}{B\left(y+\frac{1}{2},n-y+\frac{1}{2}\right)}=\frac{\Gamma\left(y+1\frac{1}{2}\right)\Gamma\left(n-y+\frac{1}{2}\right)}{\Gamma(n+2)}\bullet\frac{\Gamma(n+1)}{\Gamma\left(y+\frac{1}{2}\right)\Gamma\left(n-y+\frac{1}{2}\right)}$$

$$=\frac{y+1/2}{n+1}$$

จะได้ว่าตัวประมาณแบบเบส์ภายใต้ฟังก์ชันสูญเสียคลาดเคลื่อนกำลังสอง โดยมีการแจกแจงเบื้องต้นตามวิธีของ Hardd Jeffreys prior คือ

$$\hat{\theta}_{Bayes'Jeffrey} = \frac{\sum_{i=1}^{n} x_i + \frac{1}{2}}{n+1}$$

2. **การหาตัวประมาณค่าแบบเบส์โดยมีการแจกแจงเบื้องต้นตามวิธีของ Laplace prior** (เป็นการแจกแจงแบบ Uniform แบบหนึ่ง) มีขั้นตอนดังนี้

2.1 กำหนดการแจกแจงเบื้องต้นเป็น $p(\theta) = 1$

2.2 หาการแจกแจงโพสทีเรีย

จาก $h(\theta \mid y) = constant \bullet p(\theta) L(y \mid \theta)$

$$h(\theta \mid y) = c\left(\binom{n}{y}\theta^{y}(1-\theta)^{n-y}\right)(1)$$

$$h(\theta \mid y) = c\binom{n}{y}\theta^{y}(1-\theta)^{n-y}$$

จะสามารถหาค่า c ได้โดยใช้คุณสมบัติความน่าจะเป็น คือ

$$\int_0^1 h(\theta \mid y)\, d\theta = 1$$

$$\int_0^1 c\binom{n}{y}\theta^{y}(1-\theta)^{n-y}\, d\theta = 1$$

$$\int_0^1 c\binom{n}{y}\theta^{(y+1)-1}(1-\theta)^{(n-y+1)-1}\, d\theta = 1$$

$$c\binom{n}{y}B(y+1, n-y+1)\int_0^1 \frac{\theta^{(y+1)-1}(1-\theta)^{(n-y+1)-1}}{B(y+1, n-y+1)}\, d\theta = 1$$

เนื่องจาก $\displaystyle\int_0^1 \frac{\theta^{(y+1)-1}(1-\theta)^{(n-y+1)-1}}{B(y+1, n-y+1)}\, d\theta = 1$

จะได้ $$c\binom{n}{y}B(y+1,n-y+1) = 1$$

$$c = \frac{1}{\binom{n}{y}B(y+1,n-y+1)}$$

เมื่อ $B(a,b)$ แทนฟังก์ชันเบต้า (Beta function) = $\dfrac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$

ดังนั้น $$h(\theta \mid y) = \frac{\theta^{(y+1)-1}(1-\theta)^{(n-y+1)-1}}{B(y+1,n-y+1)}$$

นั่นคือ $$\theta \mid y \sim Beta(y+1,n-y+1)$$

$E\left(\theta-\hat{\theta}\right)^2$ จะมีค่าต่ำสุดเมื่อ $\hat{\theta} = E(\theta \mid y)$ ดังนั้น

$$\hat{\theta}_{Bayes'Laplace} = E(\theta \mid y) = \int_0^1 \theta h(\theta \mid y)\,d\theta$$

$$= \frac{1}{B(y+1,n-y+1)}\int_0^1 \theta^{(y+2)-1}(1-\theta)^{(n-y+1)-1}\,d\theta$$

$$= \frac{B(y+2,n-y+1)}{B(y+1,n-y+1)} = \frac{\Gamma(y+2)\Gamma(n-y+1)}{\Gamma(n+3)} \bullet \frac{\Gamma(n+2)}{\Gamma(y+1)\Gamma(n-y+1)}$$

$$= \frac{y+1}{n+2}$$

จะได้ว่าตัวประมาณแบบเบส์ภายใต้ฟังก์ชันสูญเสียคลาดเคลื่อนกำลังสอง โดยมีการแจกแจงเบื้องต้นตามวิธีของ Laplace prior คือ

$$\boxed{\hat{\theta}_{Bayes'Laplace} = \frac{\sum_{i=1}^{n} x_i + 1}{n+2}}$$

3. การหาตัวประมาณค่าแบบเบส์โดยมีการแจกแจงเบื้องต้นเป็นการแจกแจงแบบ Beta (a,b) มีขั้นตอนดังนี้

3.1 กำหนดการแจกแจงเบื้องต้นเป็น $p(\theta)=\dfrac{\theta^{a-1}(1-\theta)^{b-1}}{B(a,b)}$

3.2 หาการแจกแจงโพสทีเรีย

จาก $h(\theta \mid y)=constant \bullet p(\theta)L(y \mid \theta)$

$$h(\theta \mid y)=c\left(\binom{n}{y}\theta^{y}(1-\theta)^{n-y}\right)\left(\frac{\theta^{a-1}(1-\theta)^{b-1}}{B(a,b)}\right)$$

$$h(\theta \mid y)=\frac{c\binom{n}{y}\theta^{y+a-1}(1-\theta)^{n-y+b-1}}{B(a,b)}$$

จะสามารถหาค่า c ได้โดยใช้คุณสมบัติความน่าจะเป็น คือ

$$\int_0^1 h(\theta \mid y)\,d\theta = 1$$

$$\int_0^1 \frac{c\binom{n}{y}\theta^{y+a-1}(1-\theta)^{n-y+b-1}}{B(a,b)}\,d\theta = 1$$

$$\frac{c\binom{n}{y}B(y+a,n-y+b)}{B(a,b)}\int_0^1 \frac{\theta^{(y+a)-1}(1-\theta)^{(n-y+b)-1}}{B(y+a,n-y+b)}\,d\theta = 1$$

เนื่องจาก $\displaystyle\int_0^1 \frac{\theta^{(y+a)-1}(1-\theta)^{(n-y+b)-1}}{B(y+a,n-y+b)}\,d\theta = 1$

จะได้ $\displaystyle\frac{c\binom{n}{y}B(y+a,n-y+b)}{B(a,b)} = 1$

$$c = \frac{B(a,b)}{\binom{n}{y}B(y+a,n-y+b)}$$

เมื่อ $B(a,b)$ แทนฟังก์ชันเบต้า (Beta function) = $\dfrac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$

ดังนั้น $$h(\theta \mid y) = \frac{B(a,b)}{\binom{n}{y} B(y+a, n-y+b)} \bullet \frac{\binom{n}{y}\theta^{(y+a)-1}(1-\theta)^{(n-y+b)-1}}{B(a,b)}$$

$$= \frac{\theta^{(y+a)-1}(1-\theta)^{(n-y+b)-1}}{B(y+a, n-y+b)}$$

นั่นคือ $\theta \mid y \sim Beta(y+a, n-y+b)$

$E(\theta-\hat{\theta})^2$ จะมีค่าต่ำสุดเมื่อ $\hat{\theta} = E(\theta \mid y)$ ดังนั้น

$$\hat{\theta}_{Bayes'Beta} = E(\theta \mid y) = \int_0^1 \theta h(\theta \mid y) d\theta$$

$$= \frac{1}{B(y+a, n-y+b)} \int_0^1 \theta^{(y+a+1)-1}(1-\theta)^{(n-y+b)-1} d\theta$$

$$= \frac{B(y+a+1, n-y+b)}{B(y+a, n-y+b)} = \frac{\Gamma(y+a+1)\Gamma(n-y+b)}{\Gamma(n+a+b+1)} \bullet \frac{\Gamma(n+a+b)}{\Gamma(y+a)\Gamma(n-y+b)}$$

$$= \frac{y+a}{n+a+b}$$

จะได้ว่าตัวประมาณแบบเบส์ภายใต้ฟังก์ชันสูญเสียคลาดเคลื่อนกำลังสอง โดยมีการแจกแจงเบื้องต้นเป็นการแจกแจงแบบ Beta prior คือ

$$\boxed{\hat{\theta}_{Bayes'Beta} = \frac{\sum_{i=1}^{n} x_i + a}{n+a+b}}$$

ความเสี่ยง (Risk)

การพิจารณาความคลาดเคลื่อนที่เกิดขึ้นเมื่อใช้ $T = T(X_1, X_2, ..., X_n)$ เป็นตัวประมาณของ θ วิธีหนึ่งคือพิจารณาจาก ความสูญเสีย (Loss) ที่อาจเกิดขึ้นดังนี้

ฟังก์ชัน $L(\theta,T)$ เป็นฟังก์ชันสูญเสีย (Loss Function) ของการประมาณ θ ด้วย $T = T(X_1, X_2, ..., X_n)$ ก็ต่อเมื่อ $L(\theta,T)$ เป็นฟังก์ชันที่มีค่าไม่ติดลบ สำหรับทุกค่าของ θ และ $L(\theta,T)$ $= 0$ ก็ต่อเมื่อ $T = \theta$ ทุกค่าของ $\theta \in \Omega$

นั่นคือ $L(\theta,T) \geq 0$ และ $L(\theta,T) = 0 \ \ \forall \theta \in \Omega$

นอกจากนี้ยังกล่าวได้ว่า

ถ้า $L(\theta,T)$ มีค่ามาก แสดงว่า T แตกต่างจาก θ มาก

ถ้า $L(\theta,T)$ มีค่าน้อย แสดงว่า T แตกต่างจาก θ น้อย

ซึ่งวัดได้จากฟังก์ชันเสี่ยง (Risk Function) $R(\theta,T) = E[L(\theta,T)] = \int_{-\infty}^{\infty} L(\theta,T)h(\theta | T)d\theta$

ความเสี่ยงคือ ค่าคาดหวังของความสูญเสียเมื่อใช้ T ในการประมาณค่า θ (ในระยะยาว)

ความเสี่ยงของเบส์ (Bayes' Risk)

ให้ $X_1, X_2, ..., X_n$ เป็นตัวอย่างสุ่มขนาด n จากประชากรที่ฟังก์ชันความหนาแน่น $f(x | \theta)$ โดยที่ θ เป็นค่าของตัวแปรเชิงสุ่ม Θ

ความเสี่ยงของเบส์ (Bayes Risk) ผ่านฟังก์ชันสูญเสีย L และความหนาแน่นเบื้องต้น $p(\theta)$ แทนด้วย $r(T)$ ดังนี้

$$r(T) = E[R(\theta,T)] = E\left[E\{L(\theta,T)\}\right]$$

$r(T)$ มีค่าต่ำสุด ก็ต่อเมื่อ $\int_{-\infty}^{\infty} L(\theta,T)h(\theta | X)d\theta$ มีค่าต่ำที่สุด

เนื่องจาก $\int_{-\infty}^{\infty} L(\theta,T)h(\theta|X)d\theta = E[L(\theta,T)]$ ดังนั้น

เมื่อ $L(\theta,T)$ เป็น ฟังก์ชันสูญเสียคลาดเคลื่อนกำลังสอง

นั่นคือ $L[(\theta,T)] = (\theta - T)^2$

ดังนั้น ตัวประมาณแบบเบส์ของ θ ผ่านฟังก์ชันสูญเสียคลาดเคลื่อนกำลังสอง L คือ

T ที่ทำให้ $E[L(\theta,T)] = E[\theta - T]^2$ มีค่าต่ำสุด

ซึ่งเกิดขึ้นก็ต่อเมื่อ $T = E[\theta | X]$ (Rao-Blackwell Theorem)

นั่นคือ ถ้าฟังก์ชันความสูญเสียอยู่ในรูปของความคลาดเคลื่อนกำลังสอง

$$L(\theta,T) = [T - \theta]^2$$

ตัวประมาณแบบเบส์ (T) คือตัวประมาณภายหลังแบบเบส์ นั่นเอง กล่าวคือ

$$T = E[\theta | \mathrm{X}] = \int_{-\infty}^{\infty} \theta h(\theta | X) d\theta$$

โดยที่ T คือ ตัวสถิติที่ทำให้

$r(T) = \int_{-\infty}^{\infty} R(\theta,T)h(\theta)d\theta$ มีค่าต่ำสุด ก็ต่อเมื่อ $E(L) = \int_{-\infty}^{\infty} L(\theta,T)h(\theta)d\theta$ มีค่าต่ำสุด

ฟังก์ชันเสี่ยง (Risk Function) ของตัวประมาณแบบเบส์เมื่อมีการแจกแจงเบื้องต้นเป็นการแจกแจงแบบ Beta(a,b) ภายใต้ ฟังก์ชันสูญเสียคลาดเคลื่อนกำลังสอง

$$\begin{aligned} R(\theta,T) &= E[L(\theta,T)] \\ &= E[T-\theta]^2 \\ &= E[\hat{\theta}-\theta]^2 \end{aligned}$$

$$= E\left[\left(\frac{\sum_{i=1}^{n} x_i + a}{n+a+b}\right) - \theta\right]^2 \qquad ; \hat{\theta}_{Bayes'Beta} = \frac{\sum_{i=1}^{n} x_i + a}{n+a+b}$$

$$= E\left[\left(\frac{\overline{x} + \frac{a}{n}}{1 + \frac{a}{n} + \frac{b}{n}}\right) - \theta\right]^2$$

$$= \left(\frac{1}{1 + \frac{a+b}{n}}\right)^2 E\left[\overline{x} + \frac{a}{n} - \theta\left(1 + \frac{a}{n} + \frac{b}{n}\right)\right]^2$$

$$= \left(\frac{n}{n+a+b}\right)^2 E\left[\overline{x} + \frac{a}{n} - \left(\theta + \frac{a\theta}{n} + \frac{b\theta}{n}\right)\right]^2$$

$$= \left(\frac{n}{n+a+b}\right)^2 E\left[\left(\overline{x} - \theta\right) + \left(\frac{a}{n} - \frac{a\theta}{n} - \frac{b\theta}{n}\right)\right]^2$$

$$= \left(\frac{n}{n+a+b}\right)^2 E\left[\left(\overline{x} - \theta\right)^2 + 2\left(\overline{x} - \theta\right)\left(\frac{a - a\theta - b\theta}{n}\right) + \left(\frac{a - a\theta - b\theta}{n}\right)^2\right]$$

$$= \left(\frac{n}{n+a+b}\right)^2 \left[E\left(\overline{x} - \theta\right)^2 + 2\left(\frac{a - a\theta - b\theta}{n}\right)E\left(\overline{x} - \theta\right) + E\left(\frac{a - a\theta - b\theta}{n}\right)^2\right]$$

$$= \left(\frac{n}{n+a+b}\right)^2 \left[V\left(\overline{x}\right) + 0 + \left(\frac{a - a\theta - b\theta}{n}\right)^2\right]$$

$$= \left(\frac{n}{n+a+b}\right)^2 \left[\frac{1}{n^2} V\left(\sum_{i=1}^{n} x_i\right) + \left(\frac{a - a\theta - b\theta}{n}\right)^2\right]$$

$$= \left(\frac{n}{n+a+b}\right)^2 \left[\frac{n(\theta)(1-\theta) + \left(a - a\theta - b\theta\right)^2}{n^2}\right]$$

$$R(\theta, T) = \frac{n(\theta)(1-\theta) + \left(a - a\theta - b\theta\right)^2}{\left(n+a+b\right)^2}$$

Bayes' Risk: $\theta \sim Beta(a,b)$

$$r(T) = E_{\Theta}\left(R(\theta,T)\right)$$

$$= E_{\Theta}\left(\frac{n(\theta)(1-\theta)+(a-a\theta-b\theta)^2}{(n+a+b)^2}\right)$$

$$= \frac{1}{(n+a+b)^2}\left[nE\left(\theta-\theta^2\right)+E\left(a^2+a^2\theta^2+b^2\theta^2-2a^2\theta-2ab\theta+2ab\theta^2\right)\right]$$

$$= \frac{1}{(n+a+b)^2}\left[n\left[E(\theta)-E\left(\theta^2\right)\right]+a^2+a^2E\left(\theta^2\right)+b^2E\left(\theta^2\right)-2a^2E(\theta)-2abE(\theta)+2abE\left(\theta^2\right)\right]$$

$$= \frac{1}{(n+a+b)^2}\left[nE(\theta)-nE\left(\theta^2\right)+a^2-\left(2a^2+2ab\right)E(\theta)+\left(a^2+b^2+2ab\right)E\left(\theta^2\right)\right]$$

$$= \frac{1}{(n+a+b)^2}\left[a^2+\left(n-2a^2-2ab\right)E(\theta)+\left(a^2+b^2+2ab-n\right)E\left(\theta^2\right)\right]$$

$$= \frac{1}{(n+a+b)^2}\left[a^2+\left(n-2a^2-2ab\right)\left(\frac{a}{a+b}\right)+\left(a^2+b^2+2ab-n\right)\frac{a(a+1)}{(a+b)(a+b+1)}\right]$$

$$= \frac{1}{(n+a+b)^2}\left[\frac{a^2(a+b)(a+b+1)+\left(n-2a^2-2ab\right)(a)(a+b+1)+\left(a^2+b^2+2ab-n\right)a(a+1)}{(a+b)(a+b+1)}\right]$$

$$= \frac{1}{(n+a+b)^2}\left[\frac{a^2b+ab^2+nab}{(a+b)(a+b+1)}\right]$$

$$= \frac{1}{(n+a+b)^2}\left[\frac{ab(n+a+b)}{(a+b)(a+b+1)}\right]$$

$$= \frac{1}{(n+a+b)}\left[\frac{ab}{(a+b)(a+b+1)}\right]$$

$$= \frac{ab}{(n+a+b)(a+b)(a+b+1)}$$

นั่นคือ ความเสี่ยงของเบส์เมื่อตัวประมาณแบบเบส์มีการแจกแจงเบื้องต้นเป็นการแจกแจงแบบ Beta(a,b)

$$r(T)=\frac{ab}{(n+a+b)(a+b)(a+b+1)}$$

ฟังก์ชันเสี่ยงของตัวประมาณแบบเบส์เมื่อมีการแจกแจงเบื้องต้นตามวิธีของ Hardd Jeffreys

จาก $\hat{\theta}_{Bayes'Beta} = \dfrac{\sum_{i=1}^{n} x_i + a}{n+a+b}$

เมื่อ a = $1/2$ และ b = $1/2$

จะได้ $\hat{\theta}_{Bayes'Beta} = \dfrac{\sum_{i=1}^{n} x_i + \frac{1}{2}}{n+1}$ ซึ่งก็คือ $\hat{\theta}_{Bayes'J}$

จากฟังก์ชันเสี่ยงของตัวประมาณแบบเบส์เมื่อมีการแจกแจงเบื้องต้นเป็นการแจกแจงแบบ Beta (a,b) เมื่อ a = $1/2$ และ b = $1/2$ จะได้เป็นฟังก์ชันเสี่ยงของตัวประมาณแบบเบส์เมื่อมีการแจกแจงเบื้องต้นตามวิธีของ Hardd Jeffreys คือ

$$R(\theta,T)=\frac{n(\theta)(1-\theta)+\left(\frac{1}{2}-\theta\right)^2}{(n+1)^2}$$

และจะได้ว่าความเสี่ยงของเบส์เมื่อตัวประมาณแบบเบส์มีการแจกแจงเบื้องต้นตามวิธีของ Hardd Jeffreys คือ

$$r(T)=\frac{\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)}{\left(n+\frac{1}{2}+\frac{1}{2}\right)\left(\frac{1}{2}+\frac{1}{2}\right)\left(\frac{1}{2}+\frac{1}{2}+1\right)}$$

$$=\frac{1}{8(n+1)}$$

นั่นคือ ความเสี่ยงของเบส์เมื่อตัวประมาณแบบเบส์มีการแจกแจงเบื้องต้นตามวิธีของ Hardd Jeffreys

$$\boxed{r(T)=\frac{1}{8(n+1)}}$$

ฟังก์ชันเสี่ยงของตัวประมาณแบบเบส์เมื่อมีการแจกแจงเบื้องต้นตามวิธีของ Laplace

จาก $\hat{\theta}_{Bayes'Beta} = \dfrac{\sum_{i=1}^{n} x_i + a}{n+a+b}$

เมื่อ a = 1 และ b = 1

จะได้ $\hat{\theta}_{Bayes'Beta} = \dfrac{\sum_{i=1}^{n} x_i + 1}{n+2}$ ซึ่งก็คือ $\hat{\theta}_{Bayes'L}$

จากฟังก์ชันเสี่ยงของตัวประมาณแบบเบส์เมื่อมีการแจกแจงเบื้องต้นเป็นการแจกแจงแบบ Beta (a,b) เมื่อ a = 1 และ b = 1 จะได้เป็นฟังก์ชันเสี่ยงของตัวประมาณแบบเบส์เมื่อมีการแจกแจงเบื้องต้นตามวิธีของ Laplace คือ

$$R(\theta,T)=\frac{n(\theta)(1-\theta)+(1-2\theta)^2}{(n+2)^2}$$

และจะได้ว่าความเสี่ยงของเบส์เมื่อตัวประมาณแบบเบส์มีการแจกแจงเบื้องต้นตามวิธีของ Laplace คือ

$$r(T)=\frac{(1)(1)}{(n+1+1)(1+1)(1+1+1)}$$

$$=\frac{1}{6(n+2)}$$

นั่นคือ ความเสี่ยงของเบส์เมื่อตัวประมาณแบบเบส์มีการแจกแจงเบื้องต้นตามวิธีของ Laplace

$$\boxed{r(T)=\frac{1}{6(n+2)}}$$

2.2 งานวิจัยที่เกี่ยวข้อง

สำหรับการวิจัยครั้งนี้ ผู้วิจัยสนใจศึกษาการประมาณค่าพารามิเตอร์ของการแจกแจงแบบทวินาม โดยการศึกษาเกี่ยวกับการประมาณพารามิเตอร์มีผู้วิจัยหลายท่านที่ได้ทำการศึกษาในเรื่องนี้ เช่น

สาริณีย์ (2546) ได้เปรียบเทียบวิธีประมาณค่าแบบช่วงสำหรับพารามิเตอร์ในการแจกแจงแบบทวินาม โดยได้ข้อสรุปดังนี้

1. ตัวอย่างมีขนาดเล็ก ($n < 50$)
- วิธีสคอร์ เป็นวิธีที่เหมาะสมในการประมาณค่าแบบช่วงสำหรับพารามิเตอร์ในการแจกแจงทวินามได้ในทุกพารามิเตอร์ p
2. ตัวอย่างมีขนาดปานกลาง ($50 \leq n \leq 100$)
- ถ้าค่า $p \leq 0.07$ วิธีที่เหมาะสม คือ วิธีสคอร์
- ค่า $0.09 \leq p \leq 0.15$ วิธีที่เหมาะสม คือ วิธีแปลงแบบอาร์คไซน์ วิธีสคอร์
- ถ้าค่า $0.20 \leq p \leq 0.50$ วิธีที่เหมาะสม คือ วิธีสคอร์ วิธีแปลงแบบอาร์คไซน์ และ วิธีเพิ่มจำนวนลักษณะที่สนใจและจำนวนลักษณะที่ไม่สนใจอีก 2 ค่า
3. ตัวอย่างมีขนาดใหญ่ ($n > 100$)
- ถ้าค่า $0.05 \leq p \leq 0.15$ วิธีที่เหมาะสม คือ วิธีแปลงแบบอาร์คไซน์ วิธีสคอร์
- ถ้าค่า $0.20 \leq p \leq 0.50$ วิธีที่เหมาะสม คือ วิธีสคอร์ วิธีแปลงแบบอาร์คไซน์ และ วิธีเพิ่มจำนวนลักษณะที่สนใจและจำนวนลักษณะที่ไม่สนใจอีก 2 ค่า

กฤตยา (2549) ได้ศึกษาเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ของการแจกแจงปัวส์ซอง แบบช่วง 2 วิธี คือ วิธีการประมาณแบบแมกซิมัมไลลิฮูด และวิธีการประมาณแบบเบส์ ที่มีการแจกแจงก่อนเป็นแบบแกมมา โดยผลการวิจัยพบว่าวิธีการประมาณช่วงความเชื่อมั่นแบบเบส์เมื่อการแจกแจงก่อนเป็นแบบแกมมาให้ผลดีกว่าแบบแมกซิมัมไลลิฮูดในทุกกรณีที่ศึกษา

วราฤทธิ์ (2550) ได้ศึกษาเปรียบเทียบวิธีการประมาณค่าช่วงความเชื่อมั่นของพารามิเตอร์ของการแจกแจงแบบเลขชี้กำลัง เมื่อข้อมูลมีค่าผิดปกติ ด้วยวิธีการประมาณ 3 วิธี คือ วิธีการประมาณแบบปกติ (Normal method) วิธีการประมาณด้วยรากของสมการกำลังสอง (Root of Quadratic Equation method) และวิธีการประมาณด้วยช่วงแบบเบส์ (Bayesian Interval method)ผลการวิจัยสรุปได้ดังนี้ วิธีการประมาณแบบปกติ (วิธี N) เป็นวิธีที่ให้ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นต่ำกว่าวิธีอื่น เนื่องจากวิธีนี้ให้ค่าความกว้างเฉลี่ยของช่วง

ความเชื่อมั่นต่ำ และให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นต่ำที่สุด เมื่อค่าพารามิเตอร์ Ω เท่ากับ 1 ในทุกระดับของขนาดตัวอย่าง

วราฤทธิ์ (2550) ได้ศึกษาการหาค่าพารามิเตอร์ α และ β ที่ทำให้ค่าเฉลี่ยแบบเบส์ของปัวส์ซงมีค่า MSE ต่ำที่สุด สรุปได้ว่าค่า α และ β ที่เหมาะสม คือค่า α และ β ที่ทำให้ $\alpha\beta$ มีค่าใกล้เคียงกับค่าพารามิเตอร์ และค่า α มากกว่า β เสมอ และโดยทั่วไปแล้วค่า α อยู่ระหว่าง [4.0,5.0]

มัลลิกา (2551) ได้ศึกษาเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์สำหรับการแจกแจงทวินาม ผลการวิจัยสรุปได้ดังนี้

1. การประมาณค่าแบบจุด
- วิธีแบบเบส์เซียนเป็นวิธีที่เหมาะสำหรับตัวอย่างขนาดเล็ก เนื่องจากให้ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยต่ำสุดทุกค่าพารามิเตอร์ p ที่ทำการศึกษา
- วิธีภาวะน่าจะเป็นสูงสุด เหมาะสำหรับตัวอย่างขนาดปานกลางและขนาดใหญ่ เมื่อค่าพารามิเตอร์ p มีค่าน้อย
- วิธีเบส์เซียนและวิธีภาวะน่าจะเป็นสูงสุดให้ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยไม่แตกต่างกัน
- เมื่อค่าพารามิเตอร์ p เพิ่มขึ้น ทั้งสามวิธีให้ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยที่ไม่แตกต่างกัน
- และวิธีมินิแมกซ์เป็นวิธีที่ไม่ควรเลือกใช้ เพราะให้ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยสูงกว่า 2 วิธีแรก และมีการคำนวณที่ซับซ้อนกว่า

2. การประมาณค่าแบบช่วง
- วิธีปกติ เป็นวิธีที่เหมาะสำหรับตัวอย่างมีขนาดใหญ่ เมื่อค่าพารามิเตอร์ p เท่ากับ 0.05 ถึง 0.10
- วิธีโลจิส เหมาะสำหรับตัวอย่างมีขนาดเล็ก เมื่อค่าพารามิเตอร์ p เท่ากับ 0.07 ถึง 0.50 และเหมาะสำหรับตัวอย่างมีขนาดปานกลาง เมื่อค่าพารามิเตอร์ p เท่ากับ 0.05 ถึง 0.10
- วิธีสคอร์แบบปรับความต่อเนื่อง เหมาะสำหรับตัวอย่างมีขนาดเล็ก เมื่อค่าพารามิเตอร์ p เท่ากับ 0.05 และเหมาะสำหรับตัวอย่างมีขนาดปานกลางและใหญ่เมื่อค่าพารามิเตอร์ p เท่ากับ 0.30 ถึง 0.50

นฤดี (2547) ได้ศึกษาเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์แบบช่วงของการแจกแจงแบบเอ็กซ์โพเนนเชียลโดยการเปรียบเทียบค่าสัมประสิทธิ์ความเชื่อมั่นและค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นของแต่ละวิธีการประมาณ โดยวิธีการประมาณ 4 วิธี คือ วิธีการประมาณด้วยรากของสมการกำลังสอง วิธีการ

ประมาณโดยใช้ตัวอย่างขนาดใหญ่ วิธีการประมาณโดยใช้ตัวประมาณแบบจุด และวิธีการประมาณด้วยช่วงแบบเบส์ ผลการวิจัยสรุปได้ดังนี้

- **ค่าสัมประสิทธิ์ความเชื่อมั่นจากการทดลอง**

ที่ทุกระดับความเชื่อมั่น และทุกค่าพารามิเตอร์ที่ทำการศึกษา เมื่อตัวอย่างมีขนาดเล็ก ($n \leq 30$) พบว่า วิธีการประมาณโดยใช้ตัวประมาณแบบจุดกรณี n ทุกขนาด และการประมาณด้วยช่วงแบบเบส์โดยวิธี Jeffreys prior กรณี n ทุกขนาด จะมีประสิทธิภาพดีกว่าการประมาณด้วยช่วงแบบเบส์โดยวิธี Laplace prior กรณี n ทุกขนาด เพราะให้ค่าสัมประสิทธิ์ความเชื่อมั่นที่คำนวณได้จากการทดลองไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนดในทุกขนาดตัวอย่าง และเมื่อตัวอย่างมีขนาดใหญ่ ($n > 30$) พบว่า วิธีการประมาณด้วยรากของสมการกำลังสอง , วิธีการประมาณโดยใช้ตัวอย่างขนาดใหญ่ , วิธีการประมาณโดยใช้ตัวประมาณแบบจุดกรณี n ทุกขนาด , การประมาณด้วยช่วงแบบเบส์โดยวิธี Jeffreys prior กรณี n ทุกขนาด , การประมาณด้วยช่วงแบบเบส์โดยวิธี Jeffreys prior กรณี n ขนาดใหญ่ , การประมาณด้วยช่วงแบบเบส์โดยวิธี Laplace prior กรณี n ทุกขนาด และการประมาณด้วยช่วงแบบเบส์โดยวิธี Laplace prior กรณี n ขนาดใหญ่ จะมีประสิทธิภาพดีกว่าวิธีการประมาณโดยใช้ตัวประมาณแบบจุดกรณี n ขนาดใหญ่ เพราะให้ค่าสัมประสิทธิ์ความเชื่อมั่นที่คำนวณได้จากการทดลองไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนดในทุกขนาดตัวอย่าง

- **ค่าความยาวเฉลี่ยของช่วงความเชื่อมั่น**

ที่ทุกระดับความเชื่อมั่น และทุกค่าพารามิเตอร์ที่ทำการศึกษา เมื่อตัวอย่างมีขนาดเล็ก ($n \leq 30$) พบว่า การประมาณด้วยช่วงแบบเบส์โดยวิธี Laplace prio r กรณี n ทุกขนาด จะให้ค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นที่แคบที่สุดเมื่อค่าสัมประสิทธิ์ความเชื่อมั่นที่คำนวณได้จาก

การทดลองไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนดในทุกขนาดตัวอย่าง และเมื่อตัวอย่างมีขนาดใหญ่ (n >30) พบว่า วิธีการประมาณโดยใช้ตัวประมาณแบบจุด กรณี n ขนาดใหญ่ จะให้ค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นที่แคบที่สุดเมื่อค่าสัมประสิทธิ์ความเชื่อมั่นที่คำนวณได้จากการทดลองไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนดในทุกขนาดตัวอย่าง

บทที่ 3

วิธีการดำเนินการวิจัย

1. การวางแผนการวิจัย

ในการศึกษาครั้งนี้กำหนดสถานการณ์ในการศึกษาเปรียบเทียบดังนี้

1.1กำหนดขนาดตัวอย่าง (n) ดังนี้

1.1.1 ตัวอย่างมีขนาดเล็ก คือ 5, 25

1.1.2 ตัวอย่างมีขนาดกลาง คือ 50, 80

1.1.3 ตัวอย่างมีขนาดใหญ่ คือ 100, 150

1.2 กำหนดค่าพารามิเตอร์(θ) ที่ใช้ในการศึกษา ได้แก่ 0.1, 0.2, 0.3, 0.4, 0.5

2. ทำการจำลองข้อมูลโดยใช้เทคนิคการจำลอง (Simulation Technique) ด้วยโปรแกรม SAS เวอร์ชัน 9 โดยทำการจำลองข้อมูลที่มีการแจกแจงแบบทวินาม 500 รอบ ตามสถานการณ์ต่างๆที่กำหนดไว้ข้างต้น
3. หาตัวประมาณแบบเบส์โดย มีการแจกแจงเบื้องต้นของ θ ต่าง ๆดังนี้

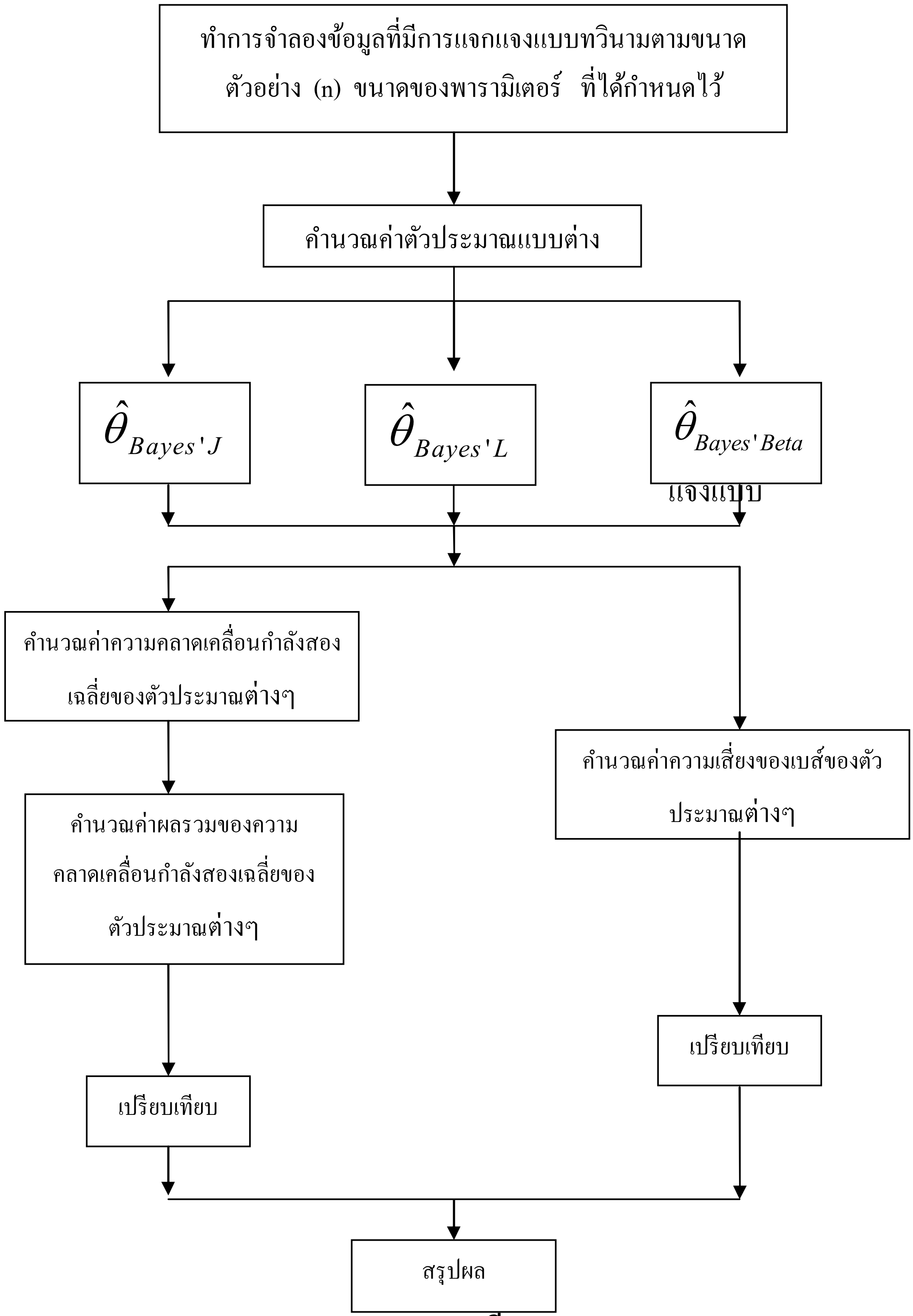
1.1 ตามวิธีของ Hardd Jeffreys

1.2 ตามวิธีของ Laplace prior (เป็นการแจกแจงแบบ Uniform แบบหนึ่ง)

1.3 เป็นการแจกแจงแบบ Beta (a,b)

4. คำนวณค่าประมาณของพารามิเตอร์ในแต่ละสถานการณ์ตามขนาดตัวอย่างและค่าความน่าจะเป็นของการเกิดผลสำเร็จในการแจกแจงทวินามที่กำหนดไว้ โดยวิธีการประมาณแบบเบส์ภายใต้ฟังก์ชันสูญเสียคลาดเคลื่อนกำลังสอง ตามการแจกแจงเบื้องต้นที่กำหนด
5. คำนวณค่าความคลาดเคลื่อนกำลังสองเฉลี่ย และค่าความเสี่ยงของเบส์ ในแต่ละสถานการณ์ตามขนาดตัวอย่างและค่าความน่าจะเป็นของการเกิดผลสำเร็จในการแจกแจงทวินามโดยวิธีการประมาณแบบเบส์ภายใต้ฟังก์ชันสูญเสียคลาดเคลื่อนกำลังสอง ตามการแจกแจงเบื้องต้นที่กำหนด
6. ทำการเปรียบเทียบค่าความคลาดเคลื่อนกำลังสองเฉลี่ย และค่าความเสี่ยงของเบส์ ของตัวประมาณแบบเบส์ภายใต้ฟังก์ชันสูญเสียคลาดเคลื่อนกำลังสองตามการแจกแจงเบื้องต้นที่กำหนด
7. สรุปผลการวิจัยว่าวิธีใดเป็นวิธีที่ดีที่สุดในการประมาณค่าพารามิเตอร์ในสถานการณ์ต่าง ๆ ที่ได้กำหนดไว้

แผนภาพแสดงขั้นตอนการดำเนินงาน



บทที่ 4

ผลการวิจัย

การวิจัยนี้มีจุดมุ่งหมายเพื่อศึกษาตัวประมาณค่าพารามิเตอร์ของการแจกแจงแบบทวินาม โดยวิธีการประมาณแบบเบส์ โดยมีการแจกแจงเบื้องต้นของ θ ต่างกัน โดยจะทำการเปรียบเทียบค่าคลาดเคลื่อนกำลังสองเฉลี่ย และค่าความเสี่ยงเบส์ที่คำนวณได้ในแต่ละสถานการณ์ ถ้าตัวประมาณใดให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ย หรือค่าความเสี่ยงของเบส์ ต่ำกว่า แสดงว่าเป็นตัวประมาณที่เหมาะสมในการนำมาอนุมานพารามิเตอร์ในแต่ละสถานการณ์ โดยจะนำเสนอเป็นตารางและกราฟ ดังนี้

4.1 <u>ตารางค่าประมาณของพารามิเตอร์ และค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณแบบเบส์ ที่มีการแจกแจงเบื้องต้นต่าง ๆ จำแนกตามขนาดตัวอย่าง และค่าพารามิเตอร์ (θ)</u>

<u>ตารางที่</u> 1. ค่าประมาณของพารามิเตอร์และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณแบบเบส์ที่มีการแจกแจงเบื้องต้น ต่าง ๆ จำแนกตามขนาดตัวอย่าง และค่าความน่าจะเป็นของการเกิดผลสำเร็จในการแจกแจงทวินาม $\theta = 0.1$

$\theta = 0.1$	**Jeffreys**	**Laplace**	**Beta (0.3,0.3)**	**Beta (1,3)**	**Beta (2,2)**	**Beta (2,3)**	**Beta (3,2)**
n=5	0.1673	0.2149	0.1435	0.1671	0.2782	0.2504	0.3504
MSE	(0.017367)	(0.022620)	(0.016630)	(0.010207)	(0.037467)	(0.027240)	(0.067320)
n=25	0.1154	0.1297	0.1094	0.1207	0.1552	0.1500	0.1834
MSE	(0.003727)	(0.004116)	(0.003687)	(0.003234)	(0.005855)	(0.005126)	(0.009575)
n=50	0.1085	0.1160	0.1054	0.1117	0.1302	0.1279	0.1461
MSE	(0.001908)	(0.002024)	(0.001894)	(0.001776)	(0.002555)	(0.002358)	(0.003704)
n=80	0.1059	0.1107	0.1039	0.1080	0.1199	0.1185	0.1303
MSE	(0.001137)	(0.001190)	(0.001128)	(0.001090)	(0.001423)	(0.001345)	(0.001920)
n=100	0.1045	0.1084	0.1030	0.1063	0.1159	0.1148	0.1244
MSE	(0.000910)	(0.000943)	(0.000905)	(0.000879)	(0.001093)	(0.001043)	(0.001418)
n=150	0.1040	0.1066	0.1029	0.1052	0.1117	0.1110	0.1174
MSE	(0.000636)	(0.000655)	(0.000632)	(0.000623)	(0.000733)	(0.000709)	(0.000893)

<u>หมายเหตุ</u> ค่าในวงเล็บคือค่าMSE

จากตารางที่ 1. พบว่าตัวประมาณแบบเบส์ ที่มีการแจกแจงเบื้องต้นเป็นแบบ Beta (1,3) ให้ค่า MSE ต่ำกว่า ตัวประมาณแบบเบส์ ที่มีการแจกแจงเบื้องต้นอื่น ๆ

ตารางที่ 2. ค่าประมาณของพารามิเตอร์และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณแบบเบส์ที่มีการแจกแจงเบื้องต้น ต่าง ๆ จำแนกตามขนาดตัวอย่าง และค่าความน่าจะเป็นของการเกิดผลสำเร็จในการแจกแจงทวินาม $\theta = 0.2$

$\theta = 0.2$	**Jeffreys**	**Laplace**	**Beta (0.3,0.3)**	**Beta (1,3)**	**Beta (2,2)**	**Beta (2,3)**	**Beta (3,2)**
n=5	0.2350	0.2728	0.2160	0.2122	0.3233	0.2910	0.3910
MSE	(0.022944)	(0.021265)	(0.025191)	(0.009802)	(0.024864)	(0.016100)	(0.044300)
n=25	0.2123	0.2229	0.2078	0.2075	0.2420	0.2340	0.2673
	(0.006452)	(0.006370)	(0.006560)	(0.005122)	(0.006834)	(0.005888)	(0.009266)
n=50	0.2068	0.2124	0.2045	0.2045	0.2231	0.2190	0.2372
	(0.003370)	(0.003352)	(0.003397)	(0.002986)	(0.003499)	(0.003221)	(0.004244)
n=80	0.2037	0.2073	0.2022	0.2024	0.2143	0.2118	0.2335
	(0.002002)	(0.001994)	(0.002013)	(0.001854)	(0.002054)	(0.001944)	(0.002361)
n=100	0.2031	0.2060	0.2019	0.2021	0.2117	0.2097	0.2192
	(0.001567)	(0.001564)	(0.001574)	(0.001473)	(0.001606)	(0.001535)	(0.001811)
n=150	0.2031	0.2050	0.2023	0.2024	0.2089	0.2075	0.2140
	(0.001132)	(0.001133)	(0.001133)	(0.001084)	(0.001158)	(0.001122)	(0.001261)

จากตารางที่ 2. พบว่าตัวประมาณแบบเบส์ ที่มีการแจกแจงเบื้องต้นเป็นแบบ Beta (1,3) ให้ค่า MSE ต่ำกว่า ตัวประมาณแบบเบส์ ที่มีการแจกแจงเบื้องต้นอื่น ๆ

ตารางที่ 3. ค่าประมาณของพารามิเตอร์และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณแบบเบส์ที่มีการแจกแจงเบื้องต้น ต่าง ๆ จำแนกตามขนาดตัวอย่าง และค่าความน่าจะเป็นของการเกิดผลสำเร็จในการแจกแจงทวินาม $\theta = 0.3$

$\theta = 0.3$	**Jeffreys**	**Laplace**	**Beta (0.3,0.3)**	**Beta (1,3)**	**Beta (2,2)**	**Beta (2,3)**	**Beta (3,2)**
n=5	0.3213	0.3468	0.3085	0.2697	0.3808	0.3428	0.4428
	(0.028144)	(0.022539)	(0.031860)	(0.013220)	(0.018849)	(0.011800)	(0.030360)
n=25	0.3086	0.3157	0.3056	0.2939	0.3284	0.3174	0.3508
	(0.007472)	(0.007107)	(0.007663)	(0.005983)	(0.006754)	(0.005862)	(0.008137)
n=50	0.3062	0.3099	0.3046	0.2984	0.3170	0.3112	0.3294
	(0.003854)	(0.003769)	(0.003898)	(0.003406)	(0.003692)	(0.003405)	(0.004146)
n=80	0.3028	0.3052	0.3018	0.2979	0.3098	0.3062	0.3179
	(0.002368)	(0.002330)	(0.002387)	(0.002199)	(0.002292)	(0.002182)	(0.002466)
n=100	0.3021	0.3041	0.3013	0.2982	0.3078	0.3049	0.3144
	(0.001886)	(0.001862)	(0.001899)	(0.001778)	(0.001837)	(0.001765)	(0.001951)
n=150	0.3023	0.3036	0.3018	0.2997	0.3061	0.3042	0.3106
	(0.001472)	(0.001460)	(0.001477)	(0.001410)	(0.001448)	(0.001409)	(0.001506)

จากตารางที่ 3. พบว่าตัวประมาณแบบเบส์ ที่มีการแจกแจงเบื้องต้นเป็นแบบ Beta (2,3) ให้ค่า MSE ต่ำกว่า ตัวประมาณแบบเบส์ ที่มีการแจกแจงเบื้องต้นอื่น ๆ

ตารางที่ 4. ค่าประมาณของพารามิเตอร์และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณแบบเบส์ที่มีการแจกแจงเบื้องต้น ต่าง ๆ จำแนกตามขนาดตัวอย่าง และค่าความน่าจะเป็นของการเกิดผลสำเร็จในการแจกแจงทวินาม $\theta = 0.4$

$\theta = 0.4$	**Jeffreys**	**Laplace**	**Beta (0.3,0.3)**	**Beta (1,3)**	**Beta (2,2)**	**Beta (2,3)**	**Beta (3,2)**
n=5	0.4043	0.4180	0.3975	0.3251	0.4362	0.3926	0.4926
	(0.032700)	(0.024335)	(0.037523)	(0.020133)	(0.015837)	(0.011820)	(0.020340)
n=25	0.4022	0.4058	0.4007	0.3778	0.4123	0.3986	0.4319
	(0.008543)	(0.007952)	(0.008808)	(0.007353)	(0.007015)	(0.006415)	(0.007433)
n=50	0.4038	0.4056	0.4030	0.3906	0.4091	0.4017	0.4199
	(0.004100)	(0.003962)	(0.004160)	(0.003731)	(0.003728)	(0.003516)	(0.003910)
n=80	0.4026	0.4038	0.4021	0.3942	0.4061	0.4013	0.4131
	(0.002553)	(0.002499)	(0.002576)	(0.002400)	(0.002404)	(0.002313)	(0.002483)
n=100	0.4028	0.4038	0.4024	0.3960	0.4056	0.4017	0.4113
	(0.001931)	(0.001900)	(0.001944)	(0.001829)	(0.001846)	(0.001782)	(0.001907)
n=150	0.4032	0.4039	0.4030	0.3986	0.4051	0.4025	0.4090
	(0.001596)	(0.001580)	(0.001603)	(0.001525)	(0.001551)	(0.001511)	(0.001585)

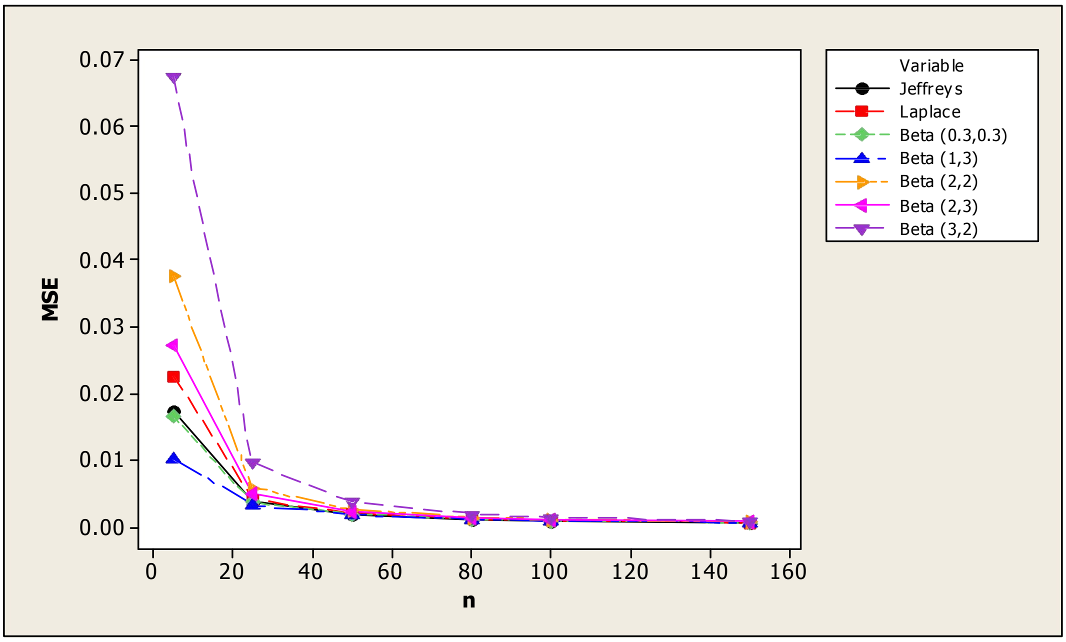
จากตารางที่ 4. พบว่าตัวประมาณแบบเบส์ ที่มีการแจกแจงเบื้องต้นเป็นแบบ Beta (2,3) ให้ค่า MSE ต่ำกว่า ตัวประมาณแบบเบส์ ที่มีการแจกแจงเบื้องต้นอื่น ๆ

ตารางที่ 5. ค่าประมาณของพารามิเตอร์และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณแบบเบส์ที่มีการแจกแจงเบื้องต้น ต่าง ๆ จำแนกตามขนาดตัวอย่าง และค่าความน่าจะเป็นของการเกิดผลสำเร็จในการแจกแจงทวินาม $\theta = 0.5$

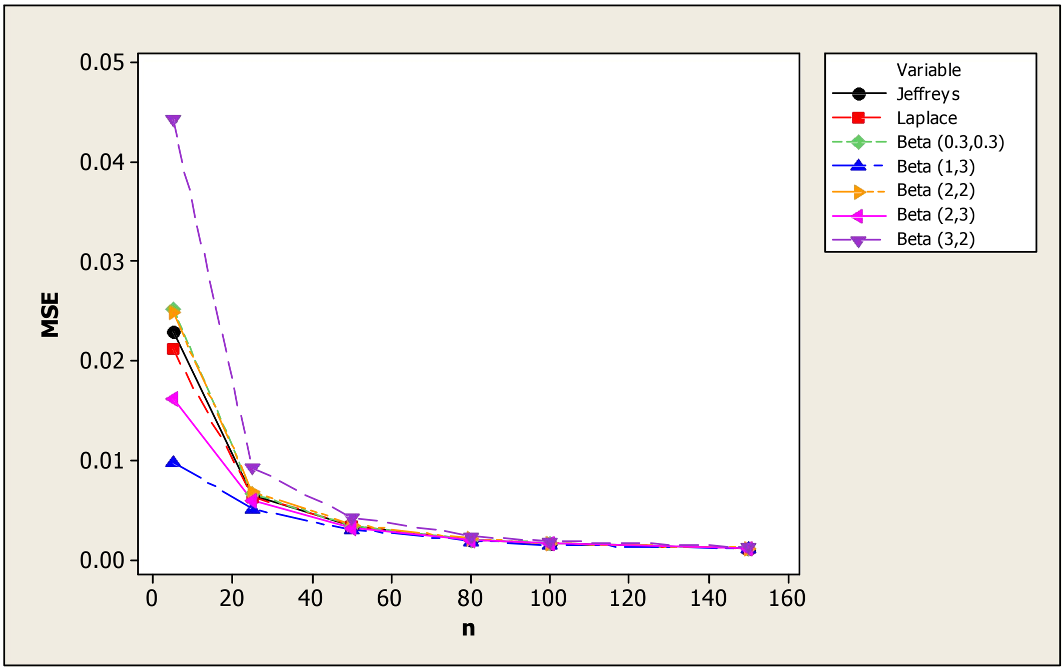
$\theta = 0.5$	**Jeffreys**	**Laplace**	**Beta (0.3,0.3)**	**Beta (1,3)**	**Beta (2,2)**	**Beta (2,3)**	**Beta (3,2)**
n=5	0.4903	0.4917	0.4896	0.3824	0.4935	0.4442	0.5442
	(0.035167)	(0.025837)	(0.040370)	(0.029407)	(0.015630)	(0.015740)	(0.01458)
n=25	0.4980	0.4981	0.4980	0.4637	0.4982	0.4816	0.5150
	(0.009310)	(0.008633)	(0.009603)	(0.008791)	(0.007483)	(0.007326)	(0.007215)
n=50	0.5025	0.5025	0.5026	0.4839	0.5024	0.4933	0.5114
	(0.004796)	(0.004613)	(0.004872)	(0.004530)	(0.004278)	(0.004163)	(0.004250)
n=80	0.5020	0.5020	0.5020	0.4900	0.5019	0.4960	0.5078
	(0.002901)	(0.002831)	(0.002930)	(0.002793)	(0.002698)	(0.002647)	(0.002692)
n=100	0.5023	0.5023	0.5023	0.4926	0.5022	0.4974	0.5070
	(0.002213)	(0.002170)	(0.002231)	(0.002136)	(0.002087)	(0.002049)	(0.002092)
n=150	0.5028	0.5028	0.5028	0.4962	0.5027	0.4995	0.5060
	(0.001675)	(0.001653)	(0.001684)	(0.001616)	(0.001610)	(0.001582)	(0.001618)

จากตารางที่ 5. พบว่าตัวประมาณแบบเบส์ ที่มีการแจกแจงเบื้องต้นเป็นแบบ Beta (2,3) ให้ค่า MSE ต่ำกว่า ตัวประมาณแบบเบส์ ที่มีการแจกแจงเบื้องต้นอื่น ๆ ยกเว้นเมื่อ n=5 และ50 Beta (3,2) ให้ค่า MSE ต่ำกว่า

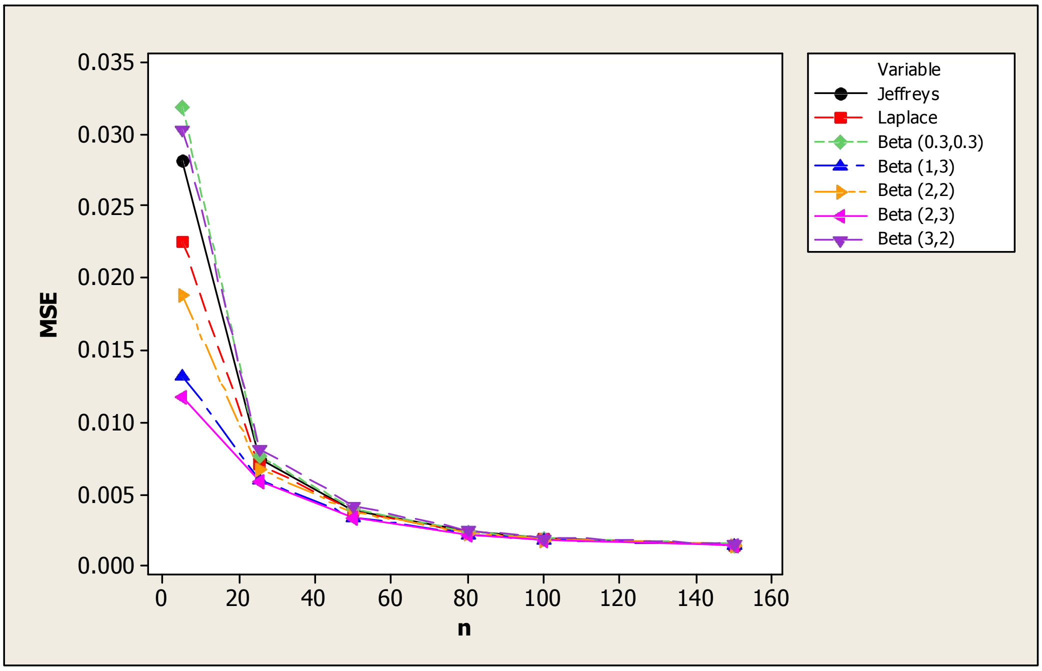
4.2 <u>กราฟค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณแบบเบส์ เมื่อพารามิเตอร์ (θ) มีขนาดต่างๆ</u>



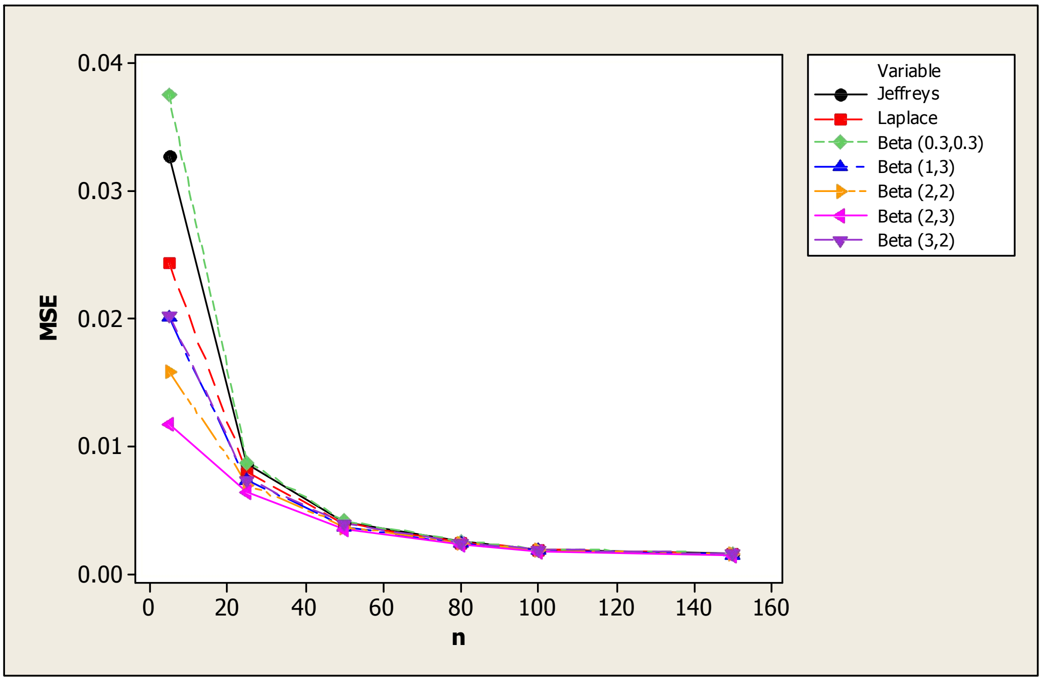
รูปที่ 1. ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณแบบเบส์เมื่อมีการแจกแจงเบื้องต้นต่าง ๆ เมื่อ $\theta = 0.1$



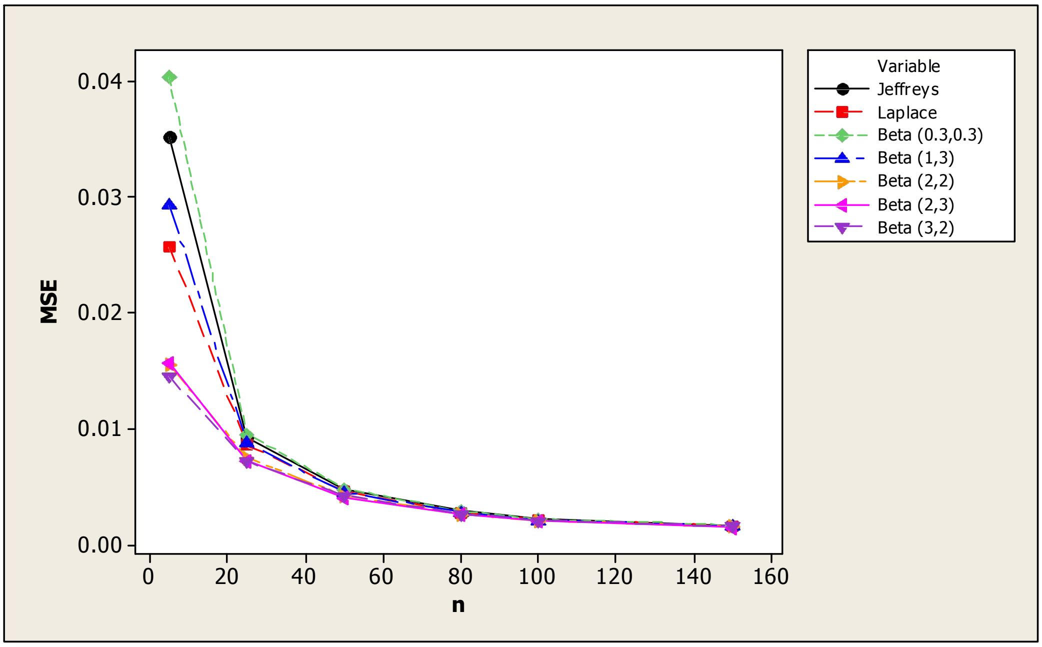
รูปที่ 2. ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณแบบเบส์เมื่อมีการแจกแจงเบื้องต้นต่าง ๆ เมื่อ $\theta = 0.2$



รูปที่ 3. ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณแบบเบส์เมื่อมีการแจกแจงเบื้องต้นต่าง ๆ เมื่อ $\theta = 0.3$



รูปที่ 4. ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณแบบเบส์เมื่อมีการแจกแจงเบื้องต้นต่าง ๆ เมื่อ $\theta = 0.4$



รูปที่ 5. ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณแบบเบส์เมื่อมีการแจกแจงเบื้องต้นต่าง ๆ เมื่อ $\theta = 0.5$

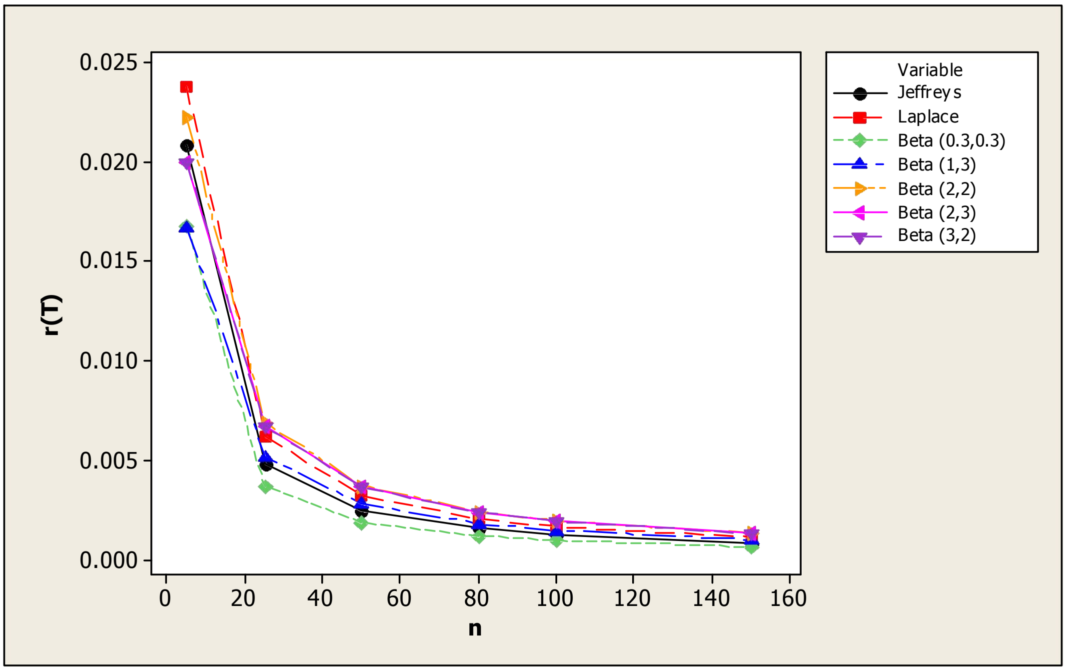
จากกราฟรูปที่ 1-5 พบว่าเมื่อพารามิเตอร์ θ มีขนาดเล็กตัวประมาณแบบเบส์ที่มีการแจกแจงเบื้องต้นเป็น Beta (1,3) เป็นตัวประมาณที่ให้ค่า MSE ต่ำกว่าตัวประมาณอื่นๆ แต่เมื่อพารามิเตอร์ θ มีขนาดเพิ่มขึ้นพบว่าตัวประมาณแบบเบส์ที่มีการแจกแจงเบื้องต้นเป็นแบบ Beta (2,3) ให้ค่า MSE ต่ำกว่าตัวประมาณแบบอื่นๆ ยกเว้นเมื่อตัวอย่างขนาดเล็ก ตัวประมาณแบบเบส์ที่มีการแจกแจงเบื้องต้นเป็นแบบ Beta (3,2) ให้ค่า MSE ต่ำกว่าตัวประมาณแบบอื่นๆ และจะพบว่าค่า MSE มีค่าใกล้เคียงกันเมื่อขนาดตัวอย่างมีขนาดเพิ่มขึ้น

4.3 ตารางค่าความเสี่ยงของเบส์จำแนกตามขนาดตัวอย่างและตามการแจกแจงเบื้องต้นที่กำหนด

ตารางที่ 6. ค่าความเสี่ยงของเบส์จำแนก ตามขนาดตัวอย่างและการแจกแจงเบื้องต้นที่กำหนด

	Jeffreys	**Laplace**	**Beta (0.3,0.3)**	**Beta (1,3)**	**Beta (2,2)**	**Beta (2,3)**	**Beta (3,2)**
n=5	0.0208333	0.0238095	0.0167411	0.0166667	0.0222222	0.0200000	0.0200000
n=25	0.0048077	0.0061728	0.0036621	0.0051724	0.0068966	0.0066667	0.0066667
n=50	0.0024510	0.0032051	0.0018528	0.0027778	0.0037037	0.0036364	0.0036364
n=80	0.0015432	0.0020325	0.0011632	0.0017857	0.0023810	0.0023529	0.0023529
n=100	0.0012376	0.0016340	0.0009319	0.0014423	0.0019231	0.0019048	0.0019048
n=150	0.0008278	0.0010965	0.0006225	0.0009740	0.0012987	0.0012903	0.0012903

4.4 กราฟค่าความเสี่ยงของเบส์ จำแนกตามขนาดตัวอย่าง และการแจกแจงเบื้องต้นที่กำหนด



รูปที่ 6. ค่าความเสี่ยงของเบส์ โดยจำแนกตามขนาดตัวอย่าง และการแจกแจงเบื้องต้นที่กำหนด

จากตารางที่ 6. และกราฟรูปที่ 6. พบว่า เมื่อตัวอย่างมีขนาดเล็ก ตัวประมาณแบบเบส์ที่มีการแจกแจงเบื้องต้น เป็นแบบ Beta (1,3) (เมื่อ n=5) และ Beta (0.3,0.3) (เมื่อ n=25) มีค่าความเสี่ยงของเบส์ต่ำที่สุด และเมื่อตัวอย่างมีขนาดกลาง และขนาดใหญ่ ตัวประมาณแบบเบส์ที่มีการแจกแจงเบื้องต้น เป็นแบบBeta (0.3,0.3) มีค่าความเสี่ยงของเบส์ต่ำที่สุด

บทที่ 5

สรุปผลการวิจัย

การวิจัยครั้งนี้เป็นการศึกษาเพื่อเปรียบเทียบค่าประมาณของพารามิเตอร์ของการแจกแจงแบบทวินาม โดยวิธีการประมาณแบบเบส์ภายใต้ฟังก์ชันสูญเสียคลาดเคลื่อนกำลังสอง โดยมีการแจกแจงเบื้องต้นต่างๆ คือ วิธี Hardd Jeffreys วิธี Laplace prior และมีการแจกแจงแบบ Beta(a,b) ที่มีพารามิเตอร์เป็น (0.3,0.3), (1,3), (2,2), (2,3) และ(3,2) ผลการวิจัยสรุปได้ดังตารางต่อไปนี้

ตารางที่7. การแจกแจงเบื้องต้นของวิธีการประมาณแบบเบส์ ที่เหมาะสมในการประมาณค่าแบบจุด จำแนกตามขนาดตัวอย่าง และค่าพารามิเตอร์ (θ)

	พารามิเตอร์ (θ)				
	0.1	**0.2**	**0.3**	**0.4**	**0.5**
n=5	Beta (1,3)	Beta (1,3)	Beta (2,3)	Beta (2,3)	Beta (3,2)
n=25	Beta (1,3)	Beta (1,3)	Beta (2,3)	Beta (2,3)	Beta (3,2)
n=50	Beta (1,3)	Beta (1,3)	Beta (2,3)	Beta (2,3)	Beta (2,3)
n=80	Beta (1,3)	Beta (1,3)	Beta (2,3)	Beta (2,3)	Beta (2,3)
n=100	Beta (1,3)	Beta (1,3)	Beta (2,3)	Beta (2,3)	Beta (2,3)
n=150	Beta (1,3)	Beta (1,3)	Beta (2,3)	Beta (2,3)	Beta (2,3)

จากตารางที่ 7. สามารถสรุปผลได้ว่าเมื่อใช้ขนาดของพารามิเตอร์ในการพิจารณาจะพบว่าตัวประมาณที่เหมาะสมสำหรับพารามิเตอร์ที่มีขนาดเล็กคือตัวประมาณแบบเบส์ ที่มีการแจกแจงเบื้องต้นเป็นแบบ Beta (1,3) และเมื่อขนาดพารามิเตอร์เพิ่มขึ้นพบว่า ตัวประมาณที่เหมาะสม คือตัวประมาณแบบเบส์ที่มีการแจกแจงเบื้องต้นเป็นแบบ Beta (2,3) ยกเว้นเมื่อตัวอย่างขนาดเล็ก ตัวประมาณแบบเบส์ที่มีการแจกแจงเบื้องต้นเป็นแบบ Beta (3,2) จะเป็นตัวประมาณที่เหมาะสม และจะสังเกตได้ว่าตัวประมาณค่าทุกตัวมีค่า MSE ลดลง เมื่อขนาดตัวอย่างมีขนาดเพิ่มขึ้น

และถ้าพิจารณาค่าความเสี่ยงของเบส์เมื่อจำแนกตามขนาดตัวอย่าง จากตารางที่ 6. สามารถสรุปผลได้ว่า ตัวประมาณที่เหมาะสมสำหรับตัวอย่างขนาดเล็กคือตัวประมาณแบบเบส์ที่มีการแจกแจงเบื้องต้นเป็นแบบ Beta (1,3) (เมื่อ n=5) และ Beta (0.3,0.3) (เมื่อ n=25) สำหรับตัวอย่างขนาดกลาง และขนาดใหญ่ตัวประมาณแบบเบส์ ที่มีการแจกแจงเบื้องต้นเป็นแบบ Beta (0.3,0.3) จะเป็นตัวประมาณที่ดีที่สุด

ข้อเสนอแนะ

1. การวิจัยในครั้งนี้สามารถนำไปประยุกต์ใช้ในการประมาณแบบเบส์ที่มีการแจกแจงเบื้องต้นแบบอื่นๆ หรือ ทำการเปรียบเทียบตัวประมาณค่าพารามิเตอร์สำหรับการแจกแจงอื่นๆ โดยมีการแจกแจงเบื้องต้นเป็นแบบที่ใช้ในการศึกษาครั้งนี้เป็นต้น

2. ศึกษาพารามิเตอร์ (a,b) ที่เหมาะสมที่ทำให้ค่า MSE ต่ำที่สุดสำหรับตัวประมาณแบบเบส์ ที่มีการแจกแจงเบื้องต้นเป็นแบบ Beta ในสถานการณ์ต่างๆ

เอกสารอ้างอิง

ประชุม สุวัตถี. **ทฤษฎีการอนุมานเชิงสถิติ.** กรุงเทพมหานคร: สถาบันบัณฑิตพัฒนบริหารศาสตร์ , 2545.

มานพ วราภักดิ์ . **ทฤษฎีความน่าจะเป็น**. พิมพ์ครั้งที่ 1. กรุงเทพมหานคร : สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2545.

รัชกมล กบิลจิตร. **สถิติอนุมาน : ทฤษฎีและการประยุกต์**. พิมพ์ครั้งที่ 1. กรุงเทพมหานคร : สำนักพิมพ์จุฬาลงกรณ์มหาวิทยาลัย, 2548.

สายชล สินสมบูรณ์ทอง. **สถิติคณิตศาสตร์1**. พิมพ์ครั้งที่ 4. กรุงเทพมหานคร : จามจุรีโปรดักท์, 2549

พาชิตชนัต ศิริพานิช. **วิธีประมาณค่าแบบเบยส์**. เอกสารประกอบการสอนวิชาสถิติอนุมาน สาขาวิชาสถิติ คณะสถิติประยุกต์ สถาบันบัณฑิตพัฒนบริหารศาสตร์, 2551.

ฟองนวล วงศ์ตะวัน. **การเปรียบเทียบวิธีการประมาณค่าตัวประมาณสัดส่วนในการแจกแจงแบบทวินามและตัวประมาณค่าเฉลี่ยของการแจกแจงแบบปัวส์ซอง**. วิทยานิพนธ์วิทยาศาสตร์มหาบัณฑิต สาขาวิชาสถิติประยุกต์ภาควิชาสถิติ มหาวิทยาลัยนเรศวร,2549.

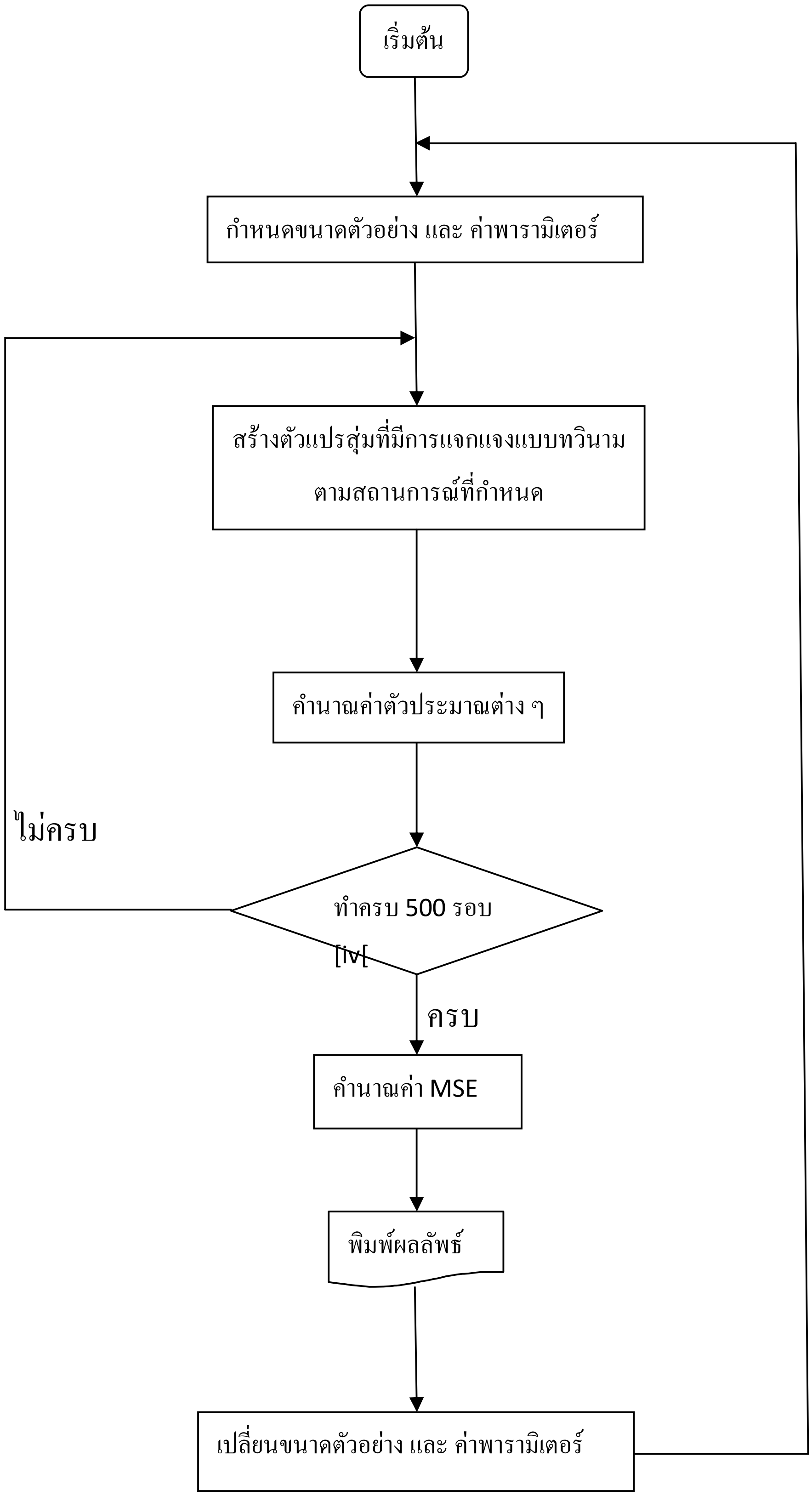
ฝันจิต แต้มทอง. **การอนุมานค่าเฉลี่ยประชากรโดยวิธีเบย์เซียน.** การค้นคว้าแบบอิสระเชิงวิทยานิพนธ์, วิทยาศาสตร์มหาบัณฑิต (สถิติประยุกต์), มหาวิทยาลัยเชียงใหม่, 2537.

Ross,S. 1994. **A First Course in Probability** . Macmillan., New York.

George Casella, Roger L. Berger. 2002. **Statistical Inference**. Second Edition. USA.

ภาคผนวก

ขั้นตอนการประมาณค่า



ตัวอย่างโปรแกรม

1. ตัวอย่างโปรแกรม SAS ที่ใช้ในการวิจัยการเปรียบเทียบตัวประมาณค่าพารามิเตอร์ สำหรับการแจกแจงแบบทวินามด้วยวิธีการประมาณแบบเบส์ภายใต้ฟังก์ชันสูญเสียคลาดเคลื่อนกำลังสอง (Squared error loss function) โดยมีการแจกแจงเบื้องต้นตามวิธี Hardd Jeffreys

....

Main program

....

N=25 Jeffreys

```
/*sample1_25*/

data bin_1;

array y[25] y_1 - y_25;
      do  r = 1 to 500;
            do s = 1 to 25 ;
            y[s] = ranbin(88888,1,0.1);

          sumy = 0;
            zeta1 = 0;
            diff = 0;
            end;

            output;
      end;
run;
data calc_1;
```

```
        set bin_1;
        keep  y_1 - y_25 sumy zeta1 diff;

        array y[25] y_1 - y_25;
              do  n = 1 to 25;
               sumy = y[n] + sumy ;
              end;
        zeta1 = (sumy+0.5)/((sumy+0.5)+(25-sumy+0.5));
  diff = (0.1*0.1-2*0.1*zeta1+zeta1*zeta1)/500;

  run;
  proc print data=calc_1;
  sum diff;
  sum zeta1;
  run;

  /*sample2_25*/

  data bin_2;

  array y[25] y_1 - y_25;
        do  r = 1 to 500;
              do s = 1 to 25 ;
              y[s] = ranbin(88888,1,0.2);

            sumy = 0;
              zeta1 = 0;
              diff = 0;
              end;
```

```
            output;
        end;
run;
data calc_2;
        set bin_2;
        keep  y_1 - y_25 sumy zeta1 diff;

        array y[25] y_1 - y_25;
            do  n = 1 to 25;
             sumy = y[n] + sumy ;
            end;
        zeta1 = (sumy+0.5)/((sumy+0.5)+(25-sumy+0.5));
diff = (0.2*0.2-2*0.2*zeta1+zeta1*zeta1)/500;
run;
proc print data=calc_2;
sum diff;
sum zeta1;
run;

/*sample3_25*/

data bin_3;

array y[25] y_1 - y_25;
        do  r = 1 to 500;
            do s = 1 to 25 ;
            y[s] = ranbin(88888,1,0.3);
```

```
            sumy = 0;
              zeta1 = 0;
              diff = 0;
              end;
              output;
        end;
run;
data calc_3;
        set bin_3;
        keep  y_1 - y_25 sumy zeta1 diff;

        array y[25] y_1 - y_25;
              do  n = 1 to 25;
               sumy = y[n] + sumy ;
              end;
        zeta1 = (sumy+0.5)/((sumy+0.5)+(25-sumy+0.5));
diff = (0.3*0.3-2*0.3*zeta1+zeta1*zeta1)/500;
run;
proc print data=calc_3;
sum diff;
sum zeta1;
run;

/*sample4_25*/

data bin_4;

array y[25] y_1 - y_25;
        do  r = 1 to 500;
```

```
            do s = 1 to 25 ;
            y[s] = ranbin(88888,1,0.4);

          sumy = 0;
            zeta1 = 0;
            diff = 0;
            end;
output;
      end;
run;
data calc_4;
      set bin_4;
      keep  y_1 - y_25 sumy zeta1 diff;

      array y[25] y_1 - y_25;
            do  n = 1 to 25;
             sumy = y[n] + sumy ;
            end;
      zeta1 = (sumy+0.5)/((sumy+0.5)+(25-sumy+0.5));
diff = (0.4*0.4-2*0.4*zeta1+zeta1*zeta1)/500;
run;
proc print data=calc_4;
sum diff;
sum zeta1;
run;

/*sample5_25*/

data bin_5;
```

```
array y[25] y_1 - y_25;
      do  r = 1 to 500;
            do s = 1 to 25 ;
            y[s] = ranbin(88888,1,0.5);

          sumy = 0;
            zeta1 = 0;
            diff = 0;
            end;
            output;
      end;
run;
data calc_5;
      set bin_5;
      keep  y_1 - y_25 sumy zeta1 diff;

      array y[25] y_1 - y_25;
            do  n = 1 to 25;
             sumy = y[n] + sumy ;
            end;
      zeta1 = (sumy+0.5)/((sumy+0.5)+(25-sumy+0.5));
     diff = (0.5*0.5-2*0.5*zeta1+zeta1*zeta1)/500;
run;
proc print data=calc_5;
sum diff;
sum zeta1;
run;
```

2. ตัวอย่างโปรแกรม SAS ที่ใช้ในการวิจัยการเปรียบเทียบตัวประมาณค่าพารามิเตอร์ สำหรับการแจกแจงแบบทวินามด้วยวิธีการประมาณแบบเบส์ภายใต้ฟังก์ชันสูญเสียคลาดเคลื่อนกำลังสอง (Squared error loss function) โดยมีการแจกแจงเบื้องต้นตามวิธี Laplace

Main program

N=25 Laplace

```
/*sample1_25*/

data bin_1;

array y[25] y_1 - y_25;
      do  r = 1 to 500;
            do s = 1 to 25 ;
            y[s] = ranbin(88888,1,0.1);

          sumy = 0;
            zeta1 = 0;
            diff = 0;
            end;

            output;
      end;
run;
data calc_1;
      set bin_1;
```

```
        keep  y_1 - y_25 sumy zeta1 diff;

        array y[25] y_1 - y_25;
              do  n = 1 to 25;
               sumy = y[n] + sumy ;
              end;
        zeta1 = (sumy+1)/((sumy+1)+(25-sumy+1));
  diff = (0.1*0.1-2*0.1*zeta1+zeta1*zeta1)/500;

  run;
  proc print data=calc_1;
  sum diff;
  sum zeta1;
  run;

  /*sample2_25*/

  data bin_2;

  array y[25] y_1 - y_25;
        do  r = 1 to 500;
              do s = 1 to 25 ;
              y[s] = ranbin(88888,1,0.2);

            sumy = 0;
              zeta1 = 0;
              diff = 0;
              end;
```

```
            output;
        end;
run;
data calc_2;
        set bin_2;
        keep  y_1 - y_25 sumy zeta1 diff;

        array y[25] y_1 - y_25;
                do  n = 1 to 25;
                 sumy = y[n] + sumy ;
                end;
        zeta1 = (sumy+1)/((sumy+1)+(25-sumy+1));
diff = (0.2*0.2-2*0.2*zeta1+zeta1*zeta1)/500;
run;
proc print data=calc_2;
sum diff;
sum zeta1;
run;

/*sample3_25*/

data bin_3;

array y[25] y_1 - y_25;
        do  r = 1 to 500;
            do s = 1 to 25 ;
            y[s] = ranbin(88888,1,0.3);
          sumy = 0;
            zeta1 = 0;
```

```
            diff = 0;
            end;
            output;
      end;
run;
data calc_3;
      set bin_3;
      keep  y_1 - y_25 sumy zeta1 diff;

      array y[25] y_1 - y_25;
            do  n = 1 to 25;
             sumy = y[n] + sumy ;
            end;
      zeta1 = (sumy+1)/((sumy+1)+(25-sumy+1));
diff = (0.3*0.3-2*0.3*zeta1+zeta1*zeta1)/500;
run;
proc print data=calc_3;
sum diff;
sum zeta1;
run;

/*sample4_25*/

data bin_4;

array y[25] y_1 - y_25;
      do  r = 1 to 500;
            do s = 1 to 25 ;
            y[s] = ranbin(88888,1,0.4);
```

```
            sumy = 0;
              zeta1 = 0;
              diff = 0;
              end;
output;
      end;
run;
data calc_4;
      set bin_4;
      keep  y_1 - y_25 sumy zeta1 diff;

      array y[25] y_1 - y_25;
            do  n = 1 to 25;
             sumy = y[n] + sumy ;
            end;
      zeta1 = (sumy+1)/((sumy+1)+(25-sumy+1));
diff = (0.4*0.4-2*0.4*zeta1+zeta1*zeta1)/500;
run;
proc print data=calc_4;
sum diff;
sum zeta1;
run;


/*sample5_25*/


data bin_5;
```

```
array y[25] y_1 - y_25;
      do  r = 1 to 500;
            do s = 1 to 25 ;
            y[s] = ranbin(88888,1,0.5);

          sumy = 0;
            zeta1 = 0;
            diff = 0;
            end;
            output;
      end;
run;
data calc_5;
      set bin_5;
      keep  y_1 - y_25 sumy zeta1 diff;

      array y[25] y_1 - y_25;
            do  n = 1 to 25;
             sumy = y[n] + sumy ;
            end;
      zeta1 = (sumy+1)/((sumy+1)+(25-sumy+1));
     diff = (0.5*0.5-2*0.5*zeta1+zeta1*zeta1)/500;
run;
proc print data=calc_5;
sum diff;
sum zeta1;
run;
```

3. ตัวอย่างโปรแกรม SAS ที่ใช้ในการวิจัยการเปรียบเทียบตัวประมาณค่าพารามิเตอร์ สำหรับการแจกแจงแบบทวินามด้วยวิธีการประมาณแบบเบส์ภายใต้ฟังก์ชันสูญเสียคลาดเคลื่อนกำลังสอง (Squared error loss function) โดยมีการแจกแจงเบื้องต้นที่มีการแจกแจงแบบ Beta(1,3)

Main program

N=25 Beta(1,3)

```
/*sample1_25*/

data bin_1;

array y[25] y_1 - y_25;
      do  r = 1 to 500;
            do s = 1 to 25 ;
            y[s] = ranbin(88888,1,0.1);

          sumy = 0;
            zeta1 = 0;
            diff = 0;
            end;

            output;
      end;
run;
data calc_1;
      set bin_1;
      keep  y_1 - y_25 sumy zeta1 diff;
```

```
      array y[25] y_1 - y_25;
            do  n = 1 to 25;
             sumy = y[n] + sumy ;
            end;
      zeta1 = (sumy+1)/29;
diff = (0.1*0.1-2*0.1*zeta1+zeta1*zeta1)/500;

run;
proc print data=calc_1;
sum diff;
sum zeta1;
run;

/*sample2_25*/

data bin_2;

array y[25] y_1 - y_25;
      do  r = 1 to 500;
            do s = 1 to 25 ;
            y[s] = ranbin(88888,1,0.2);

          sumy = 0;
            zeta1 = 0;
            diff = 0;
            end;

            output;
```

```
        end;
run;
data calc_2;
        set bin_2;
        keep  y_1 - y_25 sumy zeta1 diff;

        array y[25] y_1 - y_25;
              do  n = 1 to 25;
               sumy = y[n] + sumy ;
              end;
        zeta1 = (sumy+1)/29;
diff = (0.2*0.2-2*0.2*zeta1+zeta1*zeta1)/500;
run;
proc print data=calc_2;
sum diff;
sum zeta1;
run;

/*sample3_25*/

data bin_3;

array y[25] y_1 - y_25;
        do  r = 1 to 500;
              do s = 1 to 25 ;
              y[s] = ranbin(88888,1,0.3);
            sumy = 0;
              zeta1 = 0;
              diff = 0;
```

```
            end;
            output;
      end;
run;
data calc_3;
      set bin_3;
      keep  y_1 - y_25 sumy zeta1 diff;

      array y[25] y_1 - y_25;
            do  n = 1 to 25;
             sumy = y[n] + sumy ;
            end;
      zeta1 = (sumy+1)/29;
diff = (0.3*0.3-2*0.3*zeta1+zeta1*zeta1)/500;
run;
proc print data=calc_3;
sum diff;
sum zeta1;
run;

/*sample4_25*/

data bin_4;

array y[25] y_1 - y_25;
      do  r = 1 to 500;
            do s = 1 to 25 ;
            y[s] = ranbin(88888,1,0.4);
```

```
          sumy = 0;
            zeta1 = 0;
            diff = 0;
            end;
output;
      end;
run;
data calc_4;
      set bin_4;
      keep  y_1 - y_25 sumy zeta1 diff;

      array y[25] y_1 - y_25;
            do  n = 1 to 25;
             sumy = y[n] + sumy ;
            end;
      zeta1 = (sumy+1)/29;
diff = (0.4*0.4-2*0.4*zeta1+zeta1*zeta1)/500;
run;
proc print data=calc_4;
sum diff;
sum zeta1;
run;

/*sample5_25*/

data bin_5;

array y[25] y_1 - y_25;
```

```
       do  r = 1 to 500;
             do s = 1 to 25 ;
             y[s] = ranbin(88888,1,0.5);

           sumy = 0;
             zeta1 = 0;
             diff = 0;
             end;
             output;
       end;
run;
data calc_5;
       set bin_5;
       keep  y_1 - y_25 sumy zeta1 diff;

       array y[25] y_1 - y_25;
             do  n = 1 to 25;
              sumy = y[n] + sumy ;
             end;
       zeta1 = (sumy+1)/29;
      diff = (0.5*0.5-2*0.5*zeta1+zeta1*zeta1)/500;
run;
proc print data=calc_5;
sum diff;
sum zeta1;
run;
```

ประวัติหัวหน้าโครงการวิจัย

1. **ชื่อ-นามสกุล** (ภาษาไทย) นางสาวกาญจนา ใจบุญ

2. **รหัสประจำตัวประชาชน** 3 – 6701 – 00743 – 55 – 1

3. **ตำแหน่งปัจจุบัน** อาจารย์

4. **หน่วยงานที่อยู่ที่ติดต่อได้สะดวก** พร้อมโทรศัพท์ โทรสาร และ E-mail

สาขาวิชาคณิตศาสตร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏเพชรบูรณ์ ตั้งอยู่ อาคาร 4 ชั้น 2 ห้อง 422 เลขที่ 83 หมู่ 11 ถนนสระบุรี – หล่มสัก ตำบลสะเดียง อำเภอเมือง จังหวัดเพชรบูรณ์ 67000 โทรศัพท์ (056)717100 ต่อ 1407 โทรสาร (056)717110 อินเตอร์เน็ต http://www.pcru.ac.th โทรศัพท์มือถือ 08-6529-6629

E-mail : kanchana111@hotmail.com , kanchana111@gmail.com

5. ประวัติการศึกษา

ปริญญาโท จากสถาบันบัณฑิตพัฒนบริหารศาสตร์ วิทยาศาสตรมหาบัณฑิต สาขาสถิติ

ปริญญาโท จากมหาวิทยาลัยรามคำแหง ศึกษาศาสตรมหาบัณฑิต สาขาวิจัยการศึกษา(กำลังทำวิทยานิพนธ์)

ปริญญาตรี จากมหาวิทยาลัยราชภัฏพิบูลสงคราม วิทยาศาสตรบัณฑิต สาขาสถิติประยุกต์

6. สาขาวิชาการที่มีความชำนาญพิเศษ (แตกต่างจากวุฒิการศึกษา) ระบุ คณิตศาสตร์ วัดและประเมินผล หลักสูตรและการสอน และประกันคุณภาพการศึกษา

7. ประสบการณ์ที่เกี่ยวข้องกับการบริหารงานวิจัย ทั้งภายในและภายนอกประเทศ โดยระบุสถานภาพในการทำการวิจัยว่าเป็นผู้อำนวยการแผนงานวิจัย หัวหน้าโครงการวิจัย หรือผู้ร่วมวิจัยในแต่ละ ข้อเสนอการวิจัย เป็นต้น

7.1 ผู้อำนวยการแผนงานวิจัย : -

7.2 หัวหน้าโครงการวิจัย : -

7.3 งานวิจัยที่ทำเสร็จแล้ว :

7.3.1 การประมาณค่าพารามิเตอร์ในตัวแบบการถดถอยโลจิสติก

7.3.2 การประเมินผลหลักสูตรประกาศนียบัตรบัณฑิต สาขาวิชาชีพครูคณะครุศาสตร์อุตสาหกรรม มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร 2553

7.3.3 การประเมินโครงการบริการวิชาการ เรื่อง การพัฒนาศักยภาพบุคลากรการสอน ในกลุ่มสาระการเรียนรู้การงานอาชีพและเทคโนโลยี วิชางานบ้าน วิชางานเกษตร วิชางานประดิษฐ์ วิชางานธุรกิจ วิชางานช่าง และวิชาเทคโนโลยี 2553

7.3.4 การประเมินโครงการรณรงค์เสริมสร้างวินัยและจิตสำนึกของประชาชนเพื่อลดอุบัติเหตุจราจร 2554

7.3.5 การประเมินศักยภาพและความต้องการใช้งานหุ่นยนต์ทางการแพทย์เพื่อช่วยเหลือผู้ป่วยของสถานประกอบการในประเทศไทย 2555

สารบัญ

หน้า

บทที่	หน้า

สารบัญตาราง

สารบัญภาพ

รูปที่ **หน้า**