การเปรียบเทียบประสิทธิภาพการประมาณค่าพารามิเตอร์ของ
การแจกแจงทวินามเชิงลบ ด้วยวิธีภาวะน่าจะเป็นสูงสุด วิธีของเบส์
และวิธีมาร์คอฟ เชน มอนติคาร์โล

EFFICIENCY COMPARISON OF PARAMETER ESTIMATIONS OF NEGATIVE BINOMIAL DISTRIBUTION BY MAXIMUM LIKELIHOOD, BAYES', AND MARKOV CHAIN MONTE CARLO METHODS

อัญมณี กุมมาระกะ

UNYAMANEE KUMMARAKA

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาสถิติประยุกต์
ภาควิชาสถิติ คณะวิทยาศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
พ.ศ. 2560

KMITL-2017-SC-M-050-010

การเปรียบเทียบประสิทธิภาพการประมาณค่าพารามิเตอร์ของ
การแจกแจงทวินามเชิงลบ ด้วยวิธีภาวะน่าจะเป็นสูงสุด วิธีของเบส์
และวิธีมาร์คอฟ เชน มอนติคาร์โล

EFFICIENCY COMPARISON OF PARAMETER ESTIMATIONS OF
NEGATIVE BINOMIAL DISTRIBUTION BY MAXIMUM LIKELIHOOD,
BAYES', AND MARKOV CHAIN MONTE CARLO METHODS

อัญมณี กุมมาระกะ
UNYAMANEE KUMMARAKA

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาสถิติประยุกต์
ภาควิชาสถิติ คณะวิทยาศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
พ.ศ. 2560

KMITL-2017-SC-M-050-010

EFFICIENCY COMPARISON OF PARAMETER ESTIMATIONS OF NEGATIVE BINOMIAL DISTRIBUTION BY MAXIMUM LIKELIHOOD, BAYES', AND MARKOV CHAIN MONTE CARLO METHODS

UNYAMANEE KUMMARAKA

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE DEGREE OF MASTER IN APPLIED STATISTICS

DEPARTMENT OF STATISTICS

FACULTY OF SCIENCE

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

2017

KMITL-2017-SC-M-050-010

คณะวิทยาศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ใบรับรองวิทยานิพนธ์

หัวข้อวิทยานิพนธ์ “การเปรียบเทียบประสิทธิภาพการประมาณค่าพารามิเตอร์ของการแจกแจงทวินามเชิงลบ ด้วยวิธีภาวะน่าจะเป็นสูงสุด วิธีของเบส์ และวิธีมาร์คอฟ เชน มอนติคาร์โล” (EFFICIENCY COMPARISON OF PARAMETER ESTIMATIONS OF NEGATIVE BINOMIAL DISTRIBUTION BY MAXIMUM LIKELIHOOD, BAYES’, AND MARKOV CHAIN MONTE CARLO METHODS)

ชื่อนักศึกษา นางสาวอัญมณี กุมมาระกะ
รหัสประจำตัว 58605088
ปริญญา วิทยาศาสตรมหาบัณฑิต (สาขาสถิติประยุกต์)
ภาควิชา สถิติ
อาจารย์ที่ปรึกษาวิทยานิพนธ์ ผศ.ดร.อัชฌา อระวีพร

คณะกรรมการสอบวิทยานิพนธ์	ลายมือชื่อ
ผศ.ดร.น้อมจิต กิตติโชติพาณิชย์ ประธานกรรมการ	
ผศ.ดร.สมศรี บัณฑิตวิไล อาจารย์บัณฑิตประจำ (ในสาขาวิชาที่เกี่ยวข้อง)	
รศ.ดร.วราฤทธิ์ พานิชกิจโกศลกุล ผู้ทรงคุณวุฒิจากภายนอกสถาบันฯ	
ผศ.ดร.อัชฌา อระวีพร อาจารย์ที่ปรึกษาวิทยานิพนธ์	

วัน/ เดือน/ ปี ที่สอบ 13 มิถุนายน พ.ศ. 2560 เวลา 09.00 – 10.30 น.
สถานที่สอบ ณ ห้อง 115 ตึกจุฬาภรณ์ 1

คณะวิทยาศาสตร์รับรองแล้ว

(รองศาสตราจารย์ ดร.ดุษณี ธนะบริพัฒน์)
คณบดีคณะวิทยาศาสตร์
วันที่...11...เดือน...กค...พ.ศ...60...

หัวข้อวิทยานิพนธ์	การเปรียบเทียบประสิทธิภาพการประมาณ ค่าพารามิเตอร์ของการแจกแจงทวินามเชิงลบ ด้วยวิธีภาวะน่าจะเป็นสูงสุด วิธีของเบส์ และวิธีมาร์คอฟ เชน มอนติคาร์โล
ชื่อนักศึกษา	นางสาวอัญมณี กุมมาระกะ
รหัสประจำตัว	58605088
ปริญญา	วิทยาศาสตรมหาบัณฑิต
ภาควิชา	สถิติ
พ.ศ.	2560
อาจารย์ที่ปรึกษาวิทยานิพนธ์	ผศ.ดร.อัชฌา อระวีพร

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพการประมาณค่าพารามิเตอร์ของประชากร (p) หรือค่าความน่าจะเป็นที่จะได้รับความสำเร็จในการทดลองแต่ละครั้งของการแจกแจงทวินามเชิงลบ แบบจุดและแบบช่วง ด้วยวิธีภาวะน่าจะเป็นสูงสุด วิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา และวิธีมาร์คอฟ เชน มอนติคาร์โล โดยใช้ตัวอย่างสุ่มจากการแจกแจงทวินามเชิงลบที่มีค่าพารามิเตอร์ของประชากร (p) เท่ากับ 0.2 0.5 และ 0.8 และค่าพารามิเตอร์ r หรือจำนวนครั้งของความสำเร็จ สำหรับการประมาณค่าแบบจุด ขนาดตัวอย่าง 10 (r = 3 5 และ 7) ขนาดตัวอย่าง 30 (r = 10 15 และ 20) และขนาดตัวอย่าง 50 (r = 10 20 และ 30) สำหรับการประมาณค่าแบบช่วง ขนาดตัวอย่าง 30 (r = 10 15 และ 20) ขนาดตัวอย่าง 50 (r = 10 20 และ 30) และขนาดตัวอย่าง 70 (r = 15 30 และ 45) และกำหนดระดับความเชื่อมั่น 90% 95% และ 99% โดยใช้โปรแกรมอาร์ และทำซ้ำ 1,000 รอบในแต่ละสถานการณ์ โดยการประมาณค่าแบบจุดจะพิจารณาจากการทดสอบสมมติฐานการเท่ากันระหว่างค่าเฉลี่ยความน่าจะเป็นของความสำเร็จในแต่ละครั้งของการทดลองของตัวอย่างกับค่าพารามิเตอร์ของประชากร โดยใช้การทดสอบที สำหรับการประมาณค่าแบบช่วงจะพิจารณาจากค่าสัมประสิทธิ์ความเชื่อมั่นและค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่น

สำหรับการประมาณค่าแบบจุด พบว่าส่วนใหญ่วิธีภาวะน่าจะเป็นสูงสุดให้ค่าประมาณไม่แตกต่างจากค่าพารามิเตอร์ของประชากร ยกเว้นเมื่อค่าพารามิเตอร์ของประชากรเท่ากับ 0.8 และขนาดตัวอย่างมีค่า 30 ขึ้นไป สำหรับวิธีของเบส์ และวิธีมาร์คอฟ เชน มอนติคาร์โลส่วนใหญ่ให้ค่าประมาณไม่แตกต่างจากค่าพารามิเตอร์ของประชากร เมื่อค่าพารามิเตอร์ของประชากรเท่ากับ 0.5 และ 0.8 โดยเฉพาะเมื่อขนาดตัวอย่างมีค่า 30 ขึ้นไป สำหรับการประมาณค่าแบบช่วงในสถานการณ์ส่วนใหญ่ วิธีของเบส์จะให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นแคบที่สุด

คำสำคัญ : การแจกแจงทวินามเชิงลบ ภาวะน่าจะเป็นสูงสุด เบส์ มาร์คอฟ เชน มอนติคาร์โล

Thesis Title	Efficiency Comparison of Parameter Estimations of Negative Binomial Distribution by Maximum Likelihood, Bayes', and Markov Chain Monte Carlo Methods
Student Name	Unyamanee Kummaraka
Student ID	58605088
Degree	Master of Science
Department	Statistics
Year	2017
Thesis Advisor	Asst. Prof. Dr. Autcha Araveeporn

Abstract

The objective of this research is to compare the point and interval estimations of the population parameter (p) or probability of success in each experiment based on negative binomial distribution by maximum likelihood, Bayes' and Markov Chain Monte Carlo methods. The prior distribution of Bayes' method is defined by beta distribution. The data samples are generated from negative binomial distribution with the population parameter (p) as 0.2, 0.5 and 0.8, and the parameter r or called number of success. For point estimation, the sample sizes is set as n =10 (r =3, 5 and 7), n =30 (r =10, 15 and 20), and n =50 (r =10, 20 and 30). For interval estimation, the sample sizes is set as n =30 (r =10, 15 and 20), n =50 (r =10, 20 and 30), and n =70 (r =15, 30 and 45), and 90%, 95% and 99% confidence interval. The simulated data is generated by R program and replicated 1,000 times for each situation. The criterion of point estimation is the process of hypothesis test that the estimators from simulation data are not different from the population parameters by using t-test. The interval estimation is considered by a confidence coefficient and average interval length.

For the point estimation, the almost estimators from the maximum likelihood method is equal to population parameter except at p = 0.8 and sample sizes greater than or equal to 30. The Bayes' and Markov Chain Monte Carlo methods produce dissimilar estimation to the population parameters at population p = 0.5 and 0.8, especially sample sizes greater than or equal to 30. For the interval estimation, the Bayes' method performs very satisfactorily in most all cases when the average interval length is shown minimum value.

Keywords : Negative Binomial Distribution, Maximum Likelihood, Bayes', Markov Chain Monte Carlo

กิตติกรรมประกาศ

วิทยานิพนธ์เล่มนี้สำเร็จลุล่วงไปได้ด้วยดีและมีความถูกต้องในเนื้อหา เนื่องด้วยได้รับความอนุเคราะห์จาก ผศ.ดร.อัชฌา อระวีพร ผู้ซึ่งเป็นอาจารย์ที่ปรึกษา รวมถึงเป็นผู้ให้คำแนะนำ คำปรึกษา เอื้อเฟื้อ เอกสารที่เกี่ยวข้องต่างๆ และหนังสืออ้างอิงที่ใช้ในการจัดทำงานวิจัยนี้ ตลอดจนตรวจทานงานวิจัย และติดตามการทำวิจัยทุกขั้นตอนจนกระทั่งเสร็จสมบูรณ์ จึงขอกราบขอบพระคุณด้วยความเคารพอย่างสูง ไว้ ณ ที่นี้ด้วย

ขอขอบพระคุณ คณะกรรมการพิจารณาหัวข้อและเค้าโครงวิทยานิพนธ์ ได้แก่ ผศ.ดร.น้อมจิต กิตติโชติพานิชย์ ผู้ซึ่งเป็นประธานกรรมการ ผศ.ดร.สมศรี บัณฑิตวิไล ผู้ซึ่งเป็นกรรมการ และ รศ.ดร.วราฤทธิ์ พานิชกิจโกศลกุล ผู้ซึ่งเป็นกรรมการ (ผู้ทรงคุณวุฒิจากภายนอก) ที่กรุณาให้คำปรึกษา คำแนะนำ ชี้ให้เห็นถึงข้อบกพร่อง ตลอดจนแก้ไขข้อผิดพลาดเพิ่มเติม ทำให้วิทยานิพนธ์เล่มนี้มีความสมบูรณ์ยิ่งขึ้น

ขอขอบพระคุณ คณาจารย์และบุคลากรภาควิชาสถิติทุกท่าน ที่ได้มอบความรู้และคำแนะนำ รวมถึงความช่วยเหลือในเรื่องต่างๆ มาโดยตลอด

สุดท้ายนี้ ขอขอบพระคุณบิดามารดา และครอบครัวของข้าพเจ้า ที่ให้การสนับสนุนและเป็นกำลังใจให้เสมอมา และขอขอบคุณพี่ๆ เพื่อนๆ น้องๆ ทุกคนที่ให้คำปรึกษา ช่วยเหลือในการศึกษามาโดยตลอด จนกระทั่งวิทยานิพนธ์เล่มนี้สำเร็จไปได้ด้วยดี

นางสาวอัญมณี กุมมาระกะ

สารบัญ

สารบัญ (ต่อ)

หน้า

สารบัญตาราง

สารบัญรูป

บทที่ 1
บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

งานวิจัยส่วนมากมีจุดมุ่งหมายเพื่อศึกษาลักษณะที่สนใจในประชากร (Population) ซึ่งประชากรมักจะมีขนาดใหญ่ แต่เนื่องจากมีข้อจำกัดด้านค่าใช้จ่าย เวลา และทรัพยากรบุคคล ทำให้การศึกษาข้อมูลทั้งประชากรเป็นไปได้ยาก จึงทำการศึกษาจากกลุ่มตัวอย่าง (Sample) ที่ถูกเลือกมาอย่างสุ่มจากประชากร และนำค่าที่ได้จากกลุ่มตัวอย่างอนุมานไปยังประชากร ซึ่งเรียกว่า การอนุมานเชิงสถิติ (Statistical Inference) โดยลักษณะของประชากรที่สนใจ เรียกว่า ค่าพารามิเตอร์ (Parameter) ส่วนค่าที่ได้จากกลุ่มตัวอย่าง เรียกว่า ค่าสถิติ (Statistic)

การอนุมานเชิงสถิติ แบ่งออกได้เป็น 2 รูปแบบ ตามวัตถุประสงค์ของงานวิจัย คือ การประมาณค่าพารามิเตอร์ (Parameter Estimation) และการทดสอบสมมติฐาน (Hypothesis Testing)

สำหรับการประมาณค่าพารามิเตอร์ แบ่งได้เป็น 2 ประเภท คือ การประมาณค่าแบบจุด (Point Estimation) และการประมาณค่าแบบช่วง (Interval Estimation) สำหรับการประมาณค่าแบบจุด เป็นการประมาณค่าพารามิเตอร์ด้วยค่าสถิติเพียงค่าเดียว การประมาณค่าแบบจุดจึงอาจจะเกิดความคลาดเคลื่อนได้ ถ้ากลุ่มตัวอย่างที่สุ่มมาไม่เป็นตัวแทนที่ดีของประชากร ส่วนการประมาณค่าแบบช่วง เป็นการประมาณค่าขอบเขตความเป็นไปได้ของค่าพารามิเตอร์ โดยที่มีคุณสมบัติว่าค่าพารามิเตอร์จะอยู่ในขอบเขตหรือช่วงนั้นด้วยระดับความเชื่อมั่นที่กำหนด

การประมาณค่าแบบจุด มีวิธีการหาตัวประมาณค่าพารามิเตอร์ได้หลายวิธี เช่น วิธีโมเมนต์ (The Moments Method) วิธีภาวะน่าจะเป็นสูงสุด (The Maximum Likelihood Method) วิธีไคกำลังสองต่ำสุด (The Minimum Chi-Square Method) วิธีกำลังสองต่ำสุด (The Least Squares Method) วิธีของเบส์ (Bayes' Method) เป็นต้น โดยที่ในการประมาณค่าพารามิเตอร์ค่าหนึ่งอาจจะมีตัวประมาณค่าแบบจุดได้มากกว่าหนึ่งตัว จึงต้องทำการพิจารณาความเหมาะสมของค่าประมาณดังกล่าว โดยอาศัยคุณสมบัติของตัวประมาณ ได้แก่ ความไม่เอนเอียง (Unbiasedness) ความคงเส้นคงวา (Consistency) ความพอเพียง (Sufficiency) ความแปรปรวนต่ำสุด (Minimum Variance) ประสิทธิภาพ (Efficiency) ความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำสุด (Minimum Mean Squared Error) และความไม่แปรเปลี่ยน (Invariance) (ประชุม สุวัตถี, 2553)

การประมาณค่าแบบช่วง จะอาศัยค่าประมาณแบบจุดและการแจกแจงความน่าจะเป็นของตัวประมาณในการสร้างขอบเขตบนและขอบเขตล่าง โดยการประมาณค่าแบบช่วงสามารถบอกถึงโอกาสหรือความน่าจะเป็นที่ค่าพารามิเตอร์จะตกอยู่ในช่วงประมาณดังกล่าว ในทางสถิติเรียกค่าความน่าจะเป็น นั้นว่า สัมประสิทธิ์ความเชื่อมั่น (Confidence Coefficient) แทนด้วย $1-\alpha$ โดยที่ $0<\alpha<1$ และเรียก α ว่า ระดับนัยสำคัญ (Level of Significance)

วิธีภาวะน่าจะเป็นสูงสุด เป็นวิธีการประมาณค่าที่ได้รับความนิยมใช้อย่างแพร่หลาย โดยการหาตัวประมาณค่าของพารามิเตอร์ที่ทำให้ฟังก์ชันภาวะน่าจะเป็น (Likelihood Function) มีค่าสูงสุด และตัวประมาณที่ได้จากวิธีนี้ยังสามารถตรวจสอบคุณสมบัติของตัวประมาณที่เหมาะสมได้หลายประการ

วิธีของเบส์ (Bayes' Method) เป็นการพิจาณารูปแบบการแจกแจงของค่าพารามิเตอร์ของฟังก์ชันการแจกแจงความน่าจะเป็น หรือที่เรียกว่า ฟังก์ชันการแจกแจงก่อน (Prior Distribution Function) ร่วมกับฟังก์ชันความหนาแน่นความน่าจะเป็น (Probability Density Function) เพื่อประมาณค่าพารามิเตอร์เมื่อกำหนดข้อมูล ซึ่งเรียกว่า ฟังก์ชันการแจกแจงภายหลัง (Posterior Distribution Function) ซึ่งฟังก์ชันการแจกแจงก่อนนี้จะถูกกำหนดโดยประสบการณ์ส่วนตัวของแต่ละบุคคล ซึ่งบอกถึงระดับความเชื่อของแต่ละบุคคล เมื่อได้ทำการทดลองก็จะมีข้อสนเทศใหม่ขึ้น จะส่งผลให้การแจกแจงความน่าจะเป็นเปลี่ยนแปลงไป (ฝันจิต แต้มทอง, 2537)

Mir (2008) ได้ทำการศึกษาการประมาณค่าพารามิเตอร์แบบจุดของการแจกแจงทวินามเชิงลบทั่วไป การแจกแจงทวินามเชิงลบ และการแจกแจงทวินาม โดยวิธีของเบส์ ที่มีฟังก์ชันการแจกแจงก่อนเป็นการแจกแจงบีตา Ganji. *et al*. (2013) ได้ทำการศึกษาการประมาณค่าพารามิเตอร์ r แบบจุดของการแจกแจงทวินามเชิงลบทั่วไป โดยวิธีของเบส์ และวิธีของเบส์เชิงประจักษ์ ทั้งกรณีที่ทราบและไม่ทราบค่าพารามิเตอร์ p กรณีที่ไม่ทราบค่า p ใช้วิธีของเบส์ ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา

วิธีมาร์คอฟ เชน มอนติคาร์โล (Markov Chain Monte Carlo) เป็นวิธีที่นิยมใช้ เมื่อไม่ทราบค่าพารามิเตอร์ของฟังก์ชันการแจกแจงก่อนของตัวแปรสุ่ม โดยประกอบด้วยการสุ่มตัวอย่างตัวแปรจากวิธีมาร์คอฟ เชน (Markov Chain) จากฟังก์ชันการแจกแจงก่อน และทำการสุ่มตัวอย่างแบบกิบส์ (Gibbs Sampling) จากนั้นทำการประมาณค่าด้วยมอนติคาร์โล (Monte Carlo) จากฟังก์ชันการแจกแจงภายหลัง (อัชฌา อระวีพร, 2555)

Kumar and Tomer (2016) ได้ทำการศึกษาเปรียบเทียบวิธีการประมาณค่าความเชื่อถือได้ของการแจกแจงทวินามเชิงลบ โดยใช้ข้อมูล Type-II Censoring โดยวิธีภาวะน่าจะเป็นสูงสุด วิธีของเบส์ และวิธีมาร์คอฟ เชน มอนติคาร์โล ที่ประยุกต์ใช้กับวิธีของเบส์ Araveeporn (2014) ได้ทำการเปรียบเทียบวิธีประมาณค่าพารามิเตอร์แบบจุดของการแจกแจงปัวซง โดยวิธีภาวะน่าจะเป็นสูงสุด วิธีของเบส์ และวิธีมาร์คอฟ เชน มอนติคาร์โล ที่ประยุกต์ใช้กับวิธีของเบส์ โดยนำเทคนิคดังกล่าวมาประมาณค่าพารามิเตอร์ของการแจกแจงก่อน Araveeporn (2015) ได้ทำการเปรียบเทียบวิธีประมาณค่าพารามิเตอร์แบบช่วงของการแจกแจงปัวซง โดยวิธีภาวะน่าจะเป็นสูงสุด วิธีของเบส์ และวิธีมาร์คอฟ เชน มอนติคาร์โล ที่ประยุกต์ใช้กับวิธีของเบส์ โดยนำเทคนิคดังกล่าวมาประมาณค่าพารามิเตอร์ของการแจกแจงก่อน Thetkham and Araveeporn (2016) ได้ทำการเปรียบเทียบวิธีประมาณค่าพารามิเตอร์แบบจุดของการแจกแจงแบบเลขชี้กำลัง โดยวิธีภาวะน่าจะเป็นสูงสุด วิธีของเบส์ และวิธีมาร์คอฟ เชน มอนติคาร์โล ที่ประยุกต์ใช้กับวิธีของเบส์ โดยนำเทคนิคดังกล่าวมา

ประมาณค่าพารามิเตอร์ของการแจกแจงก่อน Araveeporn (2015) ได้ทำการเปรียบเทียบวิธีประมาณค่าเฉลี่ยแบบช่วงของการแจกแจงปกติ โดยวิธีมาร์คอฟ เชน มอนติคาร์โล กับวิธีภาวะน่าจะเป็นสูงสุด วิธีของเบส์ และวิธีมาร์คอฟ เชน มอนติคาร์โล ที่ประยุกต์ใช้กับวิธีของเบส์โดยนำเทคนิคดังกล่าวมาประมาณค่าพารามิเตอร์ของการแจกแจงก่อน

ข้อมูลโดยทั่วไปจะมีลักษณะเป็นตัวแปรสุ่ม (Random Variable) โดยตัวแปรสุ่มสามารถแบ่งออกเป็น 2 ประเภท คือ ตัวแปรสุ่มแบบไม่ต่อเนื่อง (Discrete Random Variable) และตัวแปรสุ่มแบบต่อเนื่อง (Continuous Random Variable) โดยตัวแปรสุ่มแบบไม่ต่อเนื่องจะมีการแจกแจงที่สำคัญ เช่น การแจกแจงแบร์นูลลี (The Bernoulli Distribution) และการแจกแจงทวินามเชิงลบ (Negative Binomial Distribution) เป็นต้น และการแจกแจงแบบต่อเนื่องที่สำคัญ เช่น การแจกแจงปรกติ (Normal Distribution) และการแจกแจงบีตา (Beta Distribution) เป็นต้น

สำหรับงานวิจัยนี้ผู้วิจัยสนใจศึกษาข้อมูลที่มีการแจกแจงทวินามเชิงลบ โดยจะสนใจตัวแปรสุ่ม X ที่แทนจำนวนครั้งของการพบสิ่งที่ไม่สนใจในการทดลองแบร์นูลลี ซึ่งดำเนินการซ้ำๆ กันและเป็นอิสระกันจนกระทั่งได้ความสำเร็จครบ r ครั้ง ที่มีพารามิเตอร์คือ r และ p เมื่อ r คือ จำนวนครั้งของสิ่งที่สนใจ (ผลสำเร็จ) p คือ ความน่าจะเป็นที่จะได้รับสิ่งที่สนใจในการทดลองแต่ละครั้ง สำหรับการแจกแจงทวินามเชิงลบสามารถนำไปประยุกต์ใช้ได้ในด้านวิทยาศาสตร์ ด้านอุตสาหกรรม ด้านการแพทย์ เป็นต้น ซึ่งการประมาณค่าพารามิเตอร์แบบช่วงที่ได้นั้นจะมีความแม่นยำเพียงใด ขึ้นอยู่กับวิธีการประมาณและปัจจัยอื่นๆ เช่น ขนาดของตัวอย่าง ค่าของพารามิเตอร์ และการกำหนดช่วงความเชื่อมั่น

จากแนวคิดดังกล่าวผู้วิจัยมีความสนใจที่จะศึกษาการประมาณค่าพารามิเตอร์ p แบบจุดและแบบช่วงของการแจกแจงทวินามเชิงลบด้วยวิธีภาวะน่าจะเป็นสูงสุด วิธีของเบส์ และวิธีมาร์คอฟ เชน มอนติคาร์โล โดยการจำลองข้อมูลตามสถานการณ์ต่างๆ และเปรียบเทียบการประมาณค่าพารามิเตอร์ทั้ง 3 วิธี

1.2 วัตถุประสงค์

1.2.1. เพื่อศึกษาวิธีการประมาณค่า สำหรับพารามิเตอร์ p ของการแจกแจงทวินามเชิงลบด้วยวิธีภาวะน่าจะเป็นสูงสุด วิธีของเบส์ และวิธีมาร์คอฟ เชน มอนติคาร์โล

1.2.2. เพื่อเปรียบเทียบการประมาณค่าพารามิเตอร์และช่วงความเชื่อมั่น สำหรับพารามิเตอร์ p ของการแจกแจงทวินามเชิงลบ ด้วยวิธีภาวะน่าจะเป็นสูงสุด วิธีของเบส์ และวิธีมาร์คอฟ เชน มอนติคาร์โล

1.3 สมมติฐานการวิจัย

ตัวประมาณค่าแบบจุดและแบบช่วงที่ได้จากวิธีภาวะน่าจะเป็นสูงสุด วิธีของเบส์ และวิธีมาร์คอฟ เชน มอนติคาร์โล จะมีประสิทธิภาพสูงสุด ภายใต้สถานการณ์ที่ขนาดตัวอย่าง ระดับนัยสำคัญ และค่าพารามิเตอร์แตกต่างกัน

1.4 ขอบเขตของการวิจัย

1.4.1 ศึกษาการประมาณค่าพารามิเตอร์ ด้วยวิธีการประมาณค่า 3 วิธี คือ วิธีภาวะน่าจะเป็นสูงสุด วิธีของเบส์ และวิธีมาร์คอฟ เชน มอนติคาร์โล

1.4.2 กำหนดให้ตัวแปรสุ่ม X มีการแจกแจงทวินามเชิงลบ โดยที่ X แทนจำนวนครั้งของการพบสิ่งที่ไม่สนใจในการทดลองแบร์นูลลี ซึ่งดำเนินการซ้ำๆ กันและเป็นอิสระกันจนกระทั่งได้ความสำเร็จครบ r ครั้ง โดยมีพารามิเตอร์เป็น r และ p เมื่อ r คือ จำนวนครั้งของสิ่งที่สนใจ (ผลสำเร็จ) และ p คือ ความน่าจะเป็นที่จะได้รับสิ่งที่สนใจในการทดลองแต่ละครั้ง โดยมีฟังก์ชันการแจกแจงความน่าจะเป็น ดังนี้

$$f(x;r,p)=\begin{cases}\binom{x+r-1}{x}p^r(1-p)^x & ,x=0,1,2,\ldots\\ 0 & ,x\text{ มีค่าอื่นๆ}\end{cases}$$

โดยที่ ค่าคาดหมายของตัวแปรสุ่ม X จะเท่ากับ $E(X)=\dfrac{r(1-p)}{p}$

ค่าความแปรปรวนตัวแปรสุ่ม X จะเท่ากับ $Var(X)=\dfrac{r(1-p)}{p^2}$

1.4.3 กำหนดฟังก์ชันการแจกแจงก่อน ให้มีการแจกแจงบีตา ที่มีพารามิเตอร์ a และ b โดยมีฟังก์ชันความหนาแน่นความน่าจะเป็น ดังนี้

$$f(p;a,b)=\begin{cases}\dfrac{\Gamma(a+b)\,p^{a-1}(1-p)^{b-1}}{\Gamma(a)\Gamma(b)} & ,0<p<1,a>0,b>0\\ 0 & ,p\text{ มีค่าอื่นๆ}\end{cases}$$

โดยที่ ค่าคาดหมายของตัวแปรสุ่ม P จะเท่ากับ $E(P)=\dfrac{a}{a+b}$

ค่าความแปรปรวนตัวแปรสุ่ม P จะเท่ากับ $Var(P)=\dfrac{ab}{(a+b)^2(a+b+1)}$

และกำหนดค่าพารามิเตอร์ให้มีรูปแบบสมมาตร เบ้ขวา และเบ้ซ้าย โดยมีค่าพารามิเตอร์ (a,b) เท่ากับ (3,3) (2,6) (6,2) ตามลำดับ

1.4.4 กำหนดค่าพารามิเตอร์ของประชากร (p) เท่ากับ 0.2 0.5 และ 0.8

1.4.5 กำหนดค่าพารามิเตอร์ r และขนาดตัวอย่าง (n) สำหรับการประมาณค่าแบบจุด แสดงดังตารางที่ 1.1

ตารางที่ 1.1 ค่าพารามิเตอร์สำหรับการประมาณค่าแบบจุด

n	r
10	3 5 7
30	10 15 20
50	10 20 30

1.4.6 กำหนดค่าพารามิเตอร์ r และขนาดตัวอย่าง (n) สำหรับการประมาณค่าแบบช่วง แสดงดังตารางที่ 1.2

ตารางที่ 1.2 ค่าพารามิเตอร์สำหรับการประมาณค่าแบบช่วง

n	r
30	10 15 20
50	10 20 30
70	15 30 45

1.4.7 กำหนดระดับความเชื่อมั่น 3 ระดับ คือ 90% 95% และ 99%

1.4.8 โปรแกรมที่ใช้ในการวิจัยครั้งนี้ทั้งหมดเขียนโดยโปรแกรมอาร์ เวอร์ชั่น 3.2.2 ซึ่งทำการทดลองซ้ำ 1,000 ครั้ง ในแต่ละสถานการณ์

1.5 เกณฑ์ที่ใช้ในการวิจัย

1.5.1 การประมาณค่าแบบจุด จะพิจารณาการประมาณค่าด้วยการทดสอบความแตกต่างระหว่างค่าเฉลี่ยความน่าจะเป็นของความสำเร็จในการทดลองแต่ละครั้งของตัวอย่างกับความน่าจะเป็นของความสำเร็จในการทดลองแต่ละครั้งของประชากร (p) โดยใช้การทดสอบที (t-test)

1.5.2 การประมาณค่าแบบช่วง จะทำการเปรียบเทียบค่าประมาณค่าสัมประสิทธิ์ความเชื่อมั่นกับค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด สำหรับงานวิจัยนี้ กำหนดให้เท่ากับ 90% 95% และ 99% และพิจารณาค่าประมาณค่าสัมประสิทธิ์ความเชื่อมั่นและความกว้างเฉลี่ยของช่วง โดยค่าประมาณสัมประสิทธิ์ความเชื่อมั่น คำนวณจาก

$$1-\hat{\alpha} = \frac{\text{จำนวนรอบที่ช่วงความเชื่อมั่นครอบคลุมค่าพารามิเตอร์}}{m}$$

$$\text{ความกว้างเฉลี่ยของช่วง} = \frac{\sum_{j=1}^{m}\left(U_j - L_j\right)}{m}$$

เมื่อ U_j แทน ขอบเขตบนของช่วงความเชื่อมั่น ในการทำซ้ำครั้งที่ j

L_j แทน ขอบเขตล่างของช่วงความเชื่อมั่น ในการทำซ้ำครั้งที่ j

m แทน จำนวนรอบของการทดลอง ในที่นี้เท่ากับ 1,000

1.6 ประโยชน์ที่คาดว่าจะได้รับ

1.6.1 สามารถเลือกใช้วิธีการประมาณค่าได้อย่างเหมาะสมในแต่ละสถานการณ์ของการแจกแจงทวินามเชิงลบ

1.6.2 เป็นแนวทางในการศึกษาวิจัยเพื่อเปรียบเทียบวิธีการประมาณค่า เมื่อการแจกแจงก่อนมีค่าพารามิเตอร์แตกต่างกัน ในการแจกแจงรูปแบบอื่นๆ

1.7 นิยามศัพท์

1.7.1 ช่วงความเชื่อมั่น (Confidence Interval) หมายถึง ช่วงที่ใช้ประมาณค่าพารามิเตอร์ซึ่งคำนวณได้จากค่าตัวอย่างสุ่ม

1.7.2 สัมประสิทธิ์ความเชื่อมั่น (Confidence Coefficient) เขียนแทนด้วย $(1-\alpha)$ หมายถึง ความน่าจะเป็นที่ช่วงความเชื่อมั่นจะครอบคลุมค่าพารามิเตอร์ที่กำหนด

1.7.3 ความกว้างเฉลี่ยของช่วง (Average Interval Length) หมายถึง ค่าเฉลี่ยของความกว้างของช่วงความเชื่อมั่นที่ได้จากการประมาณค่าตัวอย่างสุ่มในสถานการณ์เดียวกัน

1.7.4 วิธีภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Method) หมายถึง วิธีการประมาณค่าพารามิเตอร์ตัวหนึ่งหรือหลายตัวของประชากรที่ทำให้ฟังก์ชันภาวะน่าจะเป็นของตัวอย่างสุ่มมีค่าสูงสุด

1.7.5 ตัวประมาณภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Estimator) หมายถึง ตัวประมาณของพารามิเตอร์ตัวหนึ่งหรือหลายตัวของประชากรที่หาได้ด้วยวิธีภาวะน่าจะเป็นสูงสุด ที่ทำให้ฟังก์ชันภาวะน่าจะเป็นของตัวอย่างสุ่มมีค่าสูงสุด

1.7.6 ตัวประมาณเบส์ภายหลัง (Posterior Bayes' Estimator) ของ $h(p)$ หมายถึง ค่าคาดหมายของ $h(p)$ ที่ได้มาจากฟังก์ชันการแจกแจงภายหลัง

1.7.7 การแจกแจงก่อน (Prior Distribution) หมายถึง ฟังก์ชันความหนาแน่นน่าจะเป็นของตัวแปรสุ่ม Θ โดยที่ Θ คือพารามิเตอร์ของฟังก์ชันความหนาแน่นน่าจะเป็น

1.7.8 การแจกแจงภายหลัง (Posterior Distribution) ฟังก์ชันความหนาแน่นความน่าจะเป็นของ Θ เมื่อกำหนดให้ $X_1 = x_1, \ldots, X_n = x_n$ โดยที่ Θ คือพารามิเตอร์ของฟังก์ชันความหนาแน่นน่าจะเป็น

บทที่ 2
ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

การศึกษาครั้งนี้จะทำการประมาณค่าพารามิเตอร์แบบจุดและแบบช่วงของการแจกแจงทวินามเชิงลบ ด้วยวิธีภาวะน่าจะเป็นสูงสุด วิธีของเบส์ ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา และวิธีมาร์คอฟ เชน มอนติคาร์โล ซึ่งมีรายละเอียดของการแจกแจงทางสถิติ ทฤษฎีที่เกี่ยวข้อง และสถิติที่ใช้ในงานวิจัย ดังนี้

2.1 การแจกแจงที่ใช้ในงานวิจัย

2.1.1 การแจกแจงทวินามเชิงลบ (Negative Binomial Distribution)

กำหนดตัวแปรสุ่ม X ให้มีการแจกแจงทวินามเชิงลบ หรือเขียนได้ว่า $NB(r,p)$ โดยที่ X แทนจำนวนครั้งของการพบสิ่งที่ไม่สนใจในการทดลองแบร์นูลลี ซึ่งดำเนินการซ้ำๆ กันและเป็นอิสระกันจนกระทั่งได้ความสำเร็จครบ r ครั้ง โดยมีพารามิเตอร์เป็น r และ p เมื่อ r คือ จำนวนครั้งของสิ่งที่สนใจ (ผลสำเร็จ) และ p คือ ความน่าจะเป็นที่จะได้รับสิ่งที่สนใจในการทดลองแต่ละครั้ง (ปรางค์มาส ทรัพย์กระจ่างและยงยุทธ ไชยพงศ์, 2555)

ฟังก์ชันการแจกแจงความน่าจะเป็นของ X คือ

$$f(x;r,p)=\begin{cases}\binom{x+r-1}{x}p^r(1-p)^x & ,x=0,1,2,\ldots\\ 0 & ,x\text{ มีค่าอื่นๆ}\end{cases}$$

โดยที่ ค่าคาดหมายของตัวแปรสุ่ม X เท่ากับ $E(X)=\dfrac{r(1-p)}{p}$

ค่าความแปรปรวนของตัวแปรสุ่ม X เท่ากับ $Var(X)=\dfrac{r(1-p)}{p^2}$

2.1.2 การแจกแจงบีตา (Beta Distribution)

กำหนดตัวแปรสุ่ม X ให้มีการแจกแจงบีตา หรือเขียนได้ว่า $Beta(\alpha,\beta)$ ที่มีพารามิเตอร์เป็น α และ β

ฟังก์ชันความหนาแน่นความน่าจะเป็นของ X คือ

$$f(x;\alpha,\beta)=\begin{cases}\dfrac{\Gamma(\alpha+\beta)x^{\alpha-1}(1-x)^{\beta-1}}{\Gamma(\alpha)\Gamma(\beta)} & ,0<x<1,\alpha>0,\beta>0\\ 0 & ,x\text{ มีค่าอื่นๆ}\end{cases}$$

โดยที่ ค่าคาดหมายของตัวแปรสุ่ม X เท่ากับ $E(X)=\dfrac{\alpha}{\alpha+\beta}$

ค่าความแปรปรวนของตัวแปรสุ่ม X เท่ากับ $Var(X)=\dfrac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}$

2.2 ทฤษฎีที่ใช้ในงานวิจัย

2.2.1 การประมาณค่า (Estimation)

การประมาณค่า คือ การใช้ข้อมูลจากตัวอย่างสุ่ม ในรูปของค่าสถิติเพื่อประมาณหรือคาดหมายว่าพารามิเตอร์ θ ควรมีค่าเท่าใดหรืออยู่ในช่วงใด การประมาณค่าพารามิเตอร์มี 2 แนวทางด้วยกัน คือ การประมาณค่าแบบจุด และการประมาณค่าแบบช่วง

2.2.1.1 การประมาณค่าพารามิเตอร์แบบจุด (Point estimation)

การประมาณค่าแบบจุด หมายถึง การประมาณพารามิเตอร์ออกมาเป็นค่าเดียวหรือจุดเดียว

นิยามที่ 2.1 ให้ $X_1,\ldots,X_n$ เป็นตัวอย่างสุ่มจากประชากรที่มีฟังก์ชันความหนาแน่นความน่าจะเป็น $f(x;\theta),\theta\in\Omega$ ตัวประมาณ (Estimator) ของ θ คือ ฟังก์ชันของตัวอย่างสุ่ม $\hat{\theta}=\hat{\theta}(X_1,\ldots,X_n)$ ค่าประมาณ (Estimate) ของ θ คือ ค่าหนึ่งของตัวประมาณ $\hat{\theta}$ ของ θ

2.2.1.2 การประมาณค่าพารามิเตอร์แบบช่วง (Interval estimation)

การประมาณค่าแบบช่วง หมายถึง การประมาณพารามิเตอร์ θ ออกมาหลายค่า มักเป็นช่วงของจำนวนจริงหรือเซตของจุดหลายจุด เรียกว่า ช่วงความเชื่อมั่น (Confidence Interval) หรือเซตความเชื่อมั่น (Confidence Set) ช่วงความเชื่อมั่นหรือเซตความเชื่อมั่นดังกล่าวได้มาโดยอาศัยตัวประมาณ $\hat{\theta}=\hat{\theta}(X_1,\ldots,X_n)$ ของ θ และฟังก์ชันการแจกแจงความน่าจะเป็นของ $\hat{\theta}$ (น้อมจิต กิตติโชติพาณิชย์, 2553)

นิยามที่ 2.2 ให้ $X_1,\ldots,X_n$ เป็นตัวอย่างสุ่มจากประชากรที่มีฟังก์ชันความหนาแน่นความน่าจะเป็น $f(x;\theta)$ โดยที่ θ เป็นจำนวนจริง ให้ $L(X_1,\ldots,X_n)$ และ $U(X_1,\ldots,X_n)$ เป็นสถิติที่ $L(X_1,\ldots,X_n)\le U(X_1,\ldots,X_n)$ ทุกจุดสังเกต $x_1,\ldots,x_n$ ที่อยู่ในปริภูมิตัวอย่าง S และ $P[L(X_1,\ldots,X_n)\le g(\theta)\le U(X_1,\ldots,X_n)]=1-\alpha$ โดยที่ α ไม่ขึ้นอยู่กับ θ เราเรียกช่วงสุ่ม (Random Interval) $[L(X_1,\ldots,X_n),U(X_1,\ldots,X_n)]$ ว่า ช่วงความเชื่อมั่น $100(1-\alpha)\%$ ($100(1-\alpha)\%$ Confidence Interval) ของ $g(\theta)$ และเรียก $1-\alpha$ ว่า สัมประสิทธิ์ความเชื่อมั่น (Confidence Coefficient) หรือระดับความเชื่อมั่น (Confidence Level)

$L(X_1,\ldots,X_n)$ เป็นขีดจำกัดล่างของความเชื่อมั่น (Lower Confidence Limit) และ $U(X_1,\ldots,X_n)$ เป็นขีดจำกัดบนของความเชื่อมั่น (Upper Confidence Limit)

นิยามที่ 2.3 ถ้า $L(X_1,\ldots,X_n)$ และ $U(X_1,\ldots,X_n)$ เป็นสถิติที่

$$P[g(\theta)<L(X_1,\ldots,X_n)]=\frac{\alpha}{2}=P[g(\theta)>U(X_1,\ldots,X_n)]$$

แล้วเรียกช่วง $[L(X_1,\ldots,X_n),U(X_1,\ldots,X_n)]$ ว่า ช่วงความเชื่อมั่นศูนย์กลางขนาด $100(1-\alpha)\%$ ($100(1-\alpha)\%$ Central Confidence Interval) ของ $g(\theta)$(ประชุม สุวัตถี, 2553)

2.2.2 ทฤษฎีบทขีดจำกัดส่วนกลาง (Central Limit Throrem)

ให้ $X_1,\ldots,X_n$ เป็นตัวอย่างสุ่มจากประชากรที่มีการแจกแจงปรกติ โดยมีค่าเฉลี่ย μ และความแปรปรวน σ^2 จะได้ว่าตัวแปรสุ่ม $\sum_{i=1}^{n} X_i$ จะมีการแจกแจงปรกติ โดยมีค่าเฉลี่ย $n\mu$ และความแปรปรวน $n\sigma^2$ และ $\overline{X}$ จะมีการแจกแจงปรกติ โดยมีค่าเฉลี่ย μ และความแปรปรวน $\frac{\sigma^2}{n}$ แต่ในกรณีทั่วไป ตัวอย่างสุ่มที่ใช้ไม่จำเป็นต้องมาจากการแจกแจงปรกติ ในกรณีที่ตัวอย่างสุ่มมาจากการแจกแจงแบบอื่น การหาการแจกแจงความน่าจะเป็นของ $\overline{X}$ อาจทำได้ยาก โดยเฉพาะเมื่อ n มีค่ามาก จึงต้องใช้การแจกแจงโดยประมาณแทน ดังนั้น ถ้าสามารถหาได้ว่าการแจกแจงความน่าจะเป็นของ $\overline{X}$ เข้าใกล้การแจกแจงแบบใดเมื่อ $n \to \infty$ ก็ย่อมทำให้การอนุมานทำได้สะดวกยิ่งขึ้น

ทฤษฎีที่ 2.1 ทฤษฎีบทขีดจำกัดส่วนกลาง (Central Limit Throrem)

ให้ $X_1,\ldots,X_n$ เป็นตัวแปรสุ่มที่เป็นอิสระกันและมีการแจกแจงเหมือนกัน โดยมีค่าเฉลี่ยเป็น $E(X_i)=\mu$ และความแปรปรวนจำกัด (Finite Variance) เป็น $V(X_i)=\sigma^2$ ถ้าขนาดตัวอย่างมีขนาดใหญ่ จะได้ว่าตัวสถิติ $Y_n = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n\sigma^2}}$ หรือ $Y_n = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$ จะลู่เข้าในเชิงการแจกแจงสู่ตัวแปรสุ่ม Y ที่มีการแจกแจงปรกติ โดยมีค่าเฉลี่ย $\mu=0$ และความแปรปรวน $\sigma^2=1$ กล่าวคือ Y_n จะมีลิมิตการแจกแจงปรกติมาตรฐานโดยประมาณ (Asymptotically Standard Normal)

2.3 สถิติที่ใช้ในงานวิจัย

2.3.1 วิธีภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Method)

การประมาณพารามิเตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุดเป็นวิธีการที่ใช้กันมาก และเป็นวิธีที่ง่าย เพราะวิธีนี้เป็นการหาตัวประมาณค่าที่ทำให้ฟังก์ชันภาวะน่าจะเป็น (Likelihood Function) มีค่าสูงสุด ตัวประมาณค่าที่หาได้จากวิธีการนี้เรียกว่า ตัวประมาณภาวะน่าจะเป็นสูงสุด (น้อมจิต กิตติโชติพาณิชย์, 2553)

นิยามที่ 2.4 ให้ $X_1,\ldots,X_n$ เป็นตัวอย่างสุ่มจากประชากรที่มีฟังก์ชันความหนาแน่นความน่าจะเป็น $f(x;\theta),\theta\in\Omega$ ฟังก์ชันภาวะน่าจะเป็น (Likelihood Function) ของตัวอย่างสุ่ม คือ ฟังก์ชันความน่าจะเป็นร่วมของ $L=L(\theta;x_1,\ldots,x_n)=L(\theta)$ ของตัวอย่างสุ่มนั้น ถือว่าเป็นฟังก์ชันของพารามิเตอร์ θ นั่นคือ

$$\begin{aligned} L(\theta;x_1,\ldots,x_n) &= f(x_1,\ldots,x_n;\theta) \\ &= f(x_1;\theta)\ldots f(x_n;\theta) \\ &= \prod_{i=1}^{n} f(x_i;\theta) \end{aligned}$$

นิยามที่ 2.5 ตัวประมาณภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Estimator; MLE) ของพารามิเตอร์ θ คือค่าของ θ ที่ทำให้ฟังก์ชันภาวะน่าจะเป็น $L(\theta)$ มีค่าสูงสุด

วิธีหาตัวประมาณภาวะน่าจะเป็นสูงสุด

1. หาฟังก์ชันภาวะน่าจะเป็น $L(\theta)$

2. หา $\ln(L(\theta))$

3. หา $\dfrac{\partial \ln(L(\theta))}{\partial \theta} = 0$

4. แก้สมการในข้อ 3. หาค่า θ ที่ทำให้ $L(\theta)$ หรือ $\ln(L(\theta))$ มีค่าสูงสุด จะได้ $\hat{\theta}$ เป็นตัวประมาณภาวะน่าจะเป็นสูงสุด

ในกรณีที่หาอนุพันธ์ไม่ได้ให้เลือก θ ที่ทำให้ $L(\theta)$ หรือ $\ln(L(\theta))$ มีค่าสูงสุด จะได้ $\hat{\theta}$ เป็นตัวประมาณด้วยวิธีภาวะน่าจะเป็นสูงสุด (สายชล สินสมบูรณ์ทอง, 2548)

คุณสมบัติของตัวประมาณภาวะน่าจะเป็นสูงสุด

ตัวประมาณภาวะน่าจะเป็นสูงสุดมีสมบัติที่น่าสนใจหลายประการ ดังนี้

1) เป็นตัวประมาณคงเส้นคงวา (Consistent Estimator) ของ θ

2) เป็นตัวประมาณพอเพียง (Sufficient Estimator) ของ θ

3) ถ้าตัวประมาณไม่เอนเอียงที่มีความแปรปรวนต่ำสุด (MVUE) ของ θ หาค่าได้แล้วจะหาได้ด้วยวิธีภาวะน่าจะเป็นสูงสุด

4) ตัวประมาณภาวะน่าจะเป็นสูงสุดอาจมีได้มากกว่า 1 ตัว

5) ตัวประมาณภาวะน่าจะเป็นสูงสุดอาจจะเป็นตัวประมาณเอนเอียงหรือไม่เอนเอียงของ θ ก็ได้ ถ้า $\hat{\theta}$ เป็นตัวประมาณที่ไม่เอนเอียงของ θ แล้วจะมีคุณสมบัติเป็น MVUE ของ θ แต่ถ้า $\hat{\theta}$ เป็นตัวประมาณที่เอนเอียงของ θ แล้ว จะต้องปรับตัวประมาณภาวะน่าจะเป็นสูงสุดให้เป็นตัวประมาณที่ไม่เอนเอียงของ θ ก่อน จึงนำไปประมาณ θ ซึ่งตัวประมาณที่ปรับให้เป็นตัวประมาณไม่เอนเอียงแล้วจะเป็น MVUE ของ θ

6) ถ้า $\hat{\theta}$ เป็นตัวประมาณภาวะน่าจะเป็นของ θ และ $\hat{\theta}$ เป็นตัวประมาณที่ไม่เอนเอียงของ θ และถ้ามีตัวประมาณที่มีประสิทธิภาพสูงสุด (Efficient Estimator) ของ θ เกิดขึ้นแล้ว $\hat{\theta}$ จะเป็นตัวประมาณที่มีประสิทธิภาพสูงสุดของ θ

7) ถ้า $\hat{\theta}$ เป็นตัวประมาณภาวะน่าจะเป็นของ θ แล้วจะมีคุณสมบัติภาวะไร้การแปร (Invariance Property)

8) ถ้า $\hat{\theta}$ เป็นตัวประมาณภาวะน่าจะเป็นของ θ และ n มีค่ามากแล้ว

$$\hat{\theta} \to N\left(\theta, Var\left(\hat{\theta}\right)\right)$$

โดยที่ $Var\left(\hat{\theta}\right)$ คือ ขอบเขตล่างของอสมการคราแมร์ราว นั่นคือ

$$Var\left(\hat{\theta}\right) = \frac{1}{I(\theta)} = \frac{1}{E\left(\dfrac{\partial \ln L(\theta)}{\partial \theta}\right)^2} = \frac{1}{-E\left(\dfrac{\partial^2 \ln L(\theta)}{\partial \theta^2}\right)}$$

ให้ $X_1, \ldots, X_n$ เป็นตัวอย่างสุ่มจากประชากรที่มีการแจกแจงทวินามเชิงลบ ที่มีพารามิเตอร์ r, p จาก $X_i \sim NB(r, p)$

$$f(x; r, p) = \begin{cases} \dbinom{x+r-1}{x} p^r (1-p)^x & , x = 0, 1, 2, \ldots \\ 0 & , x \text{ มีค่าอื่นๆ} \end{cases}$$

หาตัวประมาณด้วยวิธีภาวะน่าจะเป็นสูงสุดของ p เมื่อทราบค่า r หาได้จาก

$$L(r,p) = \prod_{i=1}^{n} f(x_i; r, p)$$

$$= \prod_{i=1}^{n} \binom{x_i + r - 1}{x_i} p^{rn} (1-p)^{\sum_{i=1}^{n} x_i}$$

$$\ln L(r,p) = \ln \prod_{i=1}^{n} \binom{x_i + r - 1}{x_i} + rn \ln p + \sum_{i=1}^{n} x_i \ln(1-p)$$

$$\frac{\partial \ln L(r,p)}{\partial p} = \frac{rn}{p} - \frac{\sum_{i=1}^{n} x_i}{1-p} = 0$$

$$0 = rn(1-p) - p\left(\sum_{i=1}^{n} x_i\right)$$

$$0 = rn - rnp - p\sum_{i=1}^{n} x_i$$

$$rn = rnp + p\sum_{i=1}^{n} x_i$$

$$rn = p\left(rn + \sum_{i=1}^{n} x_i\right)$$

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

และ $\dfrac{\partial^2 \ln L(r,p)}{\partial p^2} = -\dfrac{rn}{p^2} - \dfrac{\sum_{i=1}^{n} x_i}{(1-p)^2} < 0$

ดังนั้น ตัวประมาณภาวะน่าจะเป็นสูงสุดของ p คือ $\hat{p}_{ML} = \dfrac{rn}{rn + \sum_{i=1}^{n} x_i}$

เมื่อพิจารณาตัวประมาณภาวะน่าจะเป็นสูงสุดของ p คือ $\hat{p}_{ML}$ ในรูปแบบการแจกแจงทวินามเชิงลบ มีคุณสมบัติของตัวประมาณดังต่อไปนี้

1. ความไม่เอนเอียง (Unbiasedness)

นิยามที่ 2.6 ตัวประมาณ $\hat{\theta} = t(x_1, \ldots, x_n)$ เป็นตัวประมาณไม่เอนเอียง (Unbiasedness Estimator) ของ θ ก็ต่อเมื่อ $E(\hat{\theta}) = \theta$

ให้ $X_1, \ldots, X_n$ เป็นตัวอย่างสุ่มจากประชากรที่มีการแจกแจงทวินามเชิงลบ ที่มีพารามิเตอร์ r, p จาก $X_i \sim NB(r,p)$ เมื่อ $i = 1, \ldots, n$, $0 < p < 1$ เมื่อทราบค่า r

เนื่องจาก $X_1, \ldots, X_n$ เป็นอิสระกัน จะได้ว่า $Y = \sum_{i=1}^{n} X_i \sim NB(rn, p)$

$$E\left(\frac{rn}{rn+Y}\right) = \sum_{y=0}^{\infty}\frac{rn}{rn+y}\binom{y+rn-1}{y}p^{rn}(1-p)^{y}$$

แต่ไม่สามารถประมาณค่าดังกล่าวได้ จึงทำการปรับตัวประมาณภาวะน่าจะเป็นสูงสุดให้อยู่ในรูปของ $\frac{rn-1}{rn+Y-1}$ (K. Krishnamoorthy, 2016)

$$E\left(\frac{rn-1}{rn+Y-1}\right) = \sum_{y=0}^{\infty}\frac{rn-1}{rn+y-1}\frac{(y+rn-1)!}{y!(rn-1)!}p^{rn}(1-p)^{y}$$

$$= \sum_{y=0}^{\infty}\frac{(y+rn-2)!}{y!(rn-2)!}p^{rn}(1-p)^{y}$$

$$= p\sum_{y=0}^{\infty}\binom{y+rn-2}{y}p^{rn-1}(1-p)^{y}$$

$$= p$$

พบว่า $\dfrac{rn-1}{rn+\sum_{i=1}^{n}X_i-1}$ เป็นตัวประมาณที่ไม่เอนเอียงของ p

ดังนั้นจะใช้ $\dfrac{rn-1}{rn+\sum_{i=1}^{n}X_i-1}$ เป็นตัวประมาณภาวะน่าจะเป็นสูงสุดของ p แทน $\dfrac{rn}{rn+\sum_{i=1}^{n}x_i}$

2. ความคงเส้นคงวา (Consistent)

นิยามที่ 2.7 ให้ $X_1,\ldots,X_n$ เป็นตัวอย่างสุ่มจากประชากรที่มีฟังก์ชันความหนาแน่นความน่าจะเป็น $f(x;\theta)$ $\hat{\theta}$ จะเป็นตัวประมาณที่คงเส้นคงวา (Consistent Estimator) ของ θ ก็ต่อเมื่อ $\lim_{n\to\infty}E(\hat{\theta})=\theta$ และ $\lim_{n\to\infty}Var(\hat{\theta})=0$

ให้ $X_1,\ldots,X_n$ เป็นตัวอย่างสุ่มจากประชากรที่มีการแจกแจงทวินามเชิงลบ ที่มีพารามิเตอร์ r,p จาก $X_i \sim NB(r,p)$ เมื่อ $i=1,\ldots,n$, $0<p<1$ เมื่อทราบค่า r

ค่าคาดหมายของตัวประมาณภาวะน่าจะเป็นสูงสุดของ p

$$E\left(\frac{rn-1}{rn+\sum_{i=1}^{n}X_i-1}\right) = p$$

ความแปรปรวนของตัวประมาณภาวะน่าจะเป็นสูงสุดของ p

$$Var\left(\hat{p}\right) \simeq \frac{p^2\left(1-p\right)}{2}\left(\frac{2\sum_{i=1}^{n} x_i + 2 - p}{\sum_{i=1}^{n} x_i\left(\sum_{i=1}^{n} x_i - p + 2\right)}\right)$$ (K. Krishnamoorthy, 2016)

จะได้ว่า $\lim_{n\to\infty} E\left(\hat{p}\right) = \lim_{n\to\infty} E\left(\frac{rn-1}{rn+\sum_{i=1}^{n} x_i - 1}\right) = \lim_{n\to\infty} p = p$

และ
$$\begin{aligned}\lim_{n\to\infty} Var\left(\hat{p}\right) &= \lim_{n\to\infty} Var\left(\frac{rn-1}{rn+\sum_{i=1}^{n} x_i - 1}\right)\\ &= \lim_{n\to\infty} \frac{p^2\left(1-p\right)}{2}\left(\frac{2\sum_{i=1}^{n} x_i + 2 - p}{\sum_{i=1}^{n} x_i\left(\sum_{i=1}^{n} x_i - p + 2\right)}\right)\\ &= \frac{p^2\left(1-p\right)}{2}\left(\frac{2\sum_{i=1}^{n} x_i + 2 - p}{\sum_{i=1}^{n} x_i\left(\sum_{i=1}^{n} x_i - p + 2\right)}\right) \neq 0\end{aligned}$$

ดังนั้น $\dfrac{rn-1}{rn+\sum_{i=1}^{n} X_i - 1}$ เป็นตัวประมาณที่ไม่คงเส้นคงวาของ p

3. ความพอเพียง (Sufficiency)

นิยามที่ 2.8 ให้ $X_1,\ldots,X_n$ เป็นตัวอย่างสุ่มจากประชากรที่มีฟังก์ชันความหนาแน่นความน่าจะเป็น $f\left(x;\theta\right)$ ถ้า $\hat{\theta} = t\left(x_1,\ldots,x_n\right)$ เป็นสถิติค่าหนึ่ง $\hat{\theta}$ จะเป็นสถิติพอเพียงของ θ ถ้า h, k เป็นฟังก์ชัน 2 ฟังก์ชัน ซึ่งทำให้

$$f\left(X_1,\ldots,X_n;\theta\right) = h\left(\hat{\theta};\theta\right) k\left(X_1,\ldots,X_n\right)$$

ให้ $X_1,\ldots,X_n$ เป็นตัวอย่างสุ่มจากประชากรที่มีการแจกแจงทวินามเชิงลบ ที่มีพารามิเตอร์ r, p จาก $X_i \sim NB\left(r,p\right)$ เมื่อ $i = 1,\ldots,n$, $0 < p < 1$ เมื่อทราบค่า r

ฟังก์ชันความน่าจะเป็นร่วมของ $X_1,\ldots,X_n$ คือ

$$f(X_1,\ldots,X_n;\theta) = \prod_{i=1}^{n}\binom{x_i+r-1}{x_i}p^{rn}(1-p)^{\sum_{i=1}^{n}x_i}$$

$$= k(X_1,\ldots,X_n)h(\hat{p};p)$$

โดยที่ $k(X_1,\ldots,X_n)=\prod_{i=1}^{n}\binom{x_i+r-1}{x_i}$ และ $h(\hat{p};p)=p^{rn}(1-p)^{\sum_{i=1}^{n}x_i}$

จะได้ว่า $\sum_{i=1}^{n}X_i$ เป็นสถิติพอเพียงของ p

เนื่องจาก $\dfrac{rn-1}{rn+\sum_{i=1}^{n}X_i-1}$ เป็นฟังก์ชันของ $\sum_{i=1}^{n}X_i$

ดังนั้น $\dfrac{rn-1}{rn+\sum_{i=1}^{n}X_i-1}$ จะเป็นสถิติพอเพียงของ p ด้วย

4. ความมีประสิทธิภาพ (Efficiency)

นิยามที่ 2.9 ถ้า $\hat{\theta}$ เป็นตัวประมาณไม่เอนเอียงของ θ แล้ว ประสิทธิภาพของ θ คือ

$$e(\hat{\theta})=\frac{1/I(\theta)}{Var(\hat{\theta})}$$

ให้ $X_1,\ldots,X_n$ เป็นตัวอย่างสุ่มจากประชากรที่มีการแจกแจงทวินามเชิงลบ ที่มีพารามิเตอร์ r, p จาก $X_i \sim NB(r,p)$ เมื่อ $i=1,\ldots,n,\ 0<p<1$ เมื่อทราบค่า r

จาก $$I(p) = \frac{rn}{p^2(1-p)}$$

และ $$Var\left(\frac{rn-1}{rn+\sum_{i=1}^{n}x_i-1}\right) \simeq \frac{p^2(1-p)}{2}\left(\frac{2\sum_{i=1}^{n}x_i+2-p}{\sum_{i=1}^{n}x_i\left(\sum_{i=1}^{n}x_i-p+2\right)}\right)$$

เพราะฉะนั้น $$e(\hat{p}) = \frac{1/I(\hat{p})}{Var(\hat{p})}$$

$$= \frac{1\Big/\dfrac{rn}{p^2(1-p)}}{\dfrac{p^2(1-p)}{2}\left(\dfrac{2\sum_{i=1}^{n}x_i+2-p}{\sum_{i=1}^{n}x_i\left(\sum_{i=1}^{n}x_i-p+2\right)}\right)}$$

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

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

ดังนั้น $\dfrac{rn-1}{rn+\sum_{i=1}^{n} X_i - 1}$ เป็นตัวประมาณที่ไม่มีประสิทธิภาพสูงสุดของ p

5. ตัวประมาณไม่เอนเอียงที่มีความแปรปรวนต่ำสุด (Minimum Variance Unbiased Estimator)

นิยามที่ 2.10 $\hat{\theta}$ เป็นตัวประมาณไม่เอนเอียงที่มีความแปรปรวนต่ำสุดของ θ ก็ต่อเมื่อ $\hat{\theta}$ เป็นตัวประมาณที่ไม่เอนเอียงของ θ ที่มีความแปรปรนต่ำสุดในบรรดาตัวประมาณที่ไม่เอนเอียงของ θ ด้วยกัน

นิยามที่ 2.11 ถ้ากลุ่มของฟังก์ชันความหนาแน่นความน่าจะเป็น $\{f(x;\theta), \theta \in \Omega\}$ เมื่อ $\Omega = \{\theta; \gamma < \theta < \delta\}$ โดยที่ γ และ δ เป็นค่าคงที่ที่ทราบค่า และ $f(x;\theta)$ สามารถเขียนอยู่ในรูป

$$f(x;\theta) = \begin{cases} s(x)g(\theta)\exp\left[p(\theta)t(x)\right] & , x = a_1, a_2, a_3, \ldots \\ 0 & , x \text{ มีค่าอื่นๆ} \end{cases}$$

แล้ว $f(x;\theta)$ จะเป็นสมาชิกของวงศ์เลขชี้กำลัง (Exponential Family) แบบไม่ต่อเนื่อง ถ้า

1. $\{x : x = a_1, a_2, a_3, \ldots\}$ ไม่ขึ้นกับ θ เมื่อ $\gamma < \theta < \delta$
2. $p(\theta)$ เป็นฟังก์ชันที่มีความหมายแบบต่อเนื่อง (Nontrivial Continuous Function) ของ θ เมื่อ $\gamma < \theta < \delta$
3. $t(x)$ เป็นฟังก์ชันที่มีความหมาย (Nontrivial Function) ของ X สำหรับทุกๆ ค่าของ X

ทฤษฎีบทที่ 2.1 ให้ $X_1, \ldots, X_n$ เป็นตัวอย่างสุ่มจากประชากรที่มีฟังก์ชันความหนาแน่นความน่าจะเป็น $f(x;\theta)$ ถ้าเป็นสมาชิกของวงศ์เลขชี้กำลัง แล้ว $T(X_1, \ldots, X_n) = \sum_{i=1}^{n} t(x_i)$ จะเป็นสถิติพอเพียงและสมบูรณ์ (Complete Sufficient Statistics) ของ θ

ทฤษฎีบทที่ 2.2 ให้ $X_1, \ldots, X_n$ เป็นตัวอย่างสุ่มจากประชากรที่มีฟังก์ชันความหนาแน่นความน่าจะเป็น $f(x;\theta)$ ถ้า $T(X_1, \ldots, X_n)$ จะเป็นสถิติพอเพียงและสมบูรณ์ของ θ และมีฟังก์ชัน $W(T)$

เป็นตัวประมาณไม่เอนเอียงของ $g(\theta)$ แล้ว $W(T)$ เป็นตัวประมาณไม่เอนเอียงที่มีความแปรปรวนต่ำสุด (Minimize Variance Unbiased Estimator: MVUE) เพียงตัวเดียวของ $g(\theta)$

ทฤษฎีบทที่ 2.3 ถ้า $W(T)$ เป็นตัวประมาณไม่เอนเอียงและมีความแปรปรวนต่ำสุด (MVUE) เพียงตัวเดียวของ $g(\theta)$ แล้ว $W(T)$ จะมีอยู่ตัวเดียว

ให้ $X_1,\ldots,X_n$ เป็นตัวอย่างสุ่มจากประชากรที่มีการแจกแจงทวินามเชิงลบ ที่มีพารามิเตอร์ r และ p จาก $X_i \sim NB(r,p)$ เมื่อ $i=1,\ldots,n$, $0<p<1$ เมื่อทราบค่า r

$$f(x;r,p) = \binom{x+r-1}{x} p^r (1-p)^x \qquad ,x=0,1,2,\ldots$$

$$= \binom{x+r-1}{x} p^r e^{x\ln(1-p)}$$

$$= s(x)g(\theta)\exp[p(\theta)t(x)]$$

โดยที่ $s(x)=\binom{x+r-1}{x}$, $\quad g(p)=p^r$, $\quad p(p)=\ln(1-p)$, $\quad t(x)=x$

1. $\{x : x=0,1,2,\ldots\}$ ไม่ขึ้นกับ p เมื่อ $0<p<1$

2. $p(p)=\begin{cases}\ln(1-p) & ,0<p<1\\ 0 & ,p \text{ มีค่าอื่นๆ}\end{cases}$

ซึ่งเป็นฟังก์ชันที่มีความหมายแบบต่อเนื่องของ p เมื่อ $0<p<1$

3. $t(x)=x$ เป็นฟังก์ชันที่มีความหมายของ X

โดยที่ $g(p)=\begin{cases}p^r & ,0<p<1\\ 0 & ,p \text{ มีค่าอื่นๆ}\end{cases}$

และ $s(x)=\begin{cases}\binom{x+r-1}{x} & ,x=0,1,2,\ldots\\ 0 & ,x \text{ มีค่าอื่นๆ}\end{cases}$

จึงสรุปได้ว่า $f(x;r,p)$ อยู่ในวงศ์เลขชี้กำลัง (Exponential Family) ดังนั้น $\sum_{i=1}^{n} X_i$ เป็นสถิติพอเพียงและสมบูรณ์ของ p

และ $$E\left(\frac{rn-1}{rn+\sum_{i=1}^{n} X_i - 1}\right) = p$$

ดังนั้น $\dfrac{rn-1}{rn+\sum_{i=1}^{n} X_i - 1}$ เป็นตัวประมาณไม่เอนเอียงที่มีความแปรปรวนต่ำสุดของ p

จึงสรุปได้ว่า ตัวประมาณภาวะน่าจะเป็นสูงสุดของ p คือ $\hat{p}_{ML} = \dfrac{rn}{rn+\sum_{i=1}^{n} x_i}$ ซึ่งไม่สามารถหาค่าเฉลี่ยและความแปรปรวนของตัวประมาณดังกล่าวได้ ผู้วิจัยจึงปรับ $\hat{p}_{ML}$ เป็น $\hat{p}_{ML} = \dfrac{rn-1}{rn+\sum_{i=1}^{n} x_i - 1}$ ซึ่งเป็นตัวประมาณไม่เอนเอียงที่มีความแปรปรวนต่ำสุดของ p และมีคุณสมบัติความพอเพียง

โดยที่ $E\left(\hat{p}_{ML}\right) = p$ และ $Var\left(\hat{p}_{ML}\right) \simeq \dfrac{p^2(1-p)}{2}\left(\dfrac{2\sum_{i=1}^{n} x_i + 2 - p}{\sum_{i=1}^{n} x_i \left(\sum_{i=1}^{n} x_i - p + 2\right)}\right)$

จากทฤษฎีบทขีดจำกัดส่วนกลาง (Central Limit Theorem) เมื่อ $n \to \infty$ จะได้ว่า

$$Z = \frac{\left(\dfrac{rn-1}{rn+\sum_{i=1}^{n} x_i - 1}\right) - p}{\sqrt{Var\left(\hat{p}_{ML}\right)}} \sim N(0,1)$$

$$P\left(-Z_{\alpha/2} < Z < Z_{\alpha/2}\right) = 1-\alpha$$

$$P\left(-Z_{\alpha/2} < \frac{\left(\dfrac{rn-1}{rn+\sum_{i=1}^{n} x_i - 1}\right) - p}{\sqrt{Var\left(\hat{p}_{ML}\right)}} < Z_{\alpha/2}\right) = 1-\alpha$$

$$P\left(-Z_{\alpha/2}\sqrt{Var\left(\hat{p}_{ML}\right)} < \left(\frac{rn-1}{rn+\sum_{i=1}^{n} x_i - 1}\right) - p < Z_{\alpha/2}\sqrt{Var\left(\hat{p}_{ML}\right)}\right) = 1-\alpha$$

$$P\left(\left(\frac{rn-1}{rn+\sum_{i=1}^{n}x_i-1}\right)-Z_{\alpha/2}\sqrt{Var\left(\hat{p}_{ML}\right)}<p<\left(\frac{rn-1}{rn+\sum_{i=1}^{n}x_i-1}\right)+Z_{\alpha/2}\sqrt{Var\left(\hat{p}_{ML}\right)}\right)=1-\alpha$$

นั่นคือ ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของ p คือ

ขีดจำกัดล่างของช่วงความเชื่อมั่น คือ

$$\left(\frac{rn-1}{rn+\sum_{i=1}^{n}x_i-1}\right)-Z_{\alpha/2}\sqrt{Var\left(\hat{p}_{ML}\right)}$$

ขีดจำกัดบนของช่วงความเชื่อมั่น คือ

$$\left(\frac{rn-1}{rn+\sum_{i=1}^{n}x_i-1}\right)+Z_{\alpha/2}\sqrt{Var\left(\hat{p}_{ML}\right)}$$

2.3.2 วิธีของเบส์ (Bayes' Method)

สถิติตามแนวคิดของเบส์ (Bayesian Approach) แตกต่างจากสถิติตามแนวคิดเดิม (Classical Approach) ในแนวคิดเดิม การประมาณค่าพารามิเตอร์ θ มักถือว่าเริ่มจากการสุ่มตัวอย่างจากประชากรที่มีฟังก์ชันความหนาแน่นน่าจะเป็น $f(x;\theta)$ และถือว่าพารามิเตอร์ θ เป็นค่าคงตัว แต่ไม่ทราบค่า แต่ในแนวคิดของเบส์มีความพยายามใช้ความรู้เดิมหรือข้อมูลเดิมที่เกี่ยวข้องกับ θ ให้เป็นประโยชน์ในการประมาณค่า θ ให้ดียิ่งขึ้น ดังนั้นจึงถือว่า θ เป็นค่าของตัวแปรสุ่มของ Θ มีการแจกแจงความน่าจะเป็นอยู่ในรูปใดรูปหนึ่ง

ให้ $X_1,\ldots,X_n$ เป็นตัวอย่างสุ่มจากประชากรที่มีฟังก์ชันความหนาแน่นความน่าจะเป็น $f(x;\theta)=f(x|\theta)$ โดยที่ θ เป็นค่าของตัวแปรสุ่ม Θ ในที่นี้เราถือว่า $f(x|\theta)$ เป็นฟังก์ชันความน่าจะเป็นแบบมีเงื่อนไข (Conditional Probability Density Function)

ให้ตัวแปรสุ่ม Θ มีฟังก์ชันความหนาแน่น $g(\theta)$ ซึ่งเรียกว่า ฟังก์ชันการแจกแจงก่อน (Prior Probability Density Function) ของ Θ และ $h(\theta|X_1,\ldots,X_n)$ เป็นฟังก์ชันความน่าจะเป็นแบบมีเงื่อนไขของตัวแปรสุ่ม Θ เมื่อกำหนดให้ว่า $X_1=x_1,\ldots,X_n=x_n$ หรือเรียกฟังก์ชันนี้ว่าฟังก์ชันการแจกแจงภายหลัง (Posterior Probability Density Function) ของ Θ (ประชุม สุวัตถี, 2553)

ในการประมาณค่าด้วยวิธีของเบส์จะกำหนดวิธีการประมาณค่าฟังก์ชันต่างๆ ได้ดังนี้

ฟังก์ชันความน่าจะเป็นร่วมของตัวแปรสุ่ม X เมื่อกำหนดให้ $\Theta=\theta$ คือ

$$\begin{aligned} f(x_1,\ldots,x_n|\theta) &= f(x_1|\theta)\ldots f(x_n|\theta) \\ &= \prod_{i=1}^{n} f(x_i|\theta) \end{aligned}$$

ฟังก์ชันความน่าจะเป็นร่วมของตัวแปรสุ่ม X เมื่อกำหนด θ และฟังก์ชันการแจกแจงก่อนของ θ คือ

$$f(x_1,\ldots,x_n|\theta)g(\theta) = f(x_1|\theta)\ldots f(x_n|\theta)g(\theta)$$

$$= \prod_{i=1}^{n} f(x_i|\theta)g(\theta)$$

ฟังก์ชันการแจกแจงภายหลังของ θ คือ

$$h(\theta|x_1,\ldots,x_n) = \frac{f(x_1,\ldots,x_n|\theta)g(\theta)}{\int f(x_1,\ldots,x_n|\theta)g(\theta)d\theta}$$

$$= \frac{\prod_{i=1}^{n} f(x_i|\theta)g(\theta)}{\int\prod_{i=1}^{n} f(x_i|\theta)g(\theta)d\theta}$$

ตัวประมาณเบส์ภายหลัง (Posterior Bayes' Estimator) ของฟังก์ชัน $\tau(\theta)$ เทียบกับฟังก์ชันการแจกแจงก่อน $g(\theta)$ คือ

$$\tau(\theta) = E(\hat{\tau}(\Theta)|X_1,\ldots,X_n)$$

$$= \int \tau(\theta)h(\theta|x_1,\ldots,x_n)d\theta$$

$$= \frac{\int \tau(\theta)f(x_1,\ldots,x_n|\theta)g(\theta)d\theta}{\int f(x_1,\ldots,x_n|\theta)g(\theta)d\theta}$$

สมมติตัวแปร $X_1,\ldots,X_n$ มีการแจกแจงทวินามเชิงลบ ด้วยค่าความน่าจะเป็นที่จะได้รับสิ่งที่สนใจในการทดลองแต่ละครั้ง p เป็นค่าพารามิเตอร์ที่ต้องการประมาณ สามารถเขียนได้ในรูปแบบ

$$X_i|\,p \sim NB(r,p)$$

ฟังก์ชันความน่าจะเป็นร่วมของตัวแปรสุ่ม X เมื่อกำหนด r และ p สามารถเขียนได้ดังนี้

$$f(x_1,\ldots,x_n|\,r,p) = \prod_{i=1}^{n} f(x_i|\,r,p)$$

$$= \prod_{i=1}^{n}\binom{x_i+r-1}{x_i}p^{rn}(1-p)^{\sum_{i=1}^{n}x_i}$$

เมื่อพิจารณาฟังก์ชันการแจกแจงก่อนให้มีการแจกแจงบีตา ที่มีพารามิเตอร์ a และ b

$$f(p;a,b) = \begin{cases} \dfrac{\Gamma(a+b)\,p^{a-1}(1-p)^{b-1}}{\Gamma(a)\Gamma(b)} & ,0<p<1 \\ 0 & ,p \text{ มีค่าอื่นๆ} \end{cases}$$

จะได้ฟังก์ชันการแจกแจงภายหลัง ดังนี้

$$h(p|x_1,\ldots,x_n) = \frac{\prod_{i=1}^{n}\binom{x_i+r-1}{x_i} p^{rn}(1-p)^{\sum_{i=1}^{n}x_i} \frac{\Gamma(a+b)p^{a-1}(1-p)^{b-1}}{\Gamma(a)\Gamma(b)}}{\int_0^1 \prod_{i=1}^{n}\binom{x_i+r-1}{x_i} p^{rn}(1-p)^{\sum_{i=1}^{n}x_i} \frac{\Gamma(a+b)p^{a-1}(1-p)^{b-1}}{\Gamma(a)\Gamma(b)}\,dp}$$

ค่าคงที่ไม่ขึ้นกับพารามิเตอร์ p ถูกหารทิ้ง จะเหลือแค่

$$= \frac{p^{rn+a-1}(1-p)^{\sum_{i=1}^{n}x_i+b-1}}{\int_0^1 p^{rn+a-1}(1-p)^{\sum_{i=1}^{n}x_i+b-1}\,dp}$$

พิจารณาเทอมส่วน $\int_0^1 p^{rn+a-1}(1-p)^{\sum_{i=1}^{n}x_i+b-1}\,dp$ ในรูปแบบฟังก์ชันบีตา จะได้ว่า

$$\int_0^1 p^{rn+a-1}(1-p)^{\sum_{i=1}^{n}x_i+b-1}\,dp = B\left(rn+a, \sum_{i=1}^{n}x_i+b\right) = \frac{\Gamma(rn+a)\Gamma\left(\sum_{i=1}^{n}x_i+b\right)}{\Gamma\left(rn+a+\sum_{i=1}^{n}x_i+b\right)}$$

ดังนั้นการแจกแจงภายหลัง คือ

$$h(p|x_1,\ldots,x_n) = \frac{\Gamma\left(rn+a+\sum_{i=1}^{n}x_i+b\right)}{\Gamma(rn+a)\Gamma\left(\sum_{i=1}^{n}x_i+b\right)} p^{rn+a-1}(1-p)^{\sum_{i=1}^{n}x_i+b-1}$$

นั่นคือ การแจกแจงภายหลังของ p คือ $(p|x_1,\ldots,x_n) \sim Beta\left(rn+a, \sum_{i=1}^{n}x_i+b\right)$

จะได้ว่า $E\left(\hat{p}_{Bayes}\right) = \hat{p}_{Bayes} = \dfrac{rn+a}{rn+a+\sum_{i=1}^{n}x_i+b}$ เป็นตัวประมาณเบส์ภายหลังของ P

และ $Var\left(\hat{p}_{Bayes}\right) = \dfrac{(rn+a)\left(\sum_{i=1}^{n}x_i+b\right)}{\left(rn+a+\sum_{i=1}^{n}x_i+b\right)^2\left(rn+a+\sum_{i=1}^{n}x_i+b+1\right)}$

หรือพิจารณาโดยอาศัยรูปแบบของวงศ์คู่สังยุค (Conjugate function) ของการแจกแจงบีตากับฟังก์ชันการแจกแจงทวินามเชิงลบ ดังนี้

$$h\left(p|x_1,\ldots,x_n\right) \propto L(r,p)\,f(p;a,b)$$

$$\propto \left(\prod_{i=1}^{n}\binom{x_i+r-1}{x_i}p^{rn}(1-p)^{\sum_{i=1}^{n}x_i}\right)\left(\frac{\Gamma(a+b)\,p^{a-1}(1-p)^{b-1}}{\Gamma(a)\Gamma(b)}\right)$$

$$\propto \prod_{i=1}^{n}\binom{x_i+r-1}{x_i}\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\,p^{rn+a-1}(1-p)^{\sum_{i=1}^{n}x_i+b-1}$$

ดังนั้น การแจกแจงภายหลังของ p คือ $\left(p|x_1,\ldots,x_n\right)\sim Beta\left(rn+a,\sum_{i=1}^{n}x_i+b\right)$

จะได้ว่าตัวประมาณเบส์ภายหลังของ P คือ $E\left(\hat{p}_{Bayes}\right)=\hat{p}_{Bayes}=\dfrac{rn+a}{rn+a+\sum_{i=1}^{n}x_i+b}$

และ $Var\left(\hat{p}_{Bayes}\right)=\dfrac{(rn+a)\left(\sum_{i=1}^{n}x_i+b\right)}{\left(rn+a+\sum_{i=1}^{n}x_i+b\right)^2\left(rn+a+\sum_{i=1}^{n}x_i+b+1\right)}$

จากทฤษฎีบทขีดจำกัดส่วนกลาง (Central Limit Theorem) เมื่อ $n\to\infty$ จะได้ว่า

$$Z=\frac{\left(\dfrac{rn+a}{rn+a+\sum_{i=1}^{n}X_i+b}\right)-p}{\sqrt{Var\left(\hat{p}_{Bayes}\right)}}\sim N(0,1)$$

$$P\left(-Z_{\alpha/2}<Z<Z_{\alpha/2}\right)=1-\alpha$$

$$P\left(-Z_{\alpha/2}<\frac{\left(\dfrac{rn+a}{rn+a+\sum_{i=1}^{n}X_i+b}\right)-p}{\sqrt{Var\left(\hat{p}_{Bayes}\right)}}<Z_{\alpha/2}\right)=1-\alpha$$

$$P\left(-Z_{\alpha/2}\sqrt{Var\left(\hat{p}_{Bayes}\right)}<\left(\frac{rn+a}{rn+a+\sum_{i=1}^{n}x_i+b}\right)-p<Z_{\alpha/2}\sqrt{Var\left(\hat{p}_{Bayes}\right)}\right)=1-\alpha$$

$$P\left(\left(\frac{rn+a}{rn+a+\sum_{i=1}^{n}x_i+b}\right)-Z_{\alpha/2}\sqrt{Var\left(\hat{p}_{Bayes}\right)}<p<\left(\frac{rn+a}{rn+a+\sum_{i=1}^{n}x_i+b}\right)+Z_{\alpha/2}\sqrt{Var\left(\hat{p}_{Bayes}\right)}\right)$$

$$=1-\alpha$$

นั่นคือ ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของ p คือ

ขีดจำกัดล่างของช่วงความเชื่อมั่น คือ

$$\left(\frac{rn+a}{rn+a+\sum_{i=1}^{n}X_i+b}\right)-Z_{\alpha/2}\sqrt{Var\left(\hat{p}_{Bayes}\right)}$$

ขีดจำกัดบนของช่วงความเชื่อมั่น คือ

$$\left(\frac{rn+a}{rn+a+\sum_{i=1}^{n}X_i+b}\right)+Z_{\alpha/2}\sqrt{Var\left(\hat{p}_{Bayes}\right)}$$

2.3.3 วิธีมาร์คอฟ เชน มอนติคาร์โล (Markov Chain Monte Carlo)

วิธีมาร์คอฟ เชน มอนติคาร์โล เป็นวิธีที่นิยมใช้ เมื่อไม่ทราบค่าพารามิเตอร์ของฟังก์ชันการแจกแจงก่อนของตัวแปรสุ่ม โดยประกอบด้วยการสุ่มตัวอย่างตัวแปรจากวิธีมาร์คอฟ เชน (Markov Chain) จากฟังก์ชันการแจกแจงก่อน และทำการสุ่มตัวอย่างแบบกิบส์ (Gibbs Sampling) จากนั้นทำการประมาณค่าด้วยมอนติคาร์โล (Monte Carlo) จากฟังก์ชันการแจกแจงภายหลัง (อัชฌา อระวีพร, 2555)

ขั้นตอนการสร้างสุ่มตัวอย่างจากวิธีมาร์คอฟ เชน ประกอบด้วย

1. กำหนดค่าเริ่มต้น $\theta^{(0)}$ จากฟังก์ชันการแจกแจงก่อน
2. สร้างค่าจากข้อ 1. มา T ค่า จนกระทั่งได้ลักษณะการแจกแจงที่ต้องการ
3. ตรวจสอบการแจกแจงในข้อ 2. ถ้าไม่เป็นไปตามลักษณะที่ต้องการให้สร้างค่ามากขึ้น
4. เมื่อค่าที่ได้เป็นไปตามที่ต้องการเลือกค่าสังเกตที่ค่า B เป็นต้นไป
5. พิจารณาค่า $\{\theta^{(B+1)},\theta^{(B+2)},\ldots,\theta^{(T)}\}$ ในรูปของตัวอย่างสำหรับการวิเคราะห์ฟังก์ชันการแจกแจงภายหลัง
6. สร้างกราฟเพื่อดูลักษณะการแจกแจงของฟังก์ชันการแจกแจงภายหลัง
7. คำนวณค่าเฉลี่ย ค่ากลาง และส่วนเบี่ยงเบนมาตรฐาน จากฟังก์ชันการแจกแจงภายหลัง

วิธีการสุ่มตัวอย่างอีกวิธีหนึ่งที่นิยมใช้กรณีที่ทราบการแจกแจงก่อน เมื่อกำหนดค่าของตัวแปรสุ่ม คือ วิธีกิบส์ ศึกษาโดย (Geman and Geman, 1984) มีขั้นตอนดังนี้

1. กำหนดค่าเริ่มต้น $\theta^{(0)}$ จากฟังก์ชันการแจกแจงก่อน
2. เมื่อ $t=1,2,\ldots,T$ ทำซ้ำตามขั้นตอน

2.1 กำหนด $\theta=\theta^{(t-1)}$

2.2 สร้างค่า θ_j จาก $\theta_j \sim f\left(\theta_1,\ldots,\theta_{j-1},\theta_{j+1},\ldots,\theta_d,\underset{\sim}{x}\right)$ เมื่อ $j=1,\ldots,d$

2.3 กำหนดค่าให้ $\theta^{(t)}=\theta$ โดยเก็บค่านี้ไว้สร้างพารามิเตอร์ใหม่ที่รอบ $t+1$ ตามกระบวนการนี้

$\theta_1^{(t)}$ จาก $f\left(\theta_1 \middle| \theta_2^{(t-1)},\theta_3^{(t-1)},\ldots,\theta_p^{(t-1)},\underset{\sim}{x}\right)$

$\theta_2^{(t)}$ จาก $f\left(\theta_2 \middle| \theta_1^{(t)},\theta_3^{(t-1)},\ldots,\theta_p^{(t-1)},\underset{\sim}{x}\right)$

$\theta_3^{(t)}$ จาก $f\left(\theta_3 \middle| \theta_1^{(t)}, \theta_2^{(t)}, \theta_4^{(t-1)}, \ldots, \theta_p^{(t-1)}, \underset{\sim}{x}\right)$

$\theta_j^{(t)}$ จาก $f\left(\theta_j \middle| \theta_1^{(t)}, \theta_2^{(t)}, \ldots, \theta_{j-1}^{(t-1)}, \theta_{j+1}^{(t-1)}, \ldots, \theta_p^{(t-1)}, \underset{\sim}{x}\right)$

$\theta_p^{(t)}$ จาก $f\left(\theta_p \middle| \theta_1^{(t)}, \theta_3^{(t)}, \ldots, \theta_{p-1}^{(t)}, \underset{\sim}{x}\right)$

ค่าพารามิเตอร์ที่ได้จากการสุ่มตัวอย่างแบบกิบส์สามารถสรุปได้ว่า

$$f\left(\theta_j \middle| \theta_1^{(t)}, \theta_2^{(t)}, \ldots, \theta_{j-1}^{(t-1)}, \theta_{j+1}^{(t-1)}, \ldots, \theta_p^{(t-1)}, x\right) \propto f\left(\theta \middle| \underset{\sim}{x}\right)$$

เมื่อได้ค่าพารามิเตอร์ที่สร้างจากวิธีของมาร์คอฟ เชน และทำการสุ่มตัวอย่างแบบกิบส์ เพื่อประมาณค่าพารามิเตอร์ในสถิติแบบเบส์ โดยใช้วิธีการประมาณค่าด้วยมอนติคาร์โล (Monte Carlo) เพื่อประมาณค่าจากฟังก์ชันการแจกแจงภายหลังจาก

$$E\left(g(\theta)\middle| x_1, \ldots, x_n\right) = \int f\left(\theta \middle| x_1, \ldots, x_n\right) g(\theta)\, d\theta$$

ซึ่งสามารถประมาณค่า θ ได้โดย

$$\bar{\theta} = \frac{1}{T}\sum_{t=1}^{T} \theta^{(t)}$$

กำหนดตัวแปร X มีการแจกแจงทวินามเชิงลบ เป็นจำนวนครั้งของการพบสิ่งที่ไม่สนใจในการทดลองแบร์นูลลี ซึ่งดำเนินการซ้ำๆ กันและเป็นอิสระกันจนกระทั่งได้ความสำเร็จครบ r ครั้ง ด้วยค่าความน่าจะเป็นที่จะได้รับสิ่งที่สนใจในการทดลองแต่ละครั้ง p และ p มีการแจกแจงบีตาด้วยค่าพารามิเตอร์ a และ b ซึ่งค่าพารามิเตอร์ที่ต้องการประมาณมี 3 ตัว คือ p, a และ b

ขั้นตอนการสร้างสุ่มตัวอย่างจากวิธีมาร์คอฟ เชน มอนติคาร์โล ประกอบด้วย

1. กำหนดค่าเริ่มต้น $a^{(t)}$ และ $b^{(t)}$ จากฟังก์ชันการแจกแจงแบบเลขชี้กำลัง
2. สร้างค่าจากข้อ 1. มา T ค่า เมื่อ $t = 1, 2, \ldots, T$
3. สร้างค่า $p^{(t)}$ จากฟังก์ชันการแจกแจงภายหลังของการแจกแจงบีตาที่ค่าพารามิเตอร์ $a^{(t)}$ และ $b^{(t)}$ จากข้อ 1.
4. สร้างกราฟเพื่อดูลักษณะการแจกแจงของฟังก์ชันการแจกแจงภายหลัง
5. คำนวณค่าเฉลี่ย ค่ากลาง ส่วนเบี่ยงเบนมาตรฐาน จากฟังก์ชันการแจกแจงภายหลัง

สามารถประมาณค่า p ได้โดย $\hat{p}_{MCMC} = \frac{1}{T}\sum_{t=1}^{T} p^{(t)}$

โดยมี $E\left(\hat{p}_{MCMC}\right) = \frac{1}{T}\sum_{t=1}^{T} p^{(t)}$

และ $Var\left(\hat{p}_{MCMC}\right) = \frac{1}{T-1}\sum_{t=1}^{T}\left(p^{(t)} - \bar{p}\right)^2$

จากทฤษฎีบทขีดจำกัดส่วนกลาง (Central Limit Theorem) เมื่อ $n \to \infty$ จะได้ว่า

$$Z = \frac{\frac{1}{T}\sum_{t=1}^{T} p^{(t)} - p}{Var\left(\hat{p}_{MCMC}\right)} \qquad \sim N(0,1)$$

$$P\left(-Z_{\alpha/2} < Z < Z_{\alpha/2}\right) \qquad = 1 - \alpha$$

$$P\left(-Z_{\alpha/2} < \frac{\frac{1}{T}\sum_{t=1}^{T} p^{(t)} - p}{Var\left(\hat{p}_{MCMC}\right)} < Z_{\alpha/2}\right) \qquad = 1-\alpha$$

$$P\left(-Z_{\alpha/2}\sqrt{Var\left(\hat{p}_{MCMC}\right)} < \frac{1}{T}\sum_{t=1}^{T} p^{(t)} - p < Z_{\alpha/2}\sqrt{Var\left(\hat{p}_{MCMC}\right)}\right) \qquad = 1-\alpha$$

$$P\left(\frac{1}{T}\sum_{t=1}^{T} p^{(t)} - Z_{\alpha/2}\sqrt{Var\left(\hat{p}_{MCMC}\right)} < p < \frac{1}{T}\sum_{t=1}^{T} p^{(t)} + Z_{\alpha/2}\sqrt{Var\left(\hat{p}_{MCMC}\right)}\right) = 1-\alpha$$

นั่นคือ ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของ p คือ

ขีดจำกัดล่างของช่วงความเชื่อมั่น คือ $\frac{1}{T}\sum_{t=1}^{T} p^{(t)} - Z_{\alpha/2}\sqrt{Var\left(\hat{p}_{MCMC}\right)}$

ขีดจำกัดบนของช่วงความเชื่อมั่น คือ $\frac{1}{T}\sum_{t=1}^{T} p^{(t)} + Z_{\alpha/2}\sqrt{Var\left(\hat{p}_{MCMC}\right)}$

2.4 งานวิจัยที่เกี่ยวข้อง

Mir (2008) ได้ทำการศึกษาการประมาณค่าพารามิเตอร์แบบจุดของการแจกแจงทวินามเชิงลบทั่วไป การแจกแจงทวินามเชิงลบ และการแจกแจงทวินาม โดยวิธีของเบส์ โดยที่กำหนดให้พารามิเตอร์มีการแจกแจงก่อนเป็นการแจกแจงบีตา ผลการศึกษาพบว่าค่าประมาณของเบส์จะมีค่าใกล้เคียงกับข้อมูลที่จำลอง เมื่อค่า a และ b เท่ากัน และแตกต่างเล็กน้อยเมื่อค่า a และ b ที่เท่ากันมีค่าเพิ่มขึ้น

Ganji. *et al.* (2013) ได้ทำการศึกษาการประมาณค่าพารามิเตอร์ r แบบจุดของการแจกแจงทวินามเชิงลบทั่วไป โดยวิธีของเบส์ และวิธีการของเบส์เชิงประจักษ์ ทั้งกรณีที่ทราบและไม่ทราบค่าพารามิเตอร์ p กรณีที่ไม่ทราบค่า p ใช้วิธีการประมาณของเบส์ ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา โดยกำหนดขนาดตัวอย่างเท่ากับ 10 20 และ 30 จากประชากรที่มีการแจกแจงทวินามเชิงลบด้วยค่าพารามิเตอร์ p เป็น 0.2 0.5 และ 0.8 และค่าพารามิเตอร์ r เท่ากับ 5 20 และ 50 กระทำซ้ำ 500 ครั้งในแต่ละสถานการณ์ พบว่าเมื่อค่า p เพิ่มขึ้น ค่าคลาดเคลื่อนกำลังสองเฉลี่ยจะเพิ่มขึ้น สำหรับค่า r แต่ละระดับ เมื่อขนาดตัวอย่างเพิ่มขึ้น ค่าคลาดเคลื่อนกำลังสองเฉลี่ยจะลดลง

Araveeporn (2014) ได้ทำการเปรียบเทียบวิธีประมาณค่าพารามิเตอร์แบบจุดของการแจกแจงปัวซง โดยวิธีภาวะน่าจะเป็นสูงสุด วิธีของเบส์ โดยที่กำหนดให้พารามิเตอร์มีการแจกแจงก่อนเป็นการแจกแจงแกมมา และวิธีมาร์คอฟ เชน มอนติคาร์โลที่ประยุกต์ใช้กับวิธีของเบส์ โดยนำเทคนิคดังกล่าวมาประมาณค่าพารามิเตอร์ของการแจกแจงก่อน พบว่าวิธีภาวะน่าจะเป็นสูงสุด ให้ค่าเฉลี่ย

ของตัวประมาณค่าพารามิเตอร์ไม่แตกต่างจากค่าพารามิเตอร์ทุกกรณี แต่วิธีมาร์คอฟ เชน มอนติคาร์โล และวิธีของเบส์ ให้ค่าเฉลี่ยของตัวประมาณค่าพารามิเตอร์ไม่แตกต่างจากค่าพารามิเตอร์ เมื่อค่าพารามิเตอร์และขนาดตัวอย่างมีค่ามาก แต่เมื่อพิจารณาค่าเฉลี่ยความคลาดเคลื่อนกำลังสองพบว่าวิธีภาวะน่าจะเป็นสูงสุดให้ค่าต่ำสุด เมื่อค่าพารามิเตอร์และขนาดตัวอย่างมีค่าไม่มาก ส่วนวิธีมาร์คอฟ เชน มอนติคาร์โล และวิธีการประมาณของเบส์ให้ค่าต่ำสุด เมื่อค่าพารามิเตอร์และขนาดตัวอย่างมีค่ามาก

Araveeporn (2015) ได้ทำการเปรียบเทียบวิธีประมาณค่าพารามิเตอร์แบบช่วงของการแจกแจงปัวซง โดยวิธีภาวะน่าจะเป็นสูงสุด วิธีของเบส์ โดยที่กำหนดให้พารามิเตอร์มีการแจกแจงก่อนเป็นการแจกแจงเลขชี้กำลังและการแจกแจงแกมมา และวิธีมาร์คอฟ เชน มอนติคาร์โลที่ประยุกต์ใช้กับวิธีของเบส์ โดยนำเทคนิคดังกล่าวมาประมาณค่าพารามิเตอร์ของการแจกแจงก่อน พบว่า วิธีภาวะน่าจะเป็นสูงสุดเป็นวิธีการประมาณที่ดีที่สุด ภายใต้สถานการณ์ที่พารามิเตอร์มีขนาดเล็ก (0.5) สำหรับทุกขนาดตัวอย่าง ส่วนวิธีมาร์คอฟ เชน มอนติคาร์โล จะเป็นวิธีการประมาณที่ดีที่สุดภายใต้สถานการณ์ที่พารามิเตอร์เท่ากับ 5 เมื่อขนาดตัวอย่างมีจำนวนมาก สำหรับวิธีของเบส์ให้ผลลัพธ์ที่ดีทุกกรณีเนื่องจากช่วงความเชื่อมั่นครอบคลุมพารามิเตอร์ทุกสถานการณ์ และโดยส่วนใหญ่แล้วมีความกว้างเฉลี่ยของช่วงความเชื่อมั่นแคบที่สุด

Thetkham and Araveeporn (2016) ได้ทำการเปรียบเทียบวิธีประมาณค่าพารามิเตอร์แบบจุดของการแจกแจงแบบเลขชี้กำลัง โดยวิธีภาวะน่าจะเป็นสูงสุด วิธีของเบส์ โดยที่กำหนดให้พารามิเตอร์มีการแจกแจงก่อนเป็นการแจกแจงแกมมา และวิธีมาร์คอฟ เชน มอนติคาร์โลที่ประยุกต์ใช้กับวิธีของเบส์ โดยนำเทคนิคดังกล่าวมาประมาณค่าพารามิเตอร์ของการแจกแจงก่อน พบว่า ทั้ง 3 วิธีเป็นตัวประมาณค่าที่ไม่เอนเอียง เมื่อค่าพารามิเตอร์และขนาดตัวอย่างมีค่ามาก แต่ตัวประมาณของเบส์จะเปลี่ยนไปตามค่าพารามิเตอร์ของการแจกแจงก่อน จึงควรใช้วิธีมาร์คอฟ เชน มอนติคาร์โล มาประยุกต์ใช้กับวิธีของเบส์

Araveeporn (2016) ได้ทำการเปรียบเทียบวิธีประมาณค่าเฉลี่ยแบบช่วงของการแจกแจงปรกติ โดยวิธีมาร์คอฟ เชน มอนติคาร์โลประยุกต์ใช้กับวิธีของเบส์ เปรียบเทียบกับวิธีภาวะน่าจะเป็นสูงสุด และวิธีของเบส์ โดยที่กำหนดให้พารามิเตอร์มีการแจกแจงก่อนเป็นการแจกแจงปรกติ และนำเทคนิคดังกล่าวมาประมาณค่าพารามิเตอร์ของการแจกแจงก่อน พบว่าวิธีมาร์คอฟ เชน มอนติคาร์โลที่ประยุกต์ใช้กับวิธีของเบส์เป็นวิธีการประมาณที่ดีที่สุดเมื่อตัวอย่างมีขนาดเล็ก ($n=5$) สำหรับวิธีของเบส์มีประสิทธิภาพดีกว่าวิธีการอื่นๆ ที่ขนาดตัวอย่างอื่นๆ วิธีมาร์คอฟ เชน มอนติคาร์โลมีความกว้างเฉลี่ยของช่วงความเชื่อมั่นกว้างที่สุด แต่วิธีมาร์คอฟ เชน มอนติคาร์โลมีประโยชน์ในการประมาณค่าพารามิเตอร์ของการแจกแจงก่อนในวิธีของเบส์

Kumar and Tomer (2016) ได้ทำการศึกษาเปรียบเทียบวิธีการประมาณค่าความเชื่อถือได้ของการแจกแจงทวินามเชิงลบ โดยใช้ข้อมูล Type-II Censoring โดยวิธีภาวะน่าจะเป็นสูงสุด และวิธี

ของเบส์ โดยที่กำหนดให้พารามิเตอร์มีการแจกแจงก่อนเป็นการแจกแจงบีตา และวิธีมาร์คอฟ เชน มอนติคาร์โลที่ประยุกต์ใช้กับวิธีของเบส์ โดยกำหนดขนาดตัวอย่างเท่ากับ 20 และ 50 จากประชากรที่มีการแจกแจงทวินามเชิงลบด้วยค่าพารามิเตอร์ p เป็น 0.3 และ 0.5 และค่าพารามิเตอร์ r เป็น 5 และ 8 กระทำซ้ำ 3,000 ครั้งในแต่ละสถานการณ์พบว่าค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณทั้ง 2 วิธีลดลง หากค่าพารามิเตอร์ r เพิ่มขึ้น เมื่อกำหนดขนาดตัวอย่าง n

บทที่ 3
วิธีการดำเนินงานวิจัย

การศึกษาครั้งนี้จะทำการเปรียบเทียบการประมาณค่าพารามิเตอร์แบบจุดและแบบช่วงของการแจกแจงทวินามเชิงลบ ด้วยวิธีภาวะน่าจะเป็นสูงสุด วิธีของเบส์ ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา และวิธีมาร์คอฟ เชน มอนติคาร์โล โดยการจำลองข้อมูลตามสถานการณ์ต่างๆ และทำการพิจารณาวิธีการประมาณที่เหมาะสมในแต่ละสถานการณ์ โดยมีการวางแผนการวิจัยและวิธีการดำเนินการวิจัย ดังนี้

3.1 การวางแผนการวิจัย

ในการวิจัยครั้งนี้กำหนดสถานการณ์ในการศึกษาเปรียบเทียบ ดังนี้

3.1.1 กำหนดค่าพารามิเตอร์ของประชากร (p) เท่ากับ 0.2 0.5 และ 0.8

3.1.2 กำหนดค่าพารามิเตอร์ r และขนาดตัวอย่าง (n) สำหรับการประมาณค่าแบบจุด แสดงดังตารางที่ 3.1

ตารางที่ 3.1 ค่าพารามิเตอร์สำหรับการประมาณค่าแบบจุด

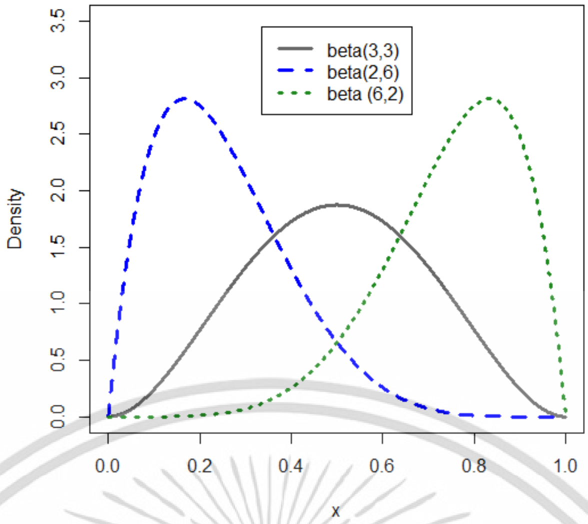
n	r
10	3 5 7
30	10 15 20
50	10 20 30

3.1.3 กำหนดค่าพารามิเตอร์ r และขนาดตัวอย่าง (n) สำหรับการประมาณค่าแบบช่วง แสดงดังตารางที่ 3.2

ตารางที่ 3.2 ค่าพารามิเตอร์สำหรับการประมาณค่าแบบช่วง

n	r
30	10 15 20
50	10 20 30
70	15 30 45

3.1.4 กำหนดค่าพารามิเตอร์ของการแจกแจงก่อน ให้มีรูปแบบสมมาตร เบ้ขวา และเบ้ซ้าย โดยมีค่าพารามิเตอร์ (a,b) เท่ากับ (3,3) (2,6) (6,2) ตามลำดับ ดังรูปที่ 3.1



รูปที่ 3.1 การแจกแจงบีตาที่มีพารามิเตอร์ (3,3) (2,6) และ (6,2)

3.1.5 กำหนดระดับความเชื่อมั่น 3 ระดับ คือ 90% 95% และ 99%

3.1.6 ทดสอบสมมติฐานความแตกต่างระหว่างค่าเฉลี่ยความน่าจะเป็นของความสำเร็จในการทดลองแต่ละครั้งของตัวอย่างกับความน่าจะเป็นของความสำเร็จในการทดลองแต่ละครั้งของประชากร

3.1.7 ตรวจสอบค่าสัมประสิทธิ์ความเชื่อมั่นและเปรียบเทียบความกว้างของการประมาณค่าพารามิเตอร์แบบช่วง เพื่อหาวิธีการประมาณที่ดีที่สุดในแต่ละสถานการณ์

3.2 วิธีการดำเนินการวิจัย

ในการวิจัยนี้มีวิธีการดำเนินการวิจัยดังนี้

3.2.1 จำลองข้อมูลที่ใช้ในการวิจัยด้วยโปรแกรมอาร์ โดยกำหนดให้ข้อมูลมีการแจกแจงทวินามเชิงลบ โดยมีค่าพารามิเตอร์ และขนาดตัวอย่างเป็นไปตามขอบเขตการวิจัย

3.2.2 ประมาณค่าพารามิเตอร์ ด้วยวิธีการประมาณทั้ง 3 วิธี ดังนี้

3.2.2.1 วิธีภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Method)

จะได้ $\hat{p}_{ML} = \dfrac{rn-1}{rn+\sum_{i=1}^{n} x_i - 1}$ คือ ตัวประมาณภาวะน่าจะเป็นสูงสุดของ p

โดยมี ค่าคาดหมาย เท่ากับ $E\left(\hat{p}_{ML}\right) = p$

ค่าความแปรปรวน เท่ากับ $Var\left(\hat{p}_{ML}\right) \simeq \dfrac{p^2(1-p)}{2}\left(\dfrac{2\sum_{i=1}^{n} x_i + 2 - p}{\sum_{i=1}^{n} x_i\left(\sum_{i=1}^{n} x_i - p + 2\right)}\right)$

3.2.2.2 วิธีของเบส์ (Bayes' Method)

การแจกแจงภายหลังของ p คือ $\left(p \mid x_1,\ldots,x_n\right) \sim Beta\left(rn+a, \sum_{i=1}^{n} x_i + b\right)$

จะได้ $\hat{p}_{Bayes} = \dfrac{rn+a}{rn+a+\sum_{i=1}^{n} x_i + b}$ เป็นตัวประมาณเบส์ภายหลังของ p

โดยมี ค่าคาดหมาย เท่ากับ $E\left(\hat{p}_{Bayes}\right) = \dfrac{rn+a}{rn+a+\sum_{i=1}^{n} x_i + b}$

ค่าความแปรปรวน เท่ากับ

$$Var\left(\hat{p}_{Bayes}\right) = \frac{(rn+a)\left(\sum_{i=1}^{n} x_i + b\right)}{\left(rn+a+\sum_{i=1}^{n} x_i + b\right)^2 \left(rn+a+\sum_{i=1}^{n} x_i + b + 1\right)}$$

3.2.2.3 วิธีมาร์คอฟ เชน มอนติคาร์โล (Markov Chain Monte Carlo Method)

จะได้ $\hat{p}_{MCMC} = \dfrac{1}{T}\sum_{t=1}^{T} p^{(t)}$ เป็นตัวประมาณมาร์คอฟ เชน มอนติคาร์โล ของ p

โดยมี ค่าคาดหมาย เท่ากับ $E\left(\hat{p}_{MCMC}\right) = \dfrac{1}{T}\sum_{t=1}^{T} p^{(t)}$

ค่าความแปรปรวน เท่ากับ $Var\left(\hat{p}_{MCMC}\right) = \dfrac{1}{T-1}\sum_{t=1}^{T}\left(p^{(t)} - \bar{p}\right)^2$

3.2.3 คำนวณช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของพารามิเตอร์ด้วยวิธีการประมาณ ทั้ง 3 วิธี ได้แก่

3.2.3.1 วิธีภาวะน่าจะเป็นสูงสุด

ขีดจำกัดล่างของช่วงความเชื่อมั่น คือ $\hat{p}_{ML} - Z_{\alpha/2}\sqrt{Var\left(\hat{p}_{ML}\right)}$

ขีดจำกัดบนของช่วงความเชื่อมั่น คือ $\hat{p}_{ML} + Z_{\alpha/2}\sqrt{Var\left(\hat{p}_{ML}\right)}$

3.2.3.2 วิธีของเบส์

ขีดจำกัดล่างของช่วงความเชื่อมั่น คือ $\hat{p}_{Bayes} - Z_{\alpha/2}\sqrt{Var\left(\hat{p}_{Bayes}\right)}$

ขีดจำกัดบนของช่วงความเชื่อมั่น คือ $\hat{p}_{Bayes} + Z_{\alpha/2}\sqrt{Var\left(\hat{p}_{Bayes}\right)}$

3.2.3.3 วิธีมาร์คอฟ เชน มอนติคาร์โล

ขีดจำกัดล่างของช่วงความเชื่อมั่น คือ $\hat{p}_{MCMC} - Z_{\alpha/2}\sqrt{Var\left(\hat{p}_{MCMC}\right)}$

ขีดจำกัดบนของช่วงความเชื่อมั่น คือ $\hat{p}_{MCMC} + Z_{\alpha/2}\sqrt{Var\left(\hat{p}_{MCMC}\right)}$

3.2.4 ทดสอบความแตกต่างระหว่างค่าเฉลี่ยความน่าจะเป็นของความสำเร็จในการทดลองแต่ละครั้งของตัวอย่างกับความน่าจะเป็นของความสำเร็จในการทดลองแต่ละครั้งประชากร โดยใช้การทดสอบที (t-test) สามารถเขียนสมมติฐานในการทดสอบ ได้ดังนี้

$$H_0 : \mu_{\hat{p}} = p$$

$$H_1 : \mu_{\hat{p}} \neq p$$

โดยที่ $$t = \frac{\overline{\hat{p}} - p}{S_{\hat{p}} / \sqrt{m}}$$

เมื่อ $\overline{\hat{p}}$ แทน ค่าเฉลี่ย (Mean) ของค่า p ที่ได้จากตัวอย่างในแต่ละสถานการณ์

$S_{\hat{p}} = \sqrt{\dfrac{\sum_{j=1}^{m}\left(\hat{p}_j - \overline{\hat{p}_j}\right)^2}{m-1}}$ แทน ส่วนเบี่ยงเบนมาตรฐาน (SD) ของค่า p ที่ได้จากตัวอย่างในแต่ละสถานการณ์

$\hat{p}_j$ แทน ค่าประมาณของค่า p ในการทำซ้ำครั้งที่ j

$\overline{\hat{p}_j}$ แทน ค่าเฉลี่ยของค่าประมาณของค่า p

m แทน จำนวนรอบของการทดลอง ในที่นี้เท่ากับ 1,000 รอบ

องศาเสรี เท่ากับ $m-1$

จะปฏิเสธสมมติฐาน H_0 เมื่อ $|t| > t_{\alpha/2, m-1}$

โดยกำหนดให้ α แทนระดับนัยสำคัญของการทดสอบ ในที่นี้เท่ากับ 0.05

3.2.5 คำนวณค่าประมาณสัมประสิทธิ์ความเชื่อมั่นและค่าเฉลี่ยความกว้างของช่วง

นำช่วงที่คำนวณได้มาพิจารณาว่าครอบคลุมค่าพารามิเตอร์ p หรือไม่ ถ้าช่วงความเชื่อมั่นใดครอบคลุมค่าพารามิเตอร์ p จะทำการนับจำนวนครั้ง และบวกสะสมค่าไว้ เมื่อทำครบทุกช่วงแล้ว ก็จะนำค่าสะสมที่ได้ของช่วงที่ครอบคลุมค่าพารามิเตอร์ p มาหารด้วยจำนวนรอบ ซึ่งเรียกค่าที่ได้ว่า ค่าประมาณสัมประสิทธิ์ความเชื่อมั่น $(1-\hat{\alpha})$ สามารถเขียนแทนได้ ดังนี้

$$1-\hat{\alpha} = \frac{\text{จำนวนรอบที่ช่วงความเชื่อมั่นครอบคลุมค่าพารามิเตอร์}}{m}$$

การประมาณค่าความกว้างเฉลี่ยของช่วง หาได้โดยการหาผลต่างระหว่างขอบเขตบนและขอบเขตล่างของช่วงความเชื่อมั่น แล้วทำการบวกค่าสะสมไว้ และนำค่าที่ได้มาหารด้วยจำนวนรอบที่ช่วงความเชื่อมั่นครอบคลุมค่าพารามิเตอร์ ค่าที่ได้คือ ค่าประมาณของความกว้างเฉลี่ยของช่วงที่คำนวณได้จากแต่ละวิธี

โดยค่าประมาณของความกว้างเฉลี่ยของช่วง เท่ากับ $\dfrac{\sum_{j=1}^{m}\left(U_j - L_j\right)}{m}$

เมื่อ U_j แทน ขอบเขตบนของช่วงความเชื่อมั่น ในการทำซ้ำครั้งที่ j

L_j แทน ขอบเขตล่างของช่วงความเชื่อมั่น ในการทำซ้ำครั้งที่ j

m แทน จำนวนรอบของการทดลอง ในที่นี้เท่ากับ 1,000 รอบ

3.2.6 การเปรียบเทียบค่าประมาณสัมประสิทธิ์ความเชื่อมั่นกับค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด

ทำการเปรียบเทียบค่าประมาณสัมประสิทธิ์ความเชื่อมั่น ($1-\hat{\alpha}$) ว่ามีค่าไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด ($1-\alpha_0$) อย่างมีนัยสำคัญหรือไม่ สามารถเขียนสมมติฐานในการทดสอบได้ดังนี้

$$H_0 : P \geq P_0$$
$$H_1 : P < P_0$$

โดยที่ $$Z = \frac{\hat{P} - P_0}{\sqrt{\frac{P_0(1-P_0)}{m}}}$$

จะปฏิเสธสมมติฐาน H_0 เมื่อ $\frac{\hat{P} - P_0}{\sqrt{\frac{P_0(1-P_0)}{m}}} < Z_{\alpha^*}$

นั่นคือ $\hat{P} < P_0 - Z_{\alpha^*}\sqrt{\frac{P_0(1-P_0)}{m}}$

เมื่อ P แทน สัมประสิทธิ์ความเชื่อมั่น หรือ $P = 1-\alpha$

P_0 แทน สัมประสิทธิ์ความเชื่อมั่นที่กำหนด หรือ $P_0 = 1-\alpha_0$

$\hat{P}$ แทน ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการทดลอง หรือ $\hat{P} = 1-\hat{\alpha}$

m แทน จำนวนรอบของการทดลอง ในที่นี้เท่ากับ 1,000 รอบ

α^* แทน ระดับนัยสำคัญของการทดสอบ ในที่นี้เท่ากับ 0.05

ที่ระดับนัยสำคัญ 90% ($P_0 = 0.90$) สามารถเขียนสมมติฐานในการทดสอบ ได้ดังนี้

$$H_0 : P \geq 0.90$$
$$H_1 : P < 0.90$$

วิธีการประมาณที่ให้ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด คือ วิธีการประมาณที่ให้ค่า

$$\hat{P} \geq 0.90 - 1.96\sqrt{\frac{0.90(1-0.90)}{1{,}000}}$$

$$\hat{P} \geq 0.8814$$

ที่ระดับนัยสำคัญ 95% ($P_0 = 0.95$) สามารถเขียนสมมติฐานในการทดสอบ ได้ดังนี้

$$H_0 : P \geq 0.95$$
$$H_1 : P < 0.95$$

วิธีการประมาณที่ให้ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด คือ วิธีการประมาณที่ให้ค่า

$$\hat{P} \geq 0.95 - 1.96\sqrt{\frac{0.95(1-0.95)}{1,000}}$$

$$\hat{P} \geq 0.9365$$

ที่ระดับนัยสำคัญ 99% ($P_0 = 0.99$) สามารถเขียนสมมติฐานในการทดสอบ ได้ดังนี้

$$H_0 : P \geq 0.99$$
$$H_1 : P < 0.99$$

วิธีการประมาณที่ให้ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด คือ วิธีการประมาณที่ให้ค่า

$$\hat{P} \geq 0.99 - 1.96\sqrt{\frac{0.99(1-0.99)}{1,000}}$$

$$\hat{P} \geq 0.9838$$

ดังนั้น เกณฑ์ตัดสินใจ คือ

(1) เมื่อ $P_0 = 0.90$ จะถือว่าวิธีการประมาณนั้นให้ค่าสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด เมื่อค่า $\hat{P} \geq 0.8814$

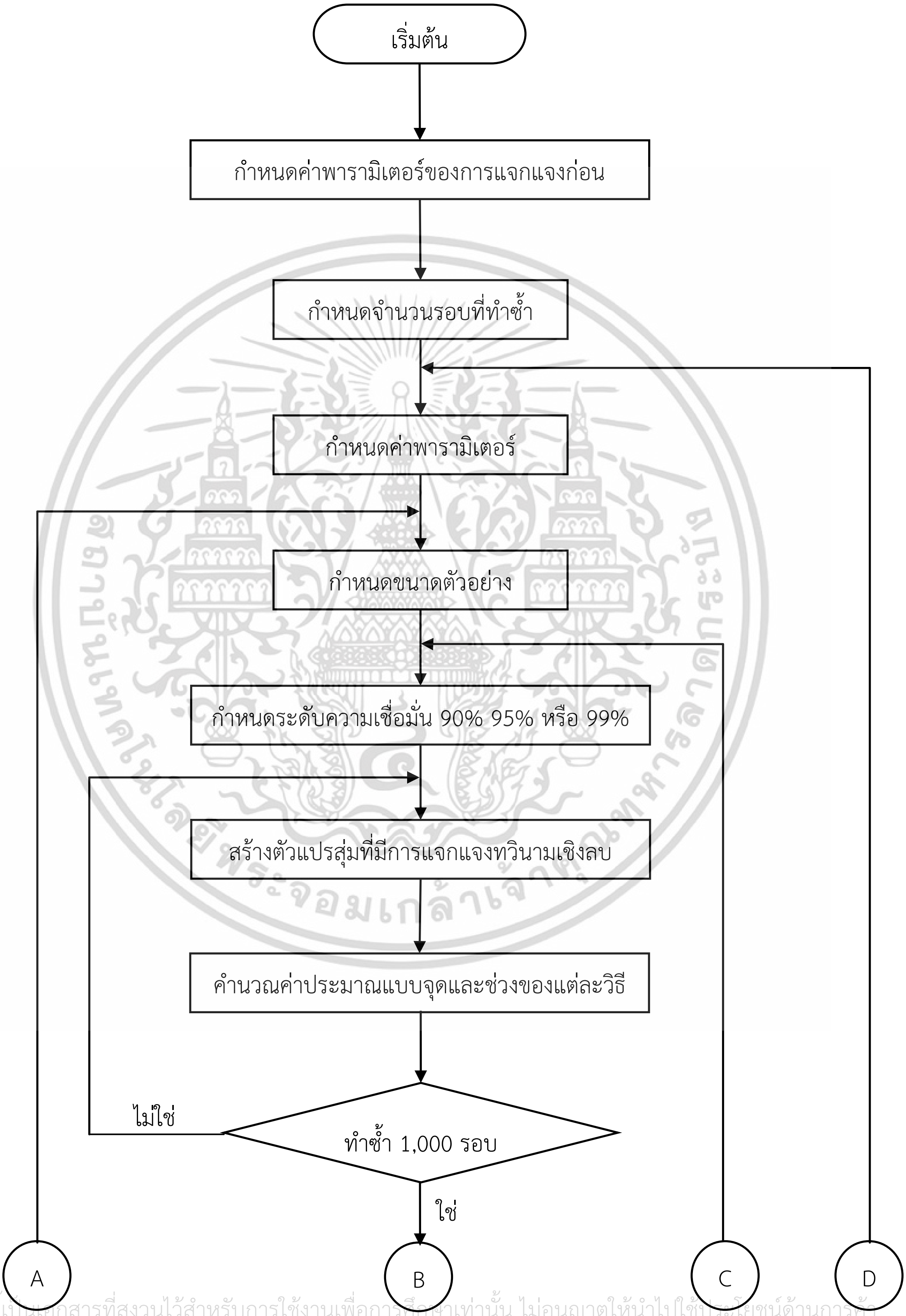
(2) เมื่อ $P_0 = 0.95$ จะถือว่าวิธีการประมาณนั้นให้ค่าสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด เมื่อค่า $\hat{P} \geq 0.9365$

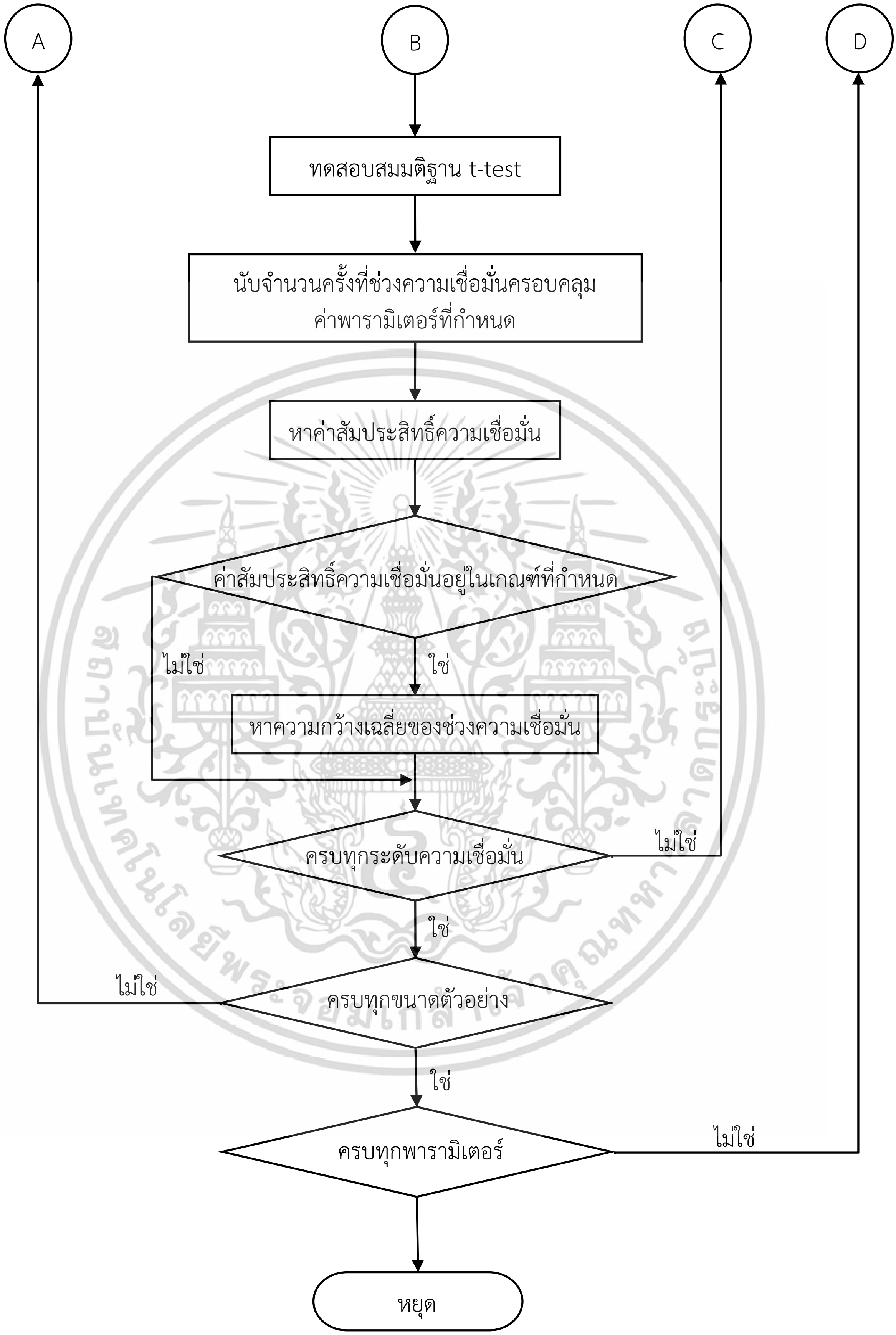
(3) เมื่อ $P_0 = 0.99$ จะถือว่าวิธีการประมาณนั้นให้ค่าสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด เมื่อค่า $\hat{P} \geq 0.9838$

ถ้าแต่ละสถานการณ์มีวิธีการประมาณค่าที่ให้ค่าสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนดหลายวิธี จะนำค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นที่ครอบคลุมค่าพารามิเตอร์มาพิจารณาด้วย โดยวิธีการประมาณใดให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นน้อยที่สุดจะถือว่าวิธีการประมาณนั้นให้ช่วงความเชื่อมั่นที่เหมาะสมที่สุด สำหรับสถานการณ์นั้น

3.3 ขั้นตอนของโปรแกรมที่ใช้ในการวิจัย

การประมวลผลข้อมูลในการวิจัยสามารถอธิบายขั้นตอนเป็นแผนภาพได้ดังนี้





รูปที่ 3.2 แผนผังแสดงลำดับการทำงานของโปรแกรม

บทที่ 4
ผลการวิจัยและการอภิปรายผล

การวิจัยนี้มีจุดมุ่งหมายเพื่อศึกษาวิธีการประมาณค่าแบบจุดและแบบช่วงสำหรับพารามิเตอร์ p ของการแจกแจงทวินามเชิงลบ ด้วยวิธีภาวะน่าจะเป็นสูงสุด วิธีของเบส์ ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา และวิธีมาร์คอฟ เชน มอนติคาร์โล โดยการประมาณค่าแบบจุดจะทำการทดสอบสมมติฐานการเท่ากันระหว่างค่าเฉลี่ยของตัวอย่างกับค่าพารามิเตอร์ของประชากรโดยใช้การทดสอบที สำหรับการประมาณค่าแบบช่วง จะพิจารณาค่าประมาณสัมประสิทธิ์ความเชื่อมั่นและค่าเฉลี่ยความกว้างของช่วงความเชื่อมั่น เพื่อหาวิธีการประมาณที่ดีที่สุดในแต่ละสถานการณ์ จะนำเสนอในรูปแบบของตารางและรูปภาพ โดยมีสัญลักษณ์แทนความหมายต่างๆ ดังนี้

- ML แทน วิธีภาวะน่าจะเป็นสูงสุด
- Bayes1 แทน วิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (3,3)
- Bayes2 แทน วิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (2,6)
- Bayes3 แทน วิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (6,2)
- MCMC แทน วิธีมาร์คอฟ เชน มอนติคาร์โล

4.1 การประมาณค่าแบบจุด (Point Estimation)

ผู้วิจัยทำการทดสอบการเท่ากันระหว่างค่าเฉลี่ยของตัวอย่างกับค่าพารามิเตอร์ของประชากรโดยใช้การทดสอบที (t-test) พิจารณาว่าแตกต่างกันอย่างมีนัยสำคัญหรือไม่ โดยใช้ค่าพี (p-value) เป็นเกณฑ์การพิจารณา นั่นคือ ถ้าค่าพีน้อยกว่าระดับนัยสำคัญที่กำหนด ซึ่งในที่นี้กำหนดให้มีค่าเท่ากับ 0.05 จะทำการปฏิเสธสมมติฐานหลัก นั่นคือจะสรุปผลได้ว่า ค่าเฉลี่ยของตัวประมาณแตกต่างไปจากค่าพารามิเตอร์ของประชากร แสดงผลดังตารางที่ 4.1 – 4.3

ตารางที่ 4.1 ค่าเฉลี่ย ค่าส่วนเบี่ยงเบนมาตรฐาน ค่าสถิติที และค่าพี ของการประมาณค่าพารามิเตอร์ เมื่อ p เท่ากับ 0.2

n	r	วิธีประมาณ	Mean	SD	t-test	p-value
10	3	ML	0.20051	0.03475	0.46385	0.64286
		Bayes1	0.21735	0.03581	15.31953	0.00000*
		Bayes2	0.20795	0.03378	7.43961	0.00000*
		Bayes3	0.23394	0.03801	28.24001	0.00000*
		MCMC	0.20984	0.03559	8.74123	0.00000*
	5	ML	0.19977	0.02637	-0.28031	0.77930
		Bayes1	0.20998	0.02691	11.72747	0.00000*
		Bayes2	0.20437	0.02598	5.32279	0.00000*
		Bayes3	0.22009	0.02797	22.71378	0.00000*
		MCMC	0.20535	0.02677	6.32160	0.00000*
	7	ML	0.20012	0.02221	0.17489	0.86120
		Bayes1	0.20746	0.02255	10.46396	0.00000*
		Bayes2	0.20345	0.02198	4.96211	0.00000*
		Bayes3	0.21475	0.02320	20.10398	0.00000*
		MCMC	0.20413	0.02246	5.81027	0.00000*
30	10	ML	0.19998	0.01040	-0.05491	0.95622
		Bayes1	0.20171	0.01044	5.18179	0.00000*
		Bayes2	0.20078	0.01038	2.36700	0.01812*
		Bayes3	0.20344	0.01051	10.33506	0.00000*
		MCMC	0.20092	0.01043	2.78970	0.00538*
	15	ML	0.20010	0.00866	0.37435	0.70822
		Bayes1	0.20126	0.00868	4.57528	0.00001*
		Bayes2	0.20063	0.00865	2.31578	0.02077*
		Bayes3	0.20241	0.00872	8.73139	0.00000*
		MCMC	0.20072	0.00868	2.62770	0.00873*
	20	ML	0.20031	0.00741	1.31638	0.18835
		Bayes1	0.20117	0.00743	5.00031	0.00000*
		Bayes2	0.20071	0.00741	3.01840	0.00261*

ตารางที่ 4.1 (ต่อ) ค่าเฉลี่ย ค่าส่วนเบี่ยงเบนมาตรฐาน ค่าสถิติที และค่าพี ของการประมาณค่าพารามิเตอร์ เมื่อ p เท่ากับ 0.2

n	r	วิธีประมาณ	Mean	SD	t-test	p-value
30	20	Bayes3	0.20204	0.00746	8.65461	0.00000*
		MCMC	0.20078	0.00742	3.30777	0.00097*
50	10	ML	0.19984	0.00792	-0.61965	0.53563
		Bayes1	0.20088	0.00794	3.51467	0.00046*
		Bayes2	0.20032	0.00791	1.29095	0.19702
		Bayes3	0.20192	0.00798	7.60902	0.00000*
		MCMC	0.20040	0.00794	1.59972	0.10998
	20	ML	0.20010	0.00560	0.53835	0.59045
		Bayes1	0.20062	0.00561	3.46916	0.00054*
		Bayes2	0.20033	0.00560	1.89191	0.05879
		Bayes3	0.20113	0.00562	6.38572	0.00000*
		MCMC	0.20038	0.00560	2.12089	0.03418*
	30	ML	0.20010	0.00462	0.69592	0.48664
		Bayes1	0.20045	0.00462	3.06638	0.00222*
		Bayes2	0.20026	0.00462	1.79044	0.07369
		Bayes3	0.20079	0.00463	5.42914	0.00000*
		MCMC	0.20029	0.00462	1.98706	0.04719*

หมายเหตุ * หมายความว่า ปฏิเสธสมมติฐานหลัก ที่ระดับนัยสำคัญเท่ากับ 0.05

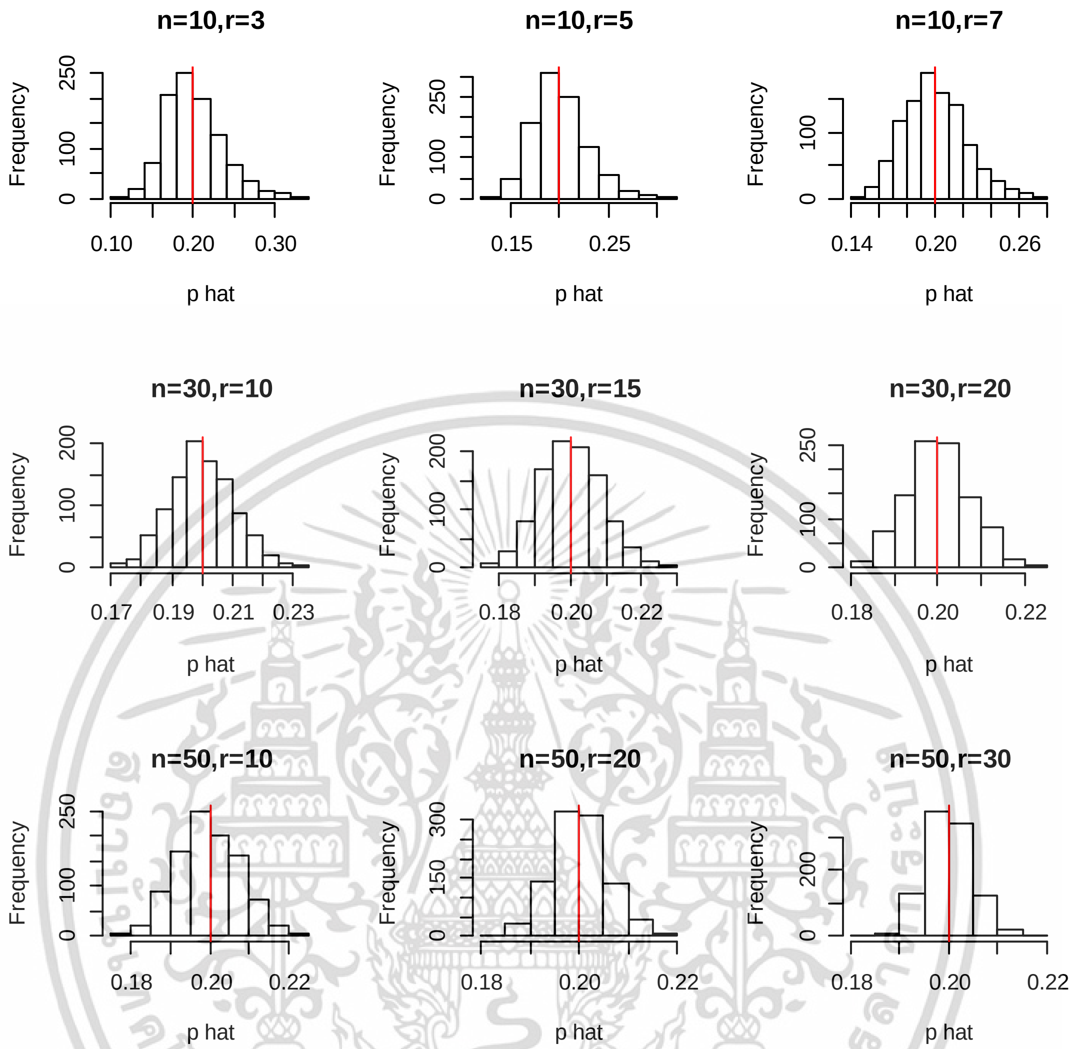
จากตารางที่ 4.1 การประมาณค่าแบบจุด เมื่อ p เท่ากับ 0.2 พบว่าวิธีการประมาณที่ให้ค่าเฉลี่ยของตัวประมาณเท่ากับค่าพารามิเตอร์ของประชากร (p) ได้แก่

- วิธีภาวะน่าจะเป็นสูงสุด ทุกสถานการณ์

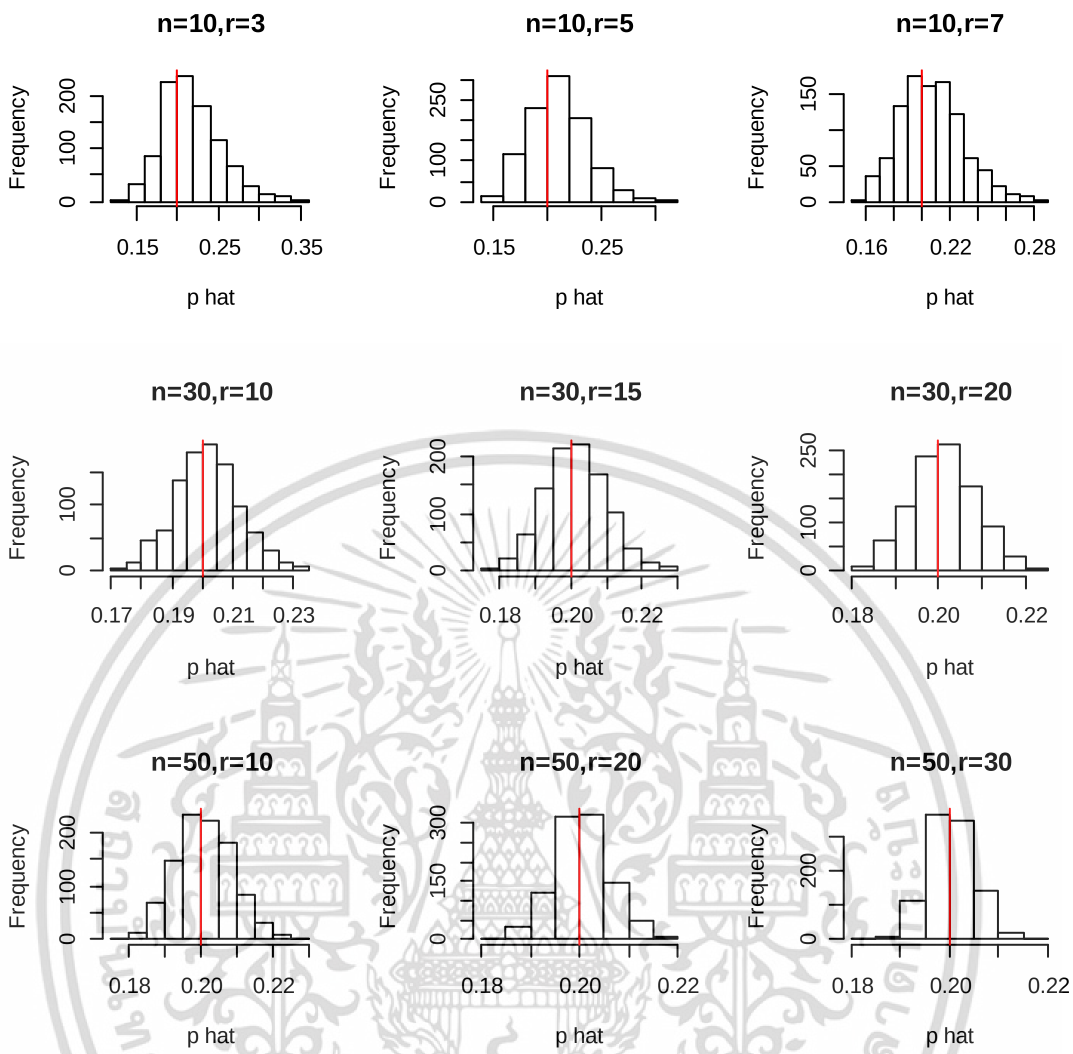
- วิธีของเบส์มีการแจกแจงบีตา (2,6) เป็นการแจกแจงก่อน ที่ขนาดตัวอย่างเท่ากับ 50 ทุกค่าพารามิเตอร์ r

- วิธีมาร์คอฟ เชน มอนติคาร์โล ที่ขนาดตัวอย่างเท่ากับ 50 และค่าพารามิเตอร์ r = 10

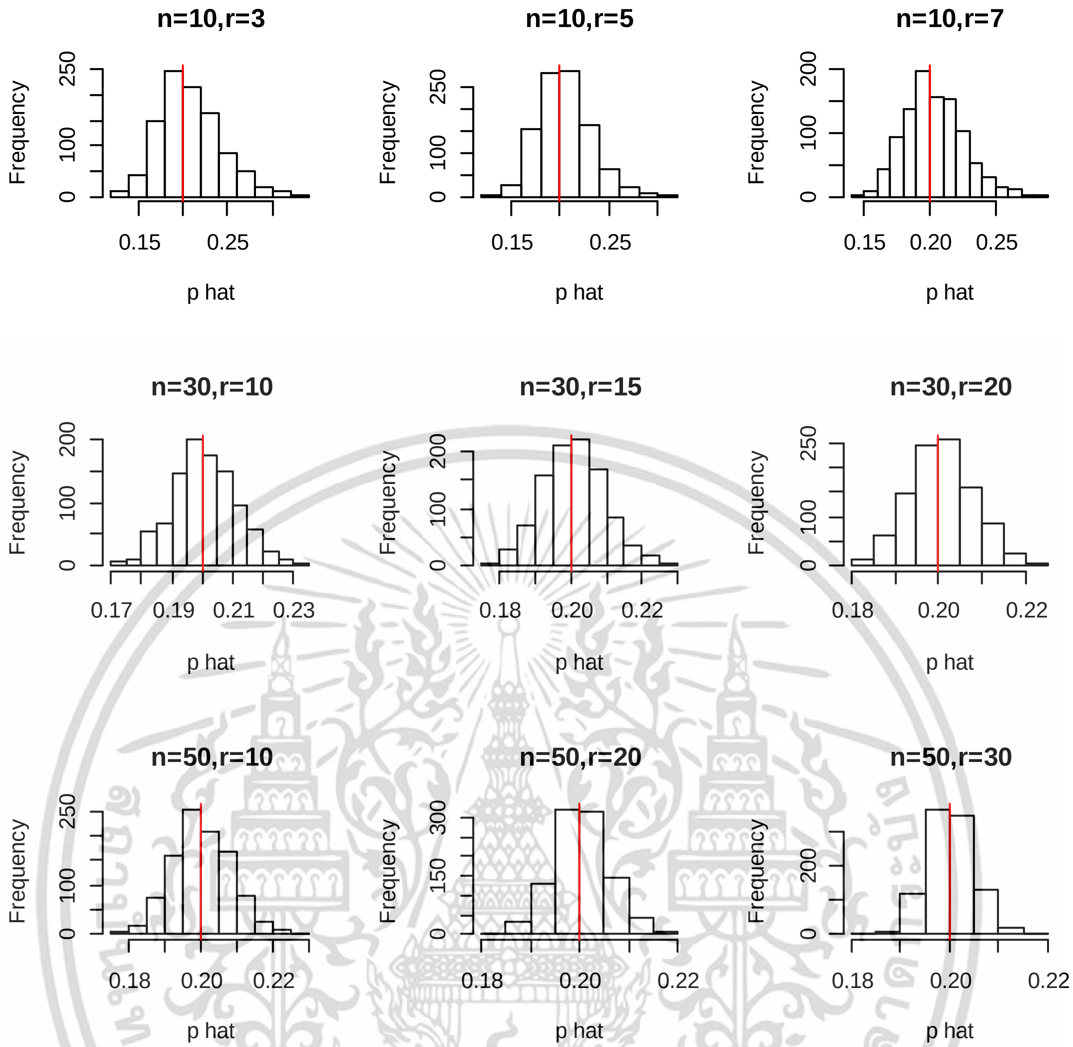
จากรูปที่ 4.1 – 4.5 แสดงให้เห็นว่า ค่าที่ประมาณได้จากทุกวิธีการประมาณจะมีแนวโน้มเข้าสู่การแจกแจงปรกติ เมื่อขนาดตัวอย่างเพิ่มขึ้น



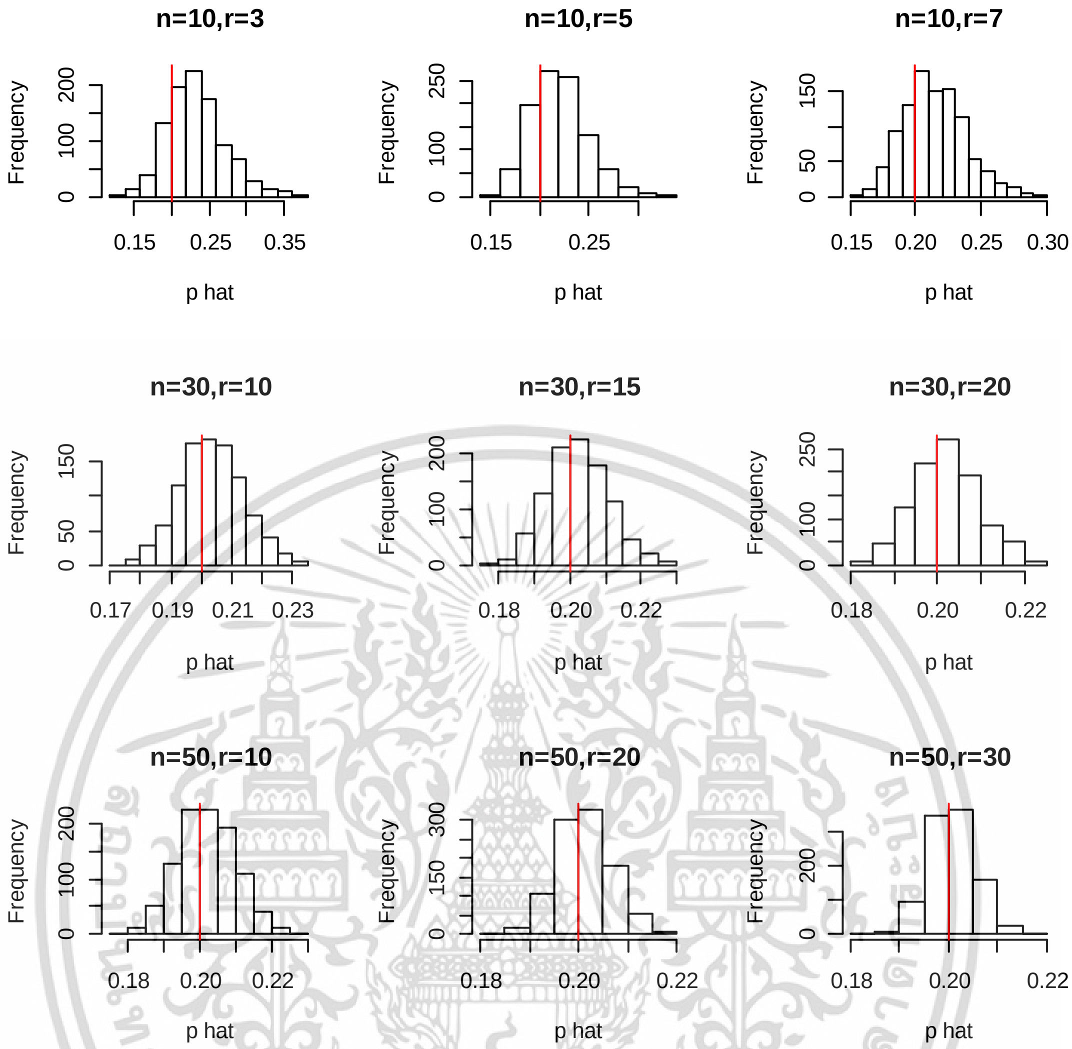
รูปที่ 4.1 แผนภาพฮิสโทแกรมแสดงค่าที่ประมาณได้ด้วยวิธีภาวะน่าจะเป็นสูงสุดเมื่อ $p = 0.2$



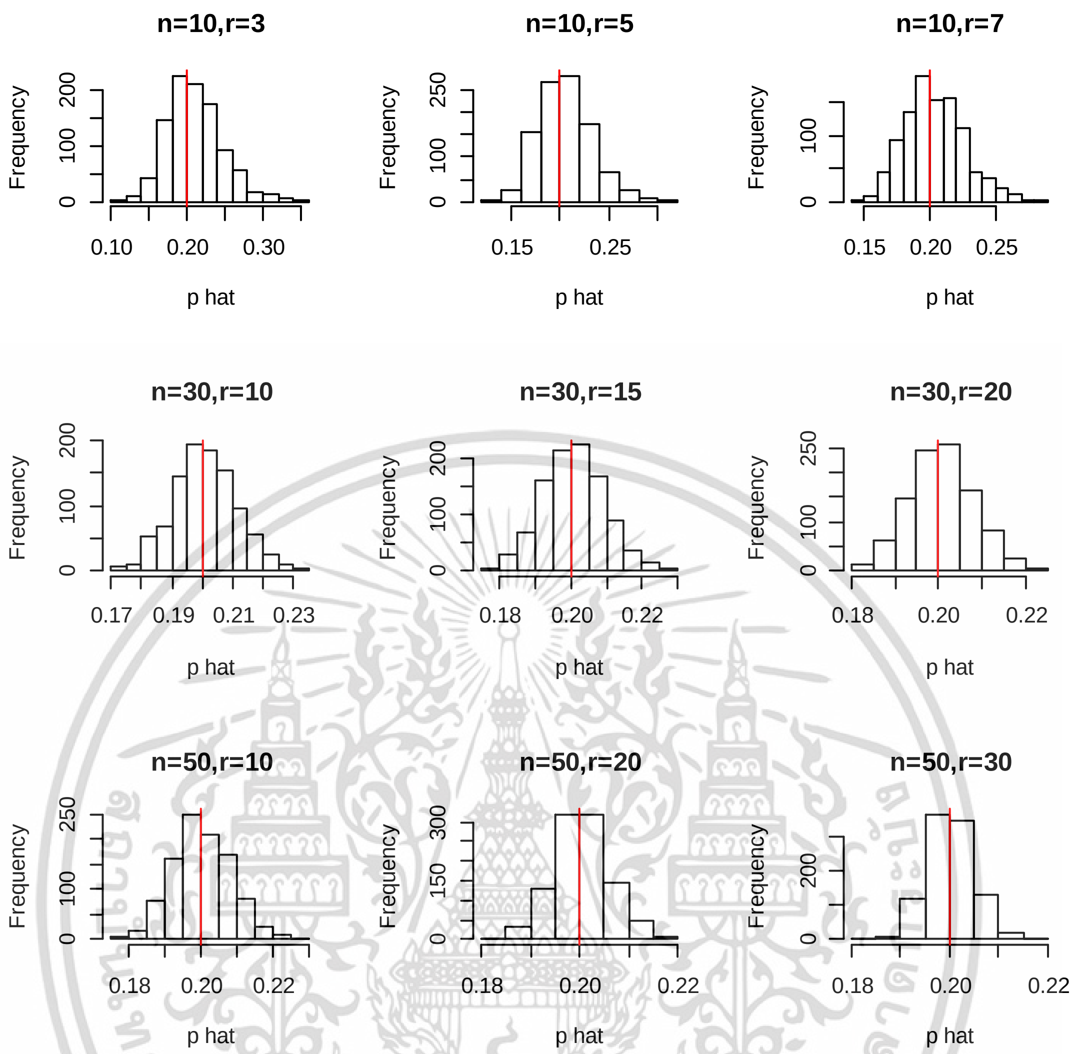
รูปที่ 4.2 แผนภาพฮิสโทแกรมแสดงค่าที่ประมาณได้ด้วยวิธีของเบส์ที่มีการแจกแจงบีตา (3,3) เป็นการแจกแจงก่อน เมื่อ $p = 0.2$



รูปที่ 4.3 แผนภาพฮิสโทแกรมแสดงค่าที่ประมาณได้ด้วยวิธีของเบส์ที่มีการแจกแจงบีตา (2,6) เป็นการแจกแจงก่อน เมื่อ $p = 0.2$



รูปที่ 4.4 แผนภาพฮิสโทแกรมแสดงค่าที่ประมาณได้ด้วยวิธีของเบส์ที่มีการแจกแจงบีตา (6,2) เป็นการแจกแจงก่อน เมื่อ $p = 0.2$



รูปที่ 4.5 แผนภาพฮิสโทแกรมแสดงค่าที่ประมาณได้ด้วยวิธีมาร์คอฟ เชน มอนติคาร์โล เมื่อ $p = 0.2$

ตารางที่ 4.2 ค่าเฉลี่ย ค่าส่วนเบี่ยงเบนมาตรฐาน ค่าสถิติที และค่าพี ของการประมาณค่าพารามิเตอร์ เมื่อ p เท่ากับ 0.5

n	r	วิธีประมาณ	Mean	SD	t-test	p-value
10	3	ML	0.50083	0.06776	0.38863	0.69763
		Bayes1	0.50756	0.06101	3.91748	0.00010*
		Bayes2	0.47729	0.05561	-12.91470	0.00000*
		Bayes3	0.53695	0.06256	18.67817	0.00000*
		MCMC	0.50852	0.06518	4.13337	0.00004*
	5	ML	0.50046	0.04914	0.29587	0.76739
		Bayes1	0.50489	0.04622	3.34742	0.00085*
		Bayes2	0.48603	0.04365	-10.12057	0.00000*
		Bayes3	0.52342	0.04701	15.75293	0.00000*
		MCMC	0.50523	0.04807	3.43802	0.00061*
	7	ML	0.50084	0.04223	0.62696	0.53083
		Bayes1	0.50409	0.04042	3.19673	0.00143*
		Bayes2	0.49037	0.03878	-7.85715	0.00000*
		Bayes3	0.51761	0.04093	13.60395	0.00000*
		MCMC	0.50430	0.04160	3.26595	0.00113*
30	10	ML	0.50015	0.01965	0.24103	0.80958
		Bayes1	0.50097	0.01946	1.56995	0.11674
		Bayes2	0.49766	0.01926	-3.83394	0.00013*
		Bayes3	0.50426	0.01952	6.89520	0.00000*
		MCMC	0.50097	0.01960	1.57193	0.11628
	15	ML	0.49978	0.01669	-0.42568	0.67043
		Bayes1	0.50033	0.01658	0.61999	0.53541
		Bayes2	0.49812	0.01647	-3.61175	0.00032*
		Bayes3	0.50253	0.01661	4.81003	0.00000*
		MCMC	0.50033	0.01666	0.62703	0.53078
	20	ML	0.50043	0.01482	0.92392	0.35575
		Bayes1	0.50084	0.01475	1.80807	0.07090
		Bayes2	0.49918	0.01468	-1.76061	0.07861

ตารางที่ 4.2 (ต่อ) ค่าเฉลี่ย ค่าส่วนเบี่ยงเบนมาตรฐาน ค่าสถิติที และค่าพี ของการประมาณค่าพารามิเตอร์ เมื่อ p เท่ากับ 0.5

n	r	วิธีประมาณ	Mean	SD	t-test	p-value
30	20	Bayes3	0.50250	0.01477	5.35014	0.00000*
		MCMC	0.50086	0.01481	1.82831	0.06780
50	10	ML	0.50055	0.01626	1.06196	0.28851
		Bayes1	0.50104	0.01616	2.02849	0.04278*
		Bayes2	0.49905	0.01606	-1.87979	0.06043
		Bayes3	0.50302	0.01619	5.90188	0.00000*
		MCMC	0.50104	0.01623	2.02917	0.04271*
	20	ML	0.49996	0.01119	-0.12385	0.90146
		Bayes1	0.50020	0.01116	0.58055	0.56168
		Bayes2	0.49921	0.01112	-2.25186	0.02455*
		Bayes3	0.50120	0.01117	3.40030	0.00070*
		MCMC	0.50020	0.01118	0.56781	0.57029
	30	ML	0.50017	0.00895	0.60618	0.54453
		Bayes1	0.50034	0.00893	1.19397	0.23277
		Bayes2	0.49967	0.00891	-1.16517	0.24423
		Bayes3	0.50100	0.00894	3.54610	0.00041*
		MCMC	0.50034	0.00894	1.19138	0.23379

<u>หมายเหตุ</u> * หมายความว่า ปฏิเสธสมมติฐานหลัก ที่ระดับนัยสำคัญเท่ากับ 0.05

จากตารางที่ 4.2 การประมาณค่าแบบจุด เมื่อ p เท่ากับ 0.5 พบว่าวิธีการประมาณที่ให้ค่าเฉลี่ยของตัวประมาณเท่ากับค่าพารามิเตอร์ของประชากร (p) ได้แก่

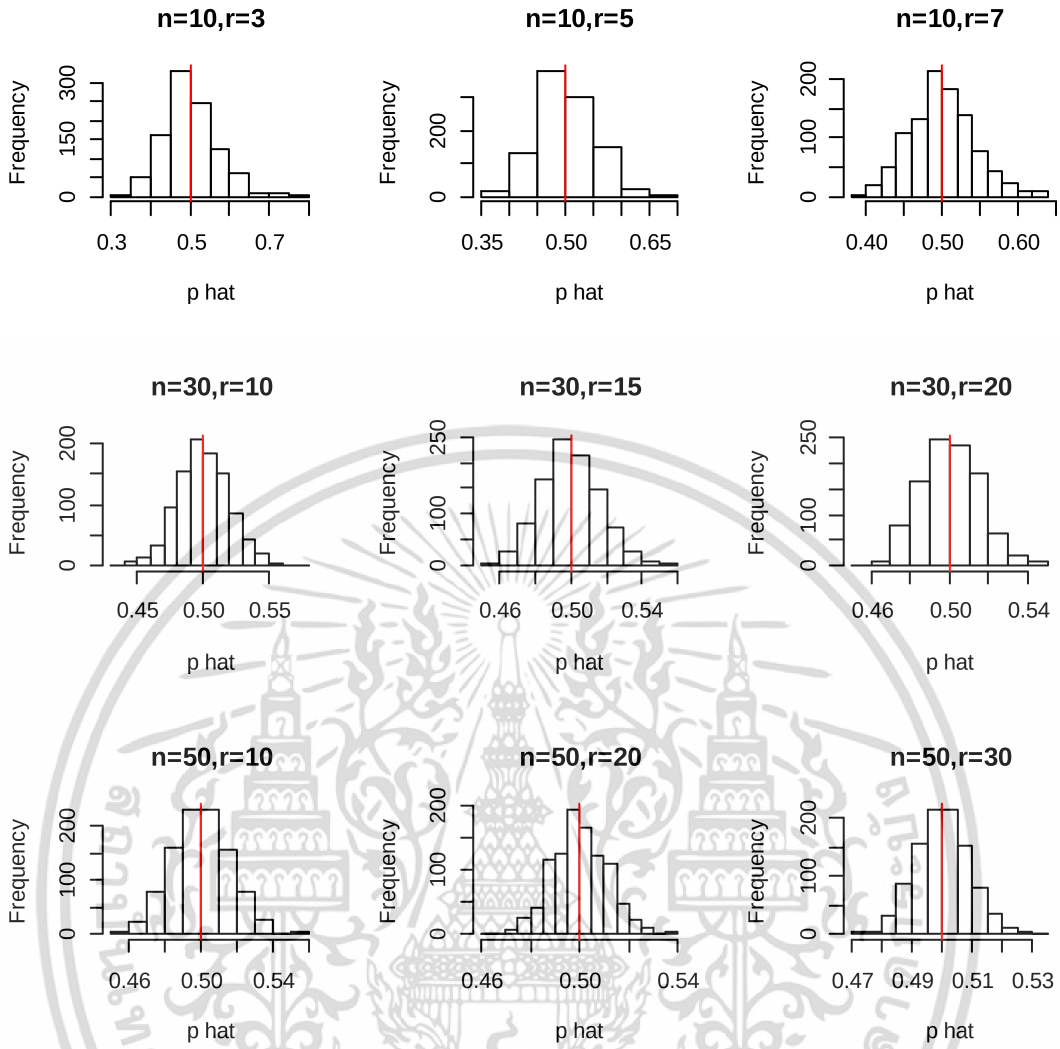
- วิธีภาวะน่าจะเป็นสูงสุด ทุกสถานการณ์

- วิธีของเบส์มีการแจก แจงบีตา (3,3) เป็นการแจกแจงก่อน ที่ขนาดตัวอย่างเท่ากับ 30 ทุกค่าพารามิเตอร์ r และที่ขนาดตัวอย่างเท่ากับ 50 ค่าพารามิเตอร์ r = 20 และ 30

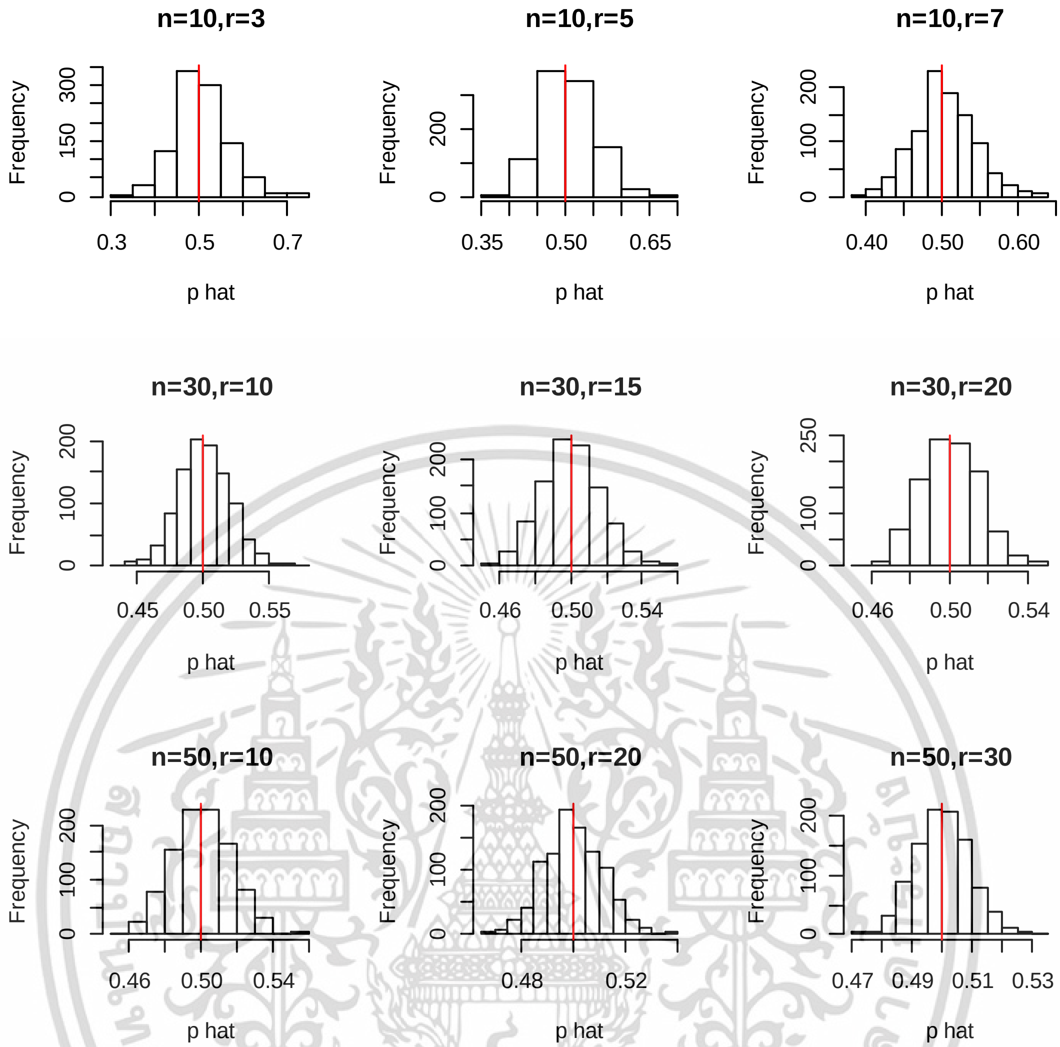
- วิธีของเบส์มีการแจก แจงบีตา (2,6) เป็นการแจกแจงก่อน ที่ขนาดตัวอย่างเท่ากับ 30 ค่าพารามิเตอร์ r = 20 และที่ขนาดตัวอย่างเท่ากับ 50 ค่าพารามิเตอร์ r = 10 และ 30

- วิธีมาร์คอฟ เชน มอนติคาร์โล ที่ขนาดตัวอย่างเท่ากับ 30 ทุกค่าพารามิเตอร์ r และที่ขนาดตัวอย่างเท่ากับ 50 ค่าพารามิเตอร์ r = 20 และ 30

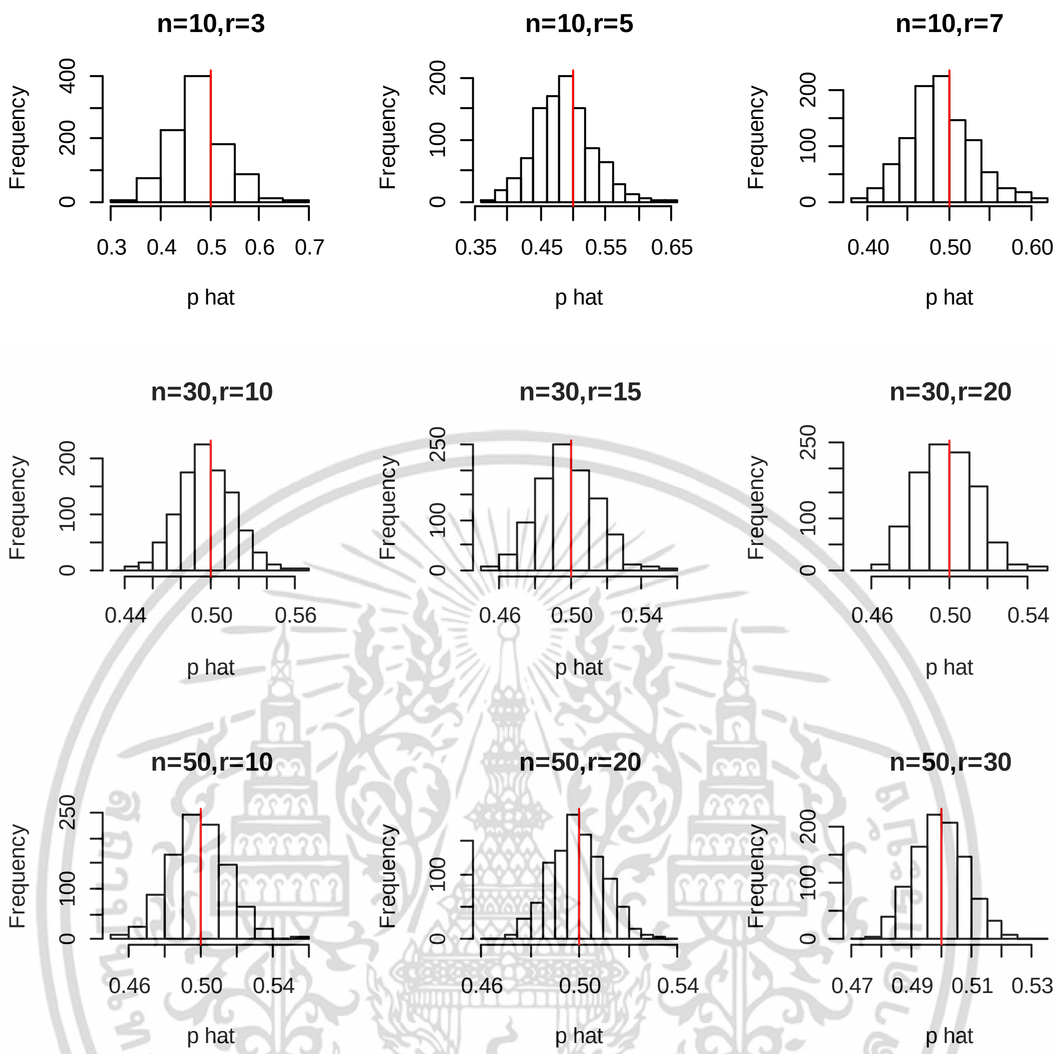
จากรูปที่ 4.6 – 4.10 แสดงให้เห็นว่า ค่าที่ประมาณได้จากทุกวิธีการประมาณจะมีแนวโน้มเข้าสู่การแจกแจงปรกติ เมื่อขนาดตัวอย่างเพิ่มขึ้น



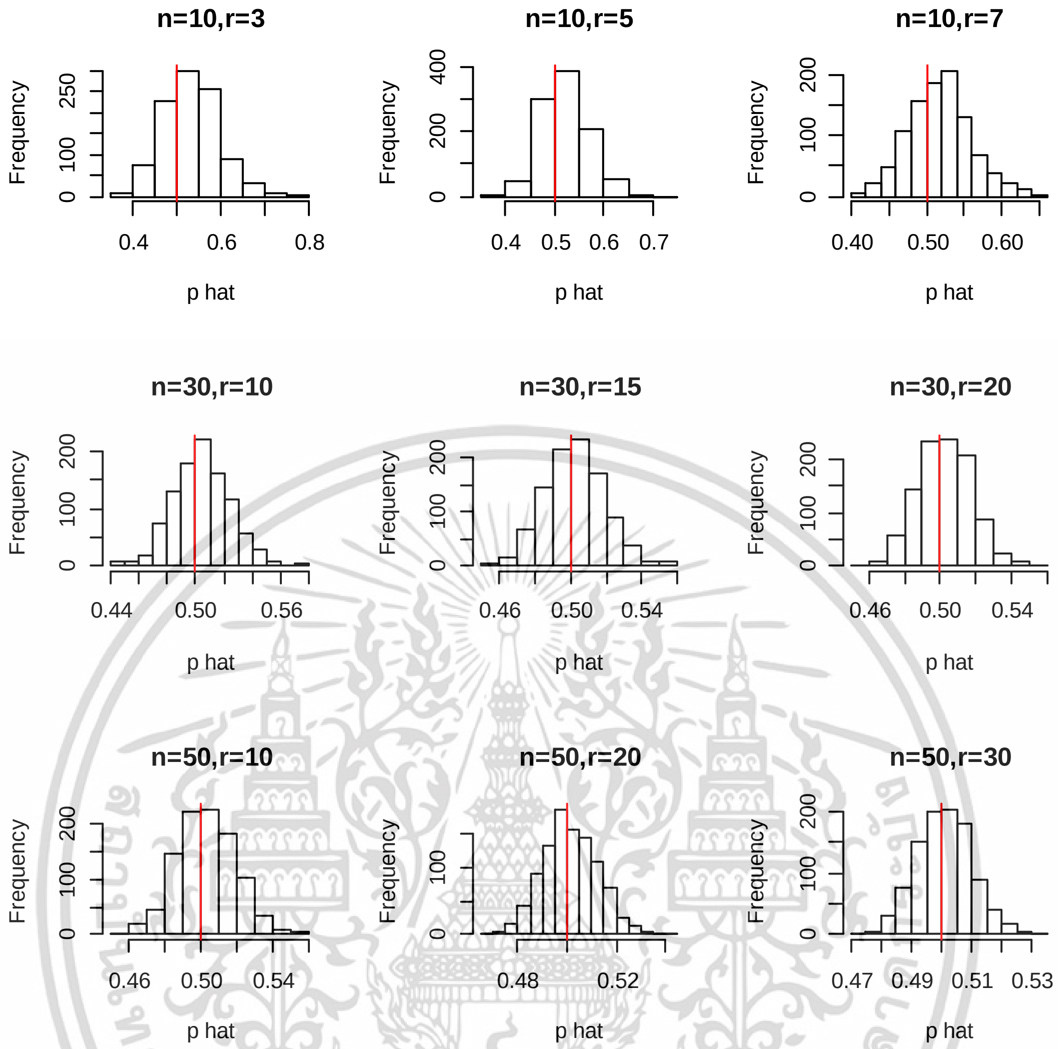
รูปที่ 4.6 แผนภาพฮิสโทแกรมแสดงค่าที่ประมาณได้ด้วยวิธีภาวะน่าจะเป็นสูงสุดเมื่อ $p = 0.5$



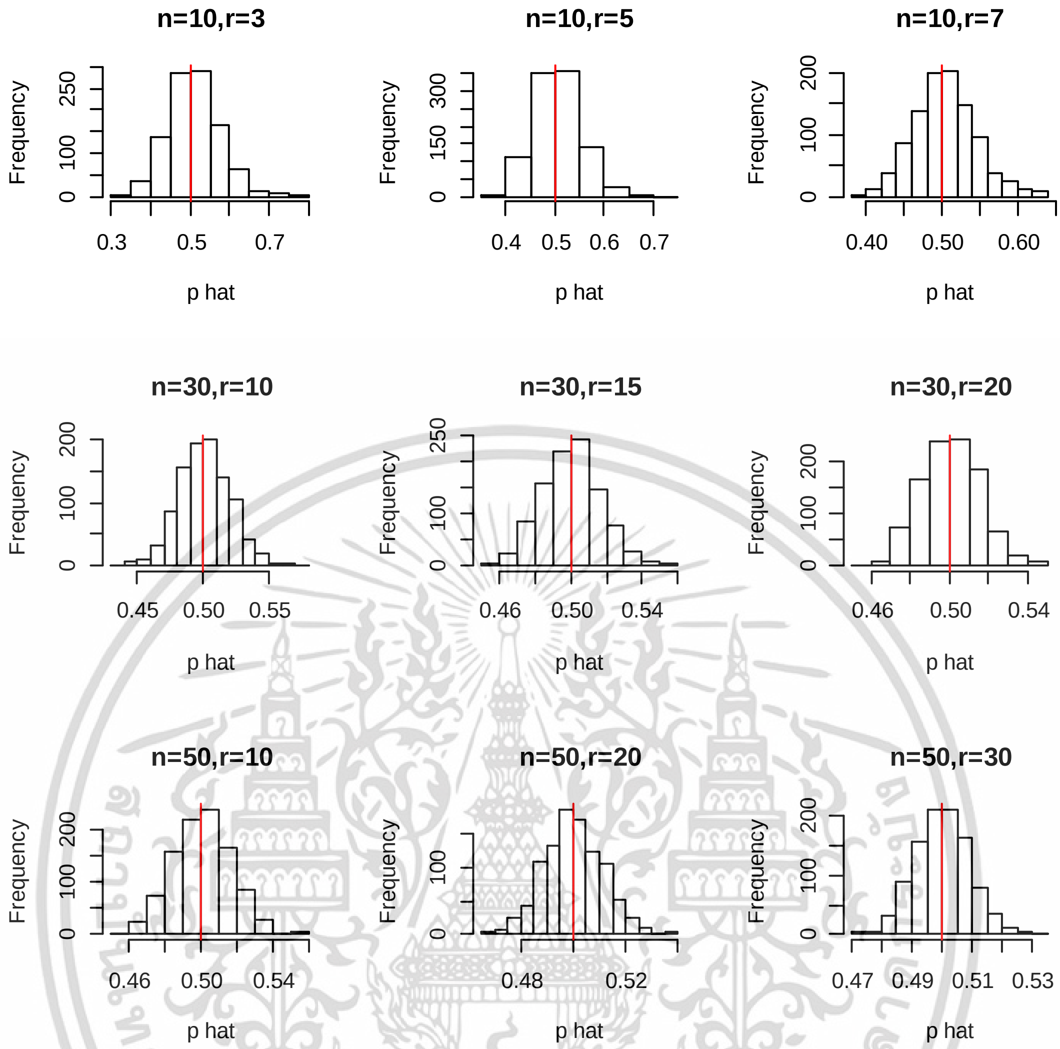
รูปที่ 4.7 แผนภาพฮิสโทแกรมแสดงค่าที่ประมาณได้ด้วยวิธีของเบส์ที่มีการแจกแจงบีตา (3,3) เป็นการแจกแจงก่อน เมื่อ $p = 0.5$



รูปที่ 4.8 แผนภาพฮิสโทแกรมแสดงค่าที่ประมาณได้ด้วยวิธีของเบส์ที่มีการแจกแจงบีตา (2,6) เป็นการแจกแจงก่อน เมื่อ $p = 0.5$



รูปที่ 4.9 แผนภาพฮิสโทแกรมแสดงค่าที่ประมาณได้ด้วยวิธีของเบส์ที่มีการแจกแจงบีตา (6,2) เป็นการแจกแจงก่อน เมื่อ $p = 0.5$



รูปที่ 4.10 แผนภาพฮิสโทแกรมแสดงค่าที่ประมาณได้ด้วยวิธีมาร์คอฟ เชน มอนติคาร์โล เมื่อ $p = 0.5$

ตารางที่ 4.3 ค่าเฉลี่ย ค่าส่วนเบี่ยงเบนมาตรฐาน ค่าสถิติที และค่าพี ของการประมาณค่าพารามิเตอร์ เมื่อ p เท่ากับ 0.8

n	r	วิธีประมาณ	Mean	SD	t-test	p-value
10	3	ML	0.80182	0.06677	0.86297	0.38836
		Bayes1	0.76374	0.05341	-21.46756	0.00000*
		Bayes2	0.70769	0.04733	-61.67751	0.00000*
		Bayes3	0.79615	0.05325	-2.28728	0.02239*
		MCMC	0.79117	0.06102	-4.57797	0.00001*
	5	ML	0.80070	0.05021	0.44210	0.65851
		Bayes1	0.77689	0.04377	-16.69892	0.00000*
		Bayes2	0.74045	0.04054	-46.44654	0.00000*
		Bayes3	0.79741	0.04366	-1.87436	0.06117
		MCMC	0.79431	0.04758	-3.77869	0.00017*
	7	ML	0.80122	0.04198	0.91821	0.35873
		Bayes1	0.78380	0.03799	-13.48860	0.00000*
		Bayes2	0.75677	0.03591	-38.06631	0.00000*
		Bayes3	0.79881	0.03791	-0.99002	0.32240
		MCMC	0.79661	0.04034	-2.65932	0.00795*
30	10	ML	0.80166	0.02069	2.54416	0.01110*
		Bayes1	0.79741	0.02020	-4.04848	0.00006*
		Bayes2	0.79062	0.01993	-14.88972	0.00000*
		Bayes3	0.80109	0.02019	1.70654	0.08822
		MCMC	0.80056	0.02050	0.86376	0.38793
	15	ML	0.80119	0.01658	2.27500	0.02312*
		Bayes1	0.79835	0.01632	-3.18972	0.00147*
		Bayes2	0.79379	0.01617	-12.14129	0.00000*
		Bayes3	0.80082	0.01631	1.58467	0.11336
		MCMC	0.80048	0.01649	0.92583	0.35476
	20	ML	0.80115	0.01419	2.55619	0.01073*
		Bayes1	0.79901	0.01402	-2.22426	0.02635*
		Bayes2	0.79558	0.01392	-10.03959	0.00000*

ตารางที่ 4.3 (ต่อ) ค่าเฉลี่ย ค่าส่วนเบี่ยงเบนมาตรฐาน ค่าสถิติที และค่าพี ของการประมาณค่าพารามิเตอร์ เมื่อ p เท่ากับ 0.8

n	r	วิธีประมาณ	Mean	SD	t-test	p-value
30	20	Bayes3	0.80087	0.01402	1.95400	0.05098
		MCMC	0.80060	0.01412	1.34576	0.17869
50	10	ML	0.80143	0.01561	2.89663	0.00385*
		Bayes1	0.79887	0.01538	-2.32464	0.02029*
		Bayes2	0.79476	0.01526	-10.87084	0.00000*
		Bayes3	0.80109	0.01538	2.23787	0.02545*
		MCMC	0.80079	0.01553	1.61119	0.10745
	20	ML	0.80101	0.01105	2.89316	0.00390*
		Bayes1	0.79973	0.01097	-0.78312	0.43374
		Bayes2	0.79766	0.01092	-6.77881	0.00000*
		Bayes3	0.80084	0.01097	2.43144	0.01521*
		MCMC	0.80069	0.01102	1.97055	0.04905*
	30	ML	0.80052	0.00886	1.86853	0.06198
		Bayes1	0.79967	0.00882	-1.18278	0.23718
		Bayes2	0.79829	0.00880	-6.15324	0.00000*
		Bayes3	0.80041	0.00882	1.48595	0.13761
		MCMC	0.80031	0.00885	1.09790	0.27251

หมายเหตุ * หมายความว่า ปฏิเสธสมมติฐานหลัก ที่ระดับนัยสำคัญเท่ากับ 0.05

จากตารางที่ 4.3 การประมาณค่าแบบจุด เมื่อ p เท่ากับ 0.8 พบว่าวิธีการประมาณที่ให้ค่าเฉลี่ยของตัวประมาณเท่ากับค่าพารามิเตอร์ของประชากร (p) ได้แก่

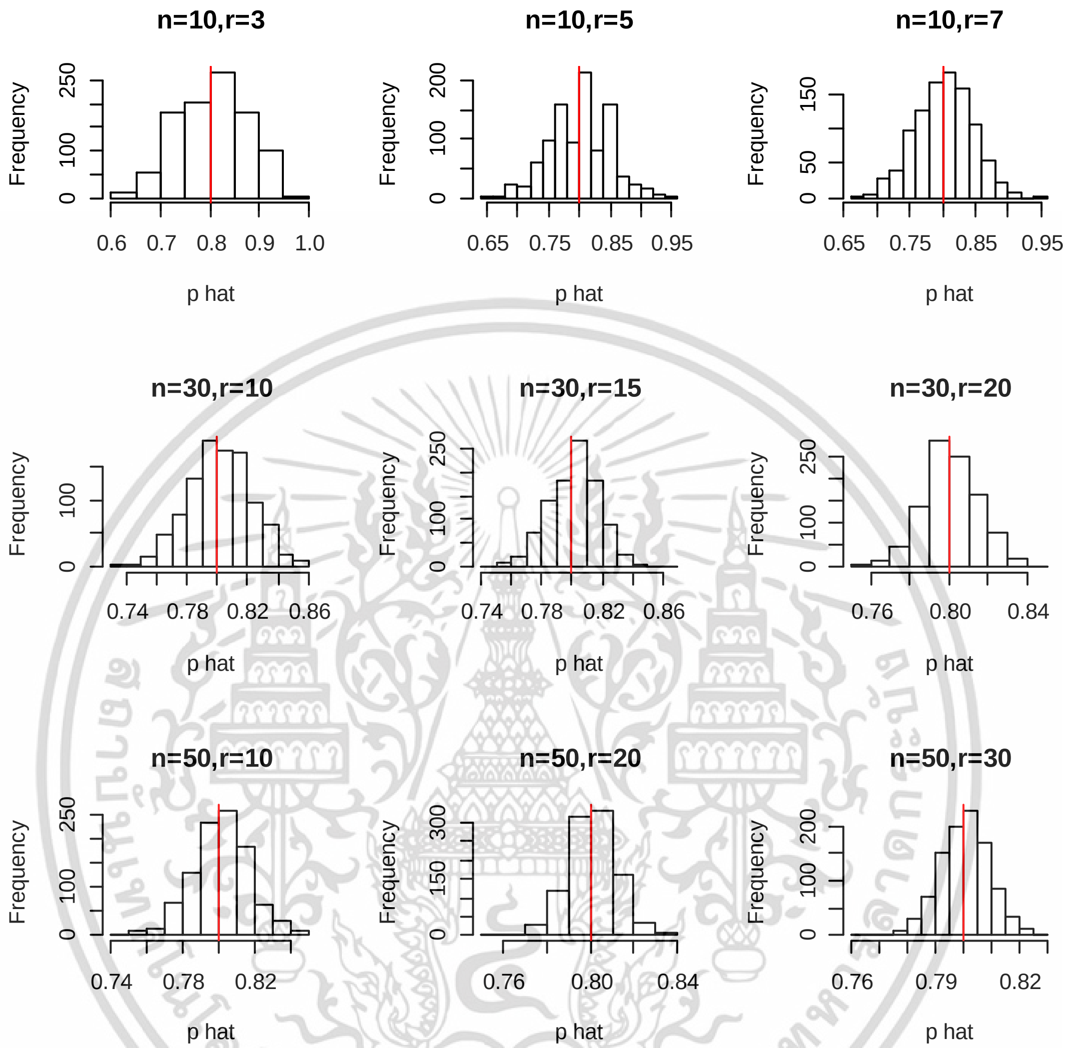
- วิธีภาวะน่าจะเป็นสูงสุด ที่ขนาดตัวอย่างเท่ากับ 10 ทุกค่าพารามิเตอร์ r และที่ขนาดตัวอย่างเท่ากับ 50 ค่าพารามิเตอร์ r = 30

- วิธีของเบส์มีการแจก แจงบีตา (3,3) เป็นการแจกแจงก่อน ที่ขนาดตัวอย่างเท่ากับ 50 ค่าพารามิเตอร์ r = 20 และ 30

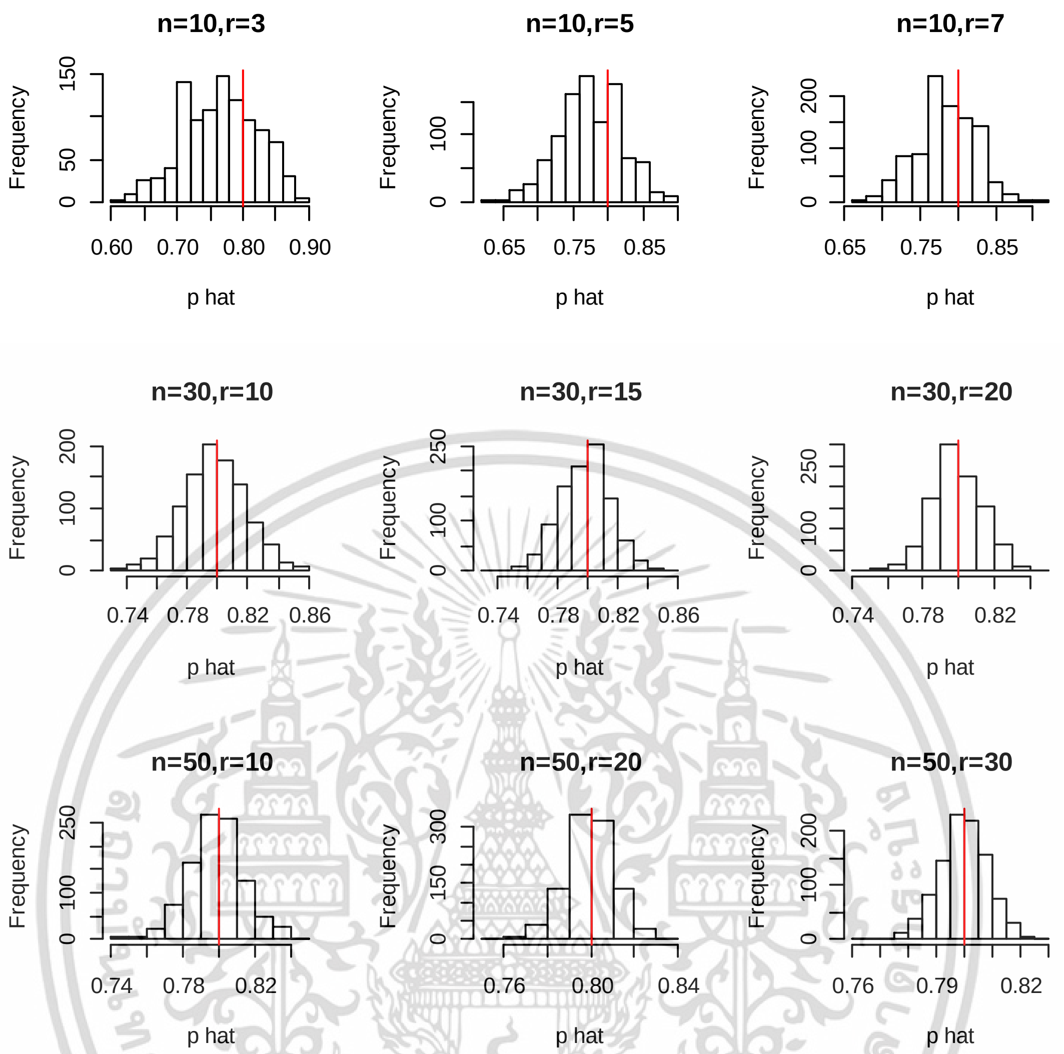
- วิธีของเบส์มีการแจก แจงบีตา (6,2) เป็นการแจกแจงก่อน ที่ขนาดตัวอย่างเท่ากับ 10 ค่าพารามิเตอร์ r = 5 และ 7 ขนาดตัวอย่างเท่ากับ 30 ทุกค่าพารามิเตอร์ r และที่ขนาดตัวอย่างเท่ากับ 50 ค่าพารามิเตอร์ r = 30

- วิธีมาร์คอฟ เชน มอนติคาร์โล ที่ขนาดตัวอย่างเท่ากับ 30 ทุกค่าพารามิเตอร์ r และขนาดตัวอย่างเท่ากับ 50 ค่าพารามิเตอร์ r = 10 และ 30

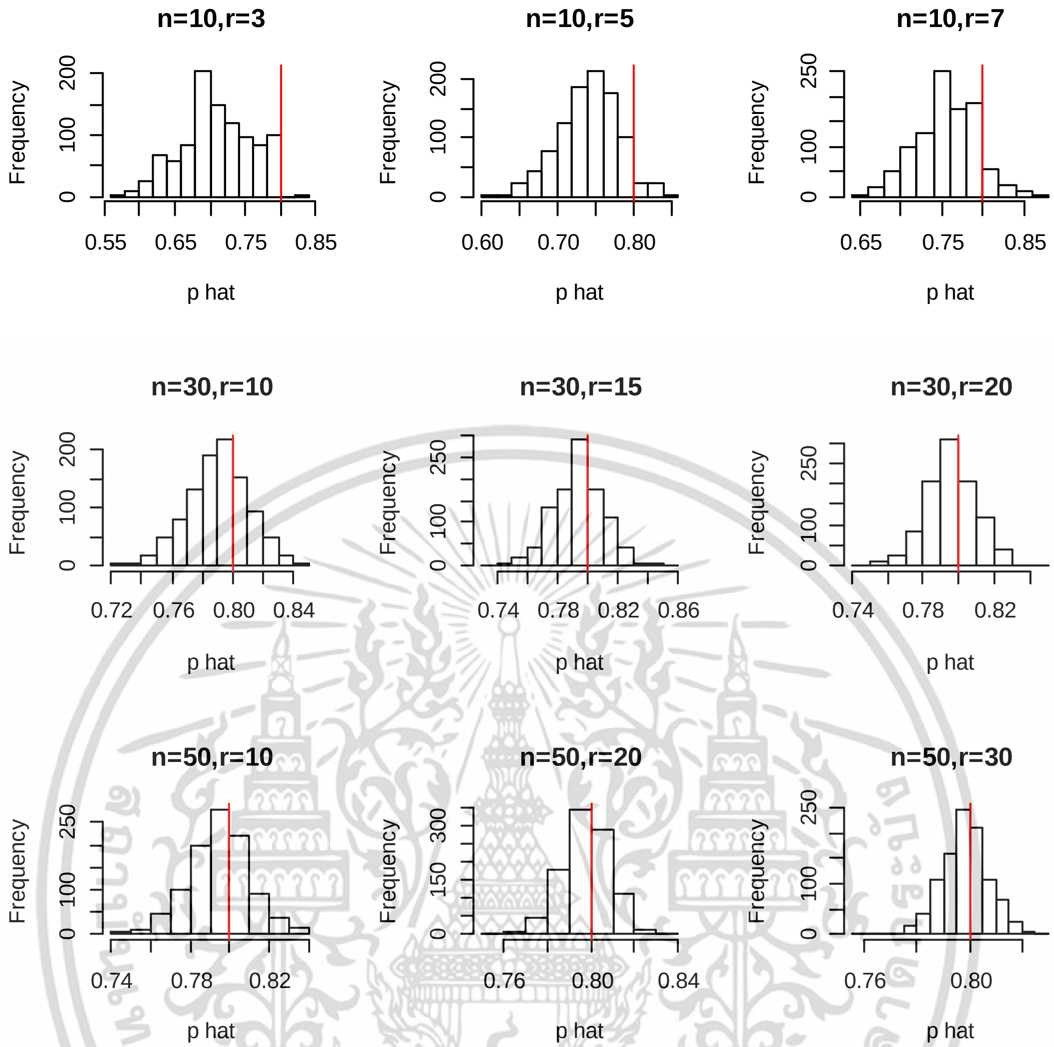
จากรูปที่ 4.11 – 4.15 แสดงให้เห็นว่า ค่าที่ประมาณได้จากทุกวิธีการประมาณจะมีแนวโน้มเข้าสู่การแจกแจงปรกติ เมื่อขนาดตัวอย่างเพิ่มขึ้น



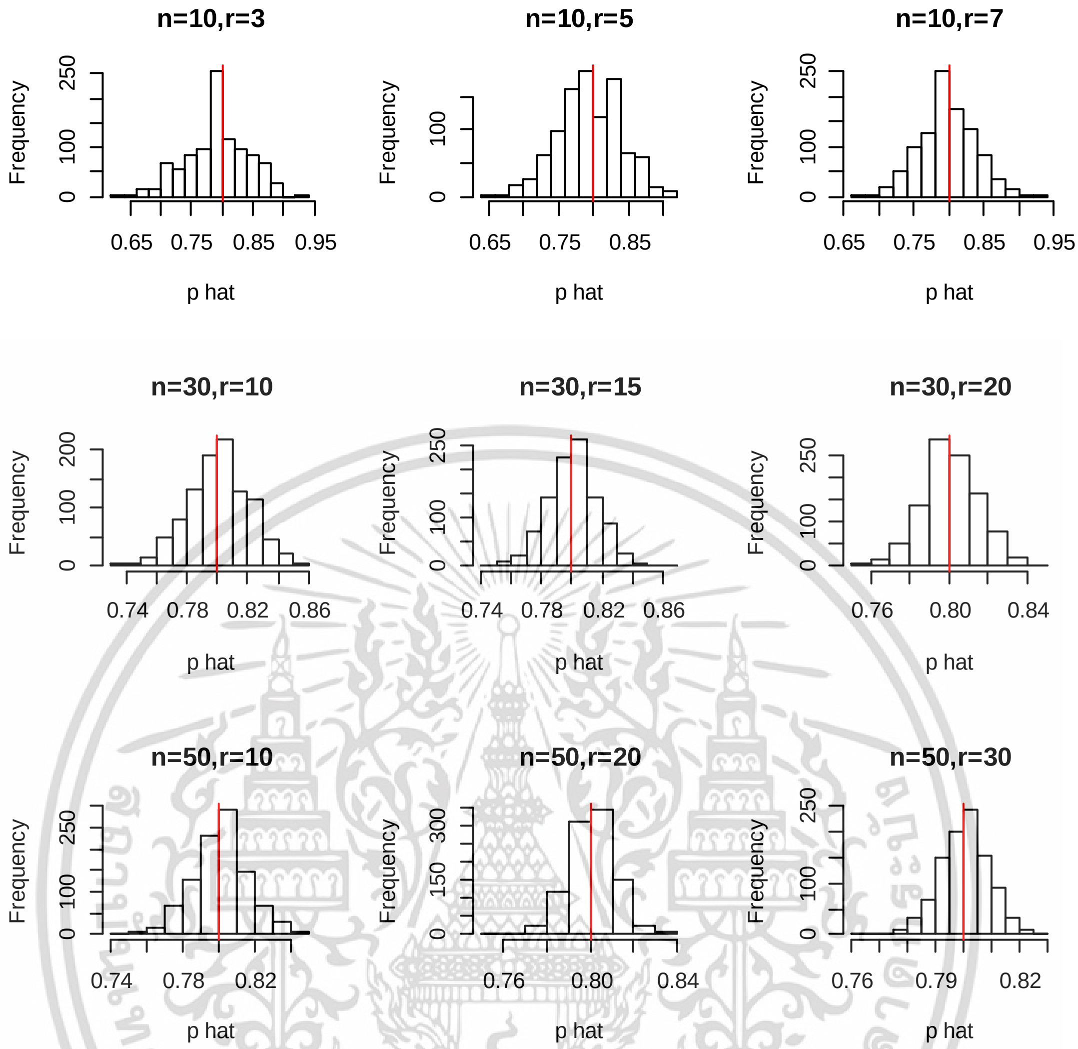
รูปที่ 4.11 แผนภาพฮิสโทแกรมแสดงค่าที่ประมาณได้ด้วยวิธีภาวะน่าจะเป็นสูงสุดเมื่อ $p = 0.8$



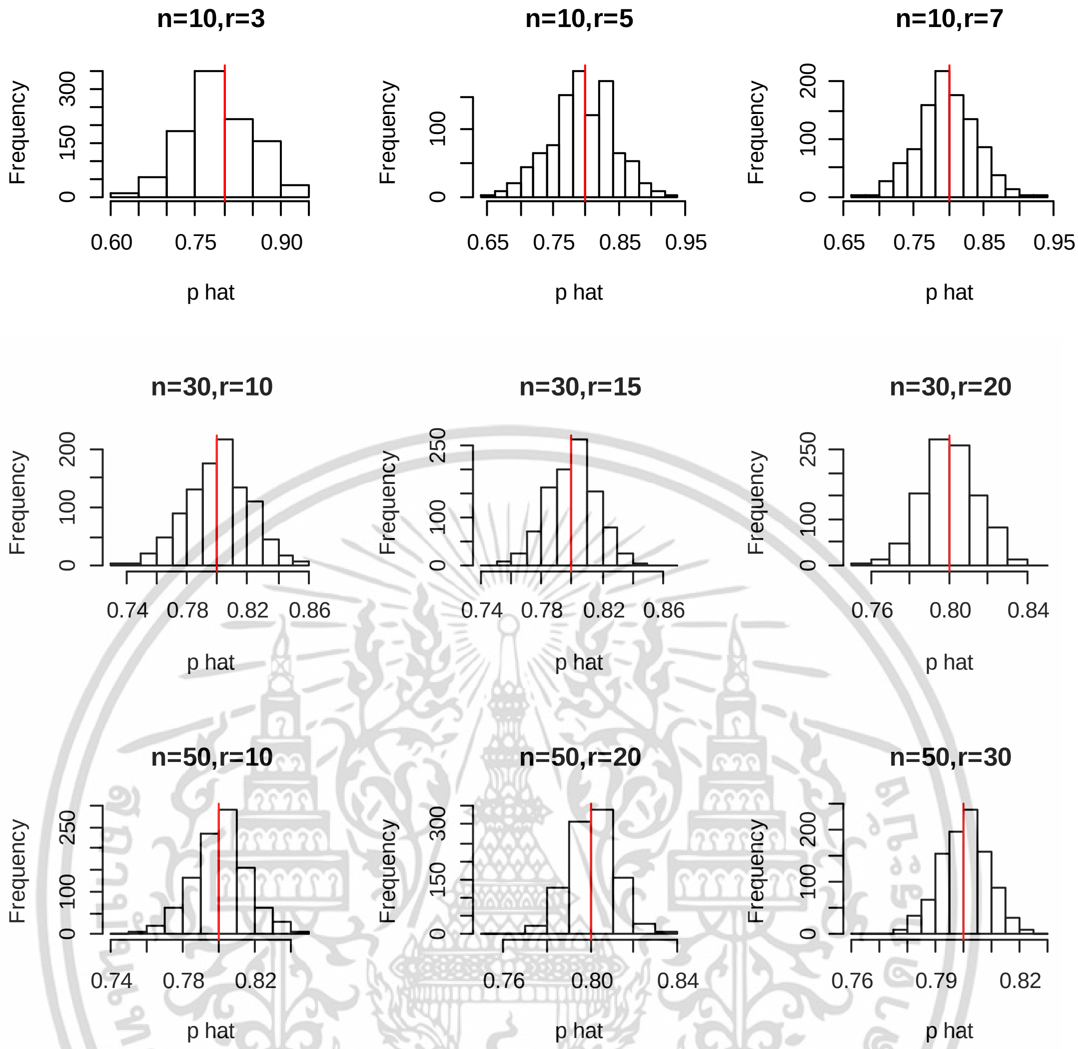
รูปที่ 4.12 แผนภาพฮิสโทแกรมแสดงค่าที่ประมาณได้ด้วยวิธีของเบส์ที่มีการแจกแจงบีตา (3,3) เป็นการแจกแจงก่อน เมื่อ $p = 0.8$



รูปที่ 4.13 แผนภาพฮิสโทแกรมแสดงค่าที่ประมาณได้ด้วยวิธีของเบส์ที่มีการแจกแจงบีตา (2,6) เป็นการแจกแจงก่อน เมื่อ $p = 0.8$



รูปที่ 4.14 แผนภาพฮิสโทแกรมแสดงค่าที่ประมาณได้ด้วยวิธีของเบส์ที่มีการแจกแจงบีตา (6,2) เป็นการแจกแจงก่อน เมื่อ $p = 0.8$



รูปที่ 4.15 แผนภาพฮิสโทแกรมแสดงค่าที่ประมาณได้ด้วยวิธีมาร์คอฟ เชน มอนติคาร์โล เมื่อ p = 0.8

การประมาณค่าแบบจุดด้วยวิธีของเบส์นั้นขึ้นอยู่กับการกำหนดฟังก์ชันการแจกแจงก่อน เมื่อรูปแบบของการแจกแจงก่อนเปลี่ยนแปลงไป ก็ส่งผลต่อการประมาณค่าด้วย จากผลการศึกษาพบว่าวิธีของเบส์จะให้ค่าประมาณที่เหมาะสม ดังรูปที่ 3.1

- เมื่อรูปแบบของการแจกแจงก่อนสมมาตร ที่ p = 0.5
- เมื่อรูปแบบของการแจกแจงก่อนเบ้ขวา ที่ p = 0.2
- เมื่อรูปแบบของการแจกแจงก่อนเบ้ซ้าย ที่ p = 0.8

4.2 การประมาณค่าแบบช่วง (Interval Estimation)

ผู้วิจัยทำการเปรียบเทียบค่าประมาณสัมประสิทธิ์ความเชื่อมั่นในแต่ละสถานการณ์กับเกณฑ์ตัดสินใจของค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด ดังนี้

- ที่ระดับนัยสำคัญ 90% ค่าสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่า 0.8814
- ที่ระดับนัยสำคัญ 95% ค่าสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่า 0.9365
- ที่ระดับนัยสำคัญ 99% ค่าสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่า 0.9838

ถ้าแต่ละสถานการณ์มีวิธีการประมาณค่าที่ให้ค่าสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนดหลายวิธี จะนำค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นที่ครอบคลุมค่าพารามิเตอร์มาพิจารณาด้วย โดยวิธีการประมาณใดให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นน้อยที่สุดจะถือว่าวิธีการประมาณนั้นให้ช่วงความเชื่อมั่นที่เหมาะสมที่สุดสำหรับสถานการณ์นั้น โดยการคำนวณค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่น จะคำนวณจากช่วงความเชื่อมั่นที่ครอบคลุมค่าพารามิเตอร์ที่แท้จริง โดยหาผลต่างระหว่างค่าขอบเขตบนและขอบเขตล่างของช่วงความเชื่อมั่นแล้วบวกสะสมไว้เพื่อหาค่าเฉลี่ยเมื่อคำนวณช่วงความเชื่อมั่นครบ 1,000 ครั้ง

ตารางที่ 4.4 ค่าสัมประสิทธิ์ความเชื่อมั่น (CC) และความกว้างเฉลี่ยของช่วงความเชื่อมั่น (AW) ที่ได้จากการประมาณ ที่ระดับความเชื่อมั่น 90% 95% และ 99% ตามลำดับ เมื่อ p เท่ากับ 0.2

n	r	วิธีประมาณ	ระดับความเชื่อมั่น					
			90%		95%		99%	
			CC	AW	CC	AW	CC	AW
30	10	ML	0.563	-	0.698	-	0.812	-
		Bayes1	0.885	0.03405	0.949	0.04045	0.993	0.05343
		Bayes2	0.889	**0.03397**	0.946	**0.04035**	0.994	**0.05331**
		Bayes3	0.878	-	0.947	0.04055	0.989	0.05357
		MCMC	0.885	0.03404	0.945	0.04045	0.994	0.05341
	15	ML	**0.553**	-	0.653	-	**0.779**	-
		Bayes1	0.895	0.02773	0.944	0.03320	0.989	0.04348
		Bayes2	0.895	**0.02769**	0.949	**0.03315**	0.988	**0.04341**
		Bayes3	0.887	0.02778	0.933	-	0.987	0.04355
		MCMC	0.896	0.02774	0.949	0.03321	0.987	0.04346
	20	ML	0.607	-	0.652	-	0.820	-
		Bayes1	0.899	0.02401	0.958	0.02866	0.996	0.03758
		Bayes2	0.900	**0.02398**	0.960	**0.02862**	0.995	**0.03753**
		Bayes3	0.889	0.02404	0.948	0.02869	0.995	0.03763
		MCMC	0.898	0.02401	0.961	0.02866	0.995	0.03757
50	10	ML	0.608	-	0.688	-	0.806	-
		Bayes1	0.885	**0.02637**	0.965	0.03142	0.991	0.04118
		Bayes2	0.880	-	0.962	**0.03137**	0.991	**0.04112**
		Bayes3	0.882	0.02641	0.961	0.03147	0.991	0.04124
		MCMC	0.881	-	0.962	0.03141	0.991	0.04118
	20	ML	0.561	-	0.673	-	0.802	-
		Bayes1	0.886	0.01863	0.961	0.02219	0.993	0.02915
		Bayes2	0.888	**0.01861**	0.960	**0.02217**	0.991	**0.02913**
		Bayes3	0.884	0.01864	0.961	0.02221	0.988	0.02917
		MCMC	0.890	0.01863	0.959	0.02218	0.991	0.02914
	30	ML	0.600	-	0.651	-	0.815	-
		Bayes1	0.906	0.01519	0.944	**0.01809**	0.988	0.02379

ตารางที่ 4.4 (ต่อ) ค่าสัมประสิทธิ์ความเชื่อมั่น (CC) และความกว้างเฉลี่ยของช่วงความเชื่อมั่น (AW) ที่ได้จากการประมาณ ที่ระดับความเชื่อมั่น 90% 95% และ 99% ตามลำดับ เมื่อ p เท่ากับ 0.2

n	r	วิธีประมาณ	ระดับความเชื่อมั่น					
			90%		95%		99%	
			CC	AW	CC	AW	CC	AW
50	30	Bayes2	0.905	**0.01518**	0.943	**0.01809**	0.988	**0.02378**
		Bayes3	0.904	0.01520	0.940	0.01810	0.988	0.02380
		MCMC	0.905	0.01519	0.944	**0.01809**	0.987	0.02379
70	15	ML	0.570	-	0.677	-	0.826	-
		Bayes1	0.876	-	0.951	0.02168	0.991	0.02850
		Bayes2	0.884	**0.01816**	0.948	**0.02166**	0.991	**0.02848**
		Bayes3	0.877	-	0.944	0.02169	0.987	0.02852
		MCMC	0.883	0.01817	0.948	0.02167	0.991	0.02851
	30	ML	0.601	-	0.686	-	0.816	-
		Bayes1	0.896	0.01285	0.962	0.01531	0.988	0.02011
		Bayes2	0.894	**0.01284**	0.961	**0.01530**	0.987	**0.02010**
		Bayes3	0.898	0.01285	0.962	0.01531	0.989	0.02012
		MCMC	0.895	0.01285	0.961	0.01531	0.988	0.02011
	45	ML	0.584	-	0.666	-	0.807	-
		Bayes1	0.905	**0.01049**	0.950	**0.01249**	0.987	**0.01643**
		Bayes2	0.905	**0.01049**	0.947	**0.01249**	0.987	**0.01643**
		Bayes3	0.904	**0.01049**	0.945	0.01250	0.987	**0.01643**
		MCMC	0.906	**0.01049**	0.950	0.01250	0.987	**0.01643**

หมายเหตุ: ขีดเส้นใต้ หมายถึง ค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการประมาณต่ำกว่าเกณฑ์ที่กำหนด

"-" หมายถึง ไม่สามารถคำนวณค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นได้ เนื่องจากให้ค่าสัมประสิทธิ์ความเชื่อมั่นต่ำกว่าเกณฑ์ที่กำหนด

ตัวหนา หมายถึง ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นแคบสุด ในแต่ละสถานการณ์และระดับความเชื่อมั่นที่กำหนด

จากตารางที่ 4.4 แสดงค่าสัมประสิทธิ์ความเชื่อมั่นและความกว้างเฉลี่ยของช่วงความเชื่อมั่นที่ได้จากการประมาณ ที่ระดับความเชื่อมั่น 90% 95% และ 99% ตามลำดับ ที่ขนาดตัวอย่าง 30 ค่าพารามิเตอร์ r เท่ากับ 10 15 และ 20 ขนาดตัวอย่าง 50 ค่าพารามิเตอร์ r เท่ากับ 10 20 และ 30 และขนาดตัวอย่าง 70 ค่าพารามิเตอร์ r เท่ากับ 15 30 และ 45 เมื่อ p เท่ากับ 0.2

พิจารณาค่าประมาณสัมประสิทธิ์ความเชื่อมั่น

- ที่ระดับความเชื่อมั่น 90% พบว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการประมาณ ส่วนใหญ่ไม่ต่ำกว่าเกณฑ์ที่กำหนด ยกเว้นค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการประมาณจากวิธีภาวะน่าจะเป็นสูงสุด ทุกสถานการณ์ วิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (3,3) ที่ขนาดตัวอย่าง 70 ค่าพารามิเตอร์ r = 15 วิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (2,6) ที่ขนาดตัวอย่าง 50 ค่าพารามิเตอร์ r = 10 วิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (6,2) ที่ขนาดตัวอย่าง 30 ค่าพารามิเตอร์ r = 10 และขนาดตัวอย่าง 70 ค่าพารามิเตอร์ r = 15 และวิธีมาร์คอฟ เชน มอนติคาร์โล ที่ขนาดตัวอย่าง 50 ค่าพารามิเตอร์ r = 10

- ที่ระดับความเชื่อมั่น 95% พบว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการประมาณ ส่วนใหญ่ไม่ต่ำกว่าเกณฑ์ที่กำหนด ยกเว้นค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการประมาณจากวิธีภาวะน่าจะเป็นสูงสุด ทุกสถานการณ์ และวิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (6,2) ที่ขนาดตัวอย่าง 30 ค่าพารามิเตอร์ r =15

- ที่ระดับความเชื่อมั่น 99% พบว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการประมาณ ส่วนใหญ่ไม่ต่ำกว่าเกณฑ์ที่กำหนด ยกเว้นค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการประมาณจากวิธีภาวะน่าจะเป็นสูงสุด ทุกสถานการณ์

พิจารณาค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่น

- ที่ระดับความเชื่อมั่น 90% พบว่า วิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (2,6) ส่วนใหญ่ให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นแคบที่สุด ยกเว้นที่ขนาดตัวอย่าง 50 และค่าพารามิเตอร์ r = 10 และวิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (3,3) วิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (6,2) และวิธีมาร์คอฟ เชน มอนติคาร์โล ให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นแคบที่สุด ที่ขนาดตัวอย่าง 70 ค่าพารามิเตอร์ r = 45

- ที่ระดับความเชื่อมั่น 95% พบว่า วิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (2,6) ให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นแคบที่สุดทุกสถานการณ์ และวิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (3,3) ที่ขนาดตัวอย่าง 50 ค่าพารามิเตอร์ r = 30 และขนาดตัวอย่าง 70 ค่าพารามิเตอร์ r = 45 และวิธีมาร์คอฟ เชน มอนติคาร์โล ที่ขนาดตัวอย่าง 50 ค่าพารามิเตอร์ r = 30 ให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นแคบที่สุด

- ที่ระดับความเชื่อมั่น 99% พบว่า วิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (2,6) ให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นแคบที่สุดทุกสถานการณ์ และวิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (3,3) วิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (6,2) และวิธีมาร์คอฟ เชน มอนติคาร์โล ให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นแคบที่สุด ที่ขนาดตัวอย่าง 70 ค่าพารามิเตอร์ r = 45

ตารางที่ 4.5 ค่าสัมประสิทธิ์ความเชื่อมั่น (CC) และความกว้างเฉลี่ยของช่วงความเชื่อมั่น (AW) ที่ได้จากการประมาณ ที่ระดับความเชื่อมั่น 90% 95% และ 99% ตามลำดับ เมื่อ p เท่ากับ 0.5

n	r	วิธีประมาณ	ระดับความเชื่อมั่น					
			90%		95%		99%	
			CC	AW	CC	AW	CC	AW
30	10	ML	0.913	0.06729	0.956	0.08007	0.990	0.10527
		Bayes1	0.915	0.06678	0.962	0.07952	0.993	0.10453
		Bayes2	0.914	**0.06667**	0.952	0.07939	0.990	**0.10435**
		Bayes3	0.897	**0.06667**	0.958	**0.07938**	0.993	**0.10435**
		MCMC	0.912	0.06698	0.959	0.07976	0.993	0.10486
	15	ML	0.909	0.05505	0.687	-	0.985	0.08596
		Bayes1	0.910	0.05469	0.954	0.03306	0.984	0.08552
		Bayes2	0.909	0.05464	0.951	**0.03300**	0.984	**0.08543**
		Bayes3	0.897	**0.05463**	0.954	0.03311	0.982	-
		MCMC	0.910	0.05478	0.951	0.03306	0.985	0.08574
	20	ML	0.886	0.04749	0.953	0.05656	0.991	0.07454
		Bayes1	0.891	0.04733	0.952	0.05639	0.989	0.07420
		Bayes2	0.887	**0.04729**	0.954	**0.05634**	0.991	**0.07414**
		Bayes3	0.888	**0.04729**	0.947	**0.05634**	0.988	**0.07414**
		MCMC	0.889	0.04740	0.954	0.05647	0.989	0.07432
50	10	ML	0.896	0.05194	0.956	0.06194	0.989	0.08144
		Bayes1	0.898	0.05177	0.960	0.06172	0.990	0.08112
		Bayes2	0.902	**0.05172**	0.955	**0.06165**	0.989	**0.08104**
		Bayes3	0.903	**0.05172**	0.952	**0.06165**	0.990	**0.08104**
		MCMC	0.897	0.05189	0.954	0.06184	0.989	0.08130
	20	ML	0.917	0.03676	0.949	0.04379	0.987	0.05766
		Bayes1	0.920	0.03670	0.949	0.04372	0.987	0.05751
		Bayes2	0.922	**0.03668**	0.949	**0.04370**	0.987	**0.05748**
		Bayes3	0.916	**0.03668**	0.948	**0.04370**	0.989	**0.05748**
		MCMC	0.919	0.03672	0.950	0.04377	0.987	0.05755
	30	ML	0.907	0.03005	0.946	0.03577	0.991	0.04704
		Bayes1	0.907	0.03000	0.947	0.03573	0.990	0.04697

ตารางที่ 4.5 (ต่อ) ค่าสัมประสิทธิ์ความเชื่อมั่น (CC) และความกว้างเฉลี่ยของช่วงความเชื่อมั่น (AW) ที่ได้จากการประมาณ ที่ระดับความเชื่อมั่น 90% 95% และ 99% ตามลำดับ เมื่อ p เท่ากับ 0.5

n	r	วิธีประมาณ	ระดับความเชื่อมั่น					
			90%		95%		99%	
			CC	AW	CC	AW	CC	AW
50	30	Bayes2	0.905	**0.02999**	0.945	**0.03572**	0.991	**0.04695**
		Bayes3	0.910	**0.02999**	0.940	**0.03572**	0.991	**0.04695**
		MCMC	0.905	0.03001	0.944	0.03576	0.990	0.04699
70	15	ML	0.898	0.03586	0.939	0.04279	0.992	0.05616
		Bayes1	0.903	0.03581	0.939	0.04270	0.991	0.05608
		Bayes2	0.898	**0.03579**	0.939	**0.04268**	0.992	**0.05605**
		Bayes3	0.900	**0.03579**	0.940	**0.04268**	0.993	**0.05605**
		MCMC	0.904	0.03583	0.939	0.04275	0.991	0.05613
	30	ML	0.895	0.02540	0.949	0.03024	0.990	0.03976
		Bayes1	0.895	0.02537	0.948	0.03021	0.988	**0.03971**
		Bayes2	0.894	**0.02536**	0.949	**0.03020**	0.989	**0.03971**
		Bayes3	0.895	**0.02536**	0.950	**0.03020**	0.988	**0.03971**
		MCMC	0.896	**0.02536**	0.949	0.03022	0.988	0.03974
	45	ML	0.892	0.02073	0.947	0.02469	0.989	0.03247
		Bayes1	0.892	**0.02071**	0.946	0.02468	0.990	**0.03244**
		Bayes2	0.890	**0.02071**	0.945	**0.02467**	0.989	**0.03244**
		Bayes3	0.895	**0.02071**	0.942	**0.02467**	0.989	**0.03244**
		MCMC	0.893	0.02072	0.945	0.02469	0.989	0.03245

หมายเหตุ: ขีดเส้นใต้ หมายถึง ค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการประมาณต่ำกว่าเกณฑ์ที่กำหนด

"-" หมายถึง ไม่สามารถคำนวณค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นได้ เนื่องจากให้ค่าสัมประสิทธิ์ความเชื่อมั่นต่ำกว่าเกณฑ์ที่กำหนด

ตัวหนา หมายถึง ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นแคบสุด ในแต่ละสถานการณ์และระดับความเชื่อมั่นที่กำหนด

จากตารางที่ 4.5 แสดงค่าสัมประสิทธิ์ความเชื่อมั่นและความกว้างเฉลี่ยของช่วงความเชื่อมั่นที่ได้จากการประมาณ ที่ระดับความเชื่อมั่น 90% 95% และ 99% ตามลำดับ ที่ขนาดตัวอย่าง 30 ค่าพารามิเตอร์ r เท่ากับ 10 15 และ 20 ขนาดตัวอย่าง 50 ค่าพารามิเตอร์ r เท่ากับ 10 20 และ 30 และขนาดตัวอย่าง 70 ค่าพารามิเตอร์ r เท่ากับ 15 30 และ 45 เมื่อ p เท่ากับ 0.5

พิจารณาค่าประมาณสัมประสิทธิ์ความเชื่อมั่น

- ที่ระดับความเชื่อมั่น 90% พบว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการประมาณไม่ต่ำกว่าเกณฑ์ที่กำหนด ทุกสถานการณ์

- ที่ระดับความเชื่อมั่น 95% พบว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการประมาณ ส่วนใหญ่ไม่ต่ำกว่าเกณฑ์ที่กำหนด ยกเว้นค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการประมาณจากวิธีภาวะน่าจะเป็นสูงสุด ที่ขนาดตัวอย่าง 30 ค่าพารามิเตอร์ r = 15

- ที่ระดับความเชื่อมั่น 99% พบว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการประมาณ ส่วนใหญ่ไม่ต่ำกว่าเกณฑ์ที่กำหนด ยกเว้นค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการประมาณจากวิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (6,2) ที่ขนาดตัวอย่าง 30 ค่าพารามิเตอร์ r = 15

พิจารณาค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่น

- ที่ระดับความเชื่อมั่น 90% พบว่า วิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (6,2) ให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นแคบที่สุดทุกสถานการณ์ สำหรับวิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (2,6) ส่วนใหญ่ให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นแคบที่สุด ยกเว้นที่ขนาดตัวอย่าง 30 และค่าพารามิเตอร์ r = 15 และวิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (3,3) ที่ขนาดตัวอย่าง 70 ค่าพารามิเตอร์ r = 45 และวิธีมาร์คอฟ เชน มอนติคาร์โล ที่ขนาดตัวอย่าง 70 ค่าพารามิเตอร์ r = 30 ให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นแคบที่สุด

- ที่ระดับความเชื่อมั่น 95% พบว่า วิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (2,6) และ (6,2) ส่วนใหญ่ให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นแคบที่สุด ยกเว้นที่ขนาดตัวอย่าง 30 ค่าพารามิเตอร์ r = 10 และขนาดตัวอย่าง 30 ค่าพารามิเตอร์ r = 15 ตามลำดับ

- ที่ระดับความเชื่อมั่น 99% พบว่า วิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (2,6) ให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นแคบที่สุดทุกสถานการณ์ และวิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (6,2) ส่วนใหญ่ให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นแคบที่สุด ยกเว้นที่ขนาดตัวอย่าง 30 ค่าพารามิเตอร์ r =15 และวิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (3,3) ให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นแคบที่สุด ที่ขนาดตัวอย่าง 70 ค่าพารามิเตอร์ r = 30 และ 45

ตารางที่ 4.6 ค่าสัมประสิทธิ์ความเชื่อมั่น (CC) และความกว้างเฉลี่ยของช่วงความเชื่อมั่น (AW) ที่ได้จากการประมาณ ที่ระดับความเชื่อมั่น 90% 95% และ 99% ตามลำดับ เมื่อ p เท่ากับ 0.8

n	r	วิธีประมาณ	ระดับความเชื่อมั่น					
			90%		95%		99%	
			CC	AW	CC	AW	CC	AW
30	10	ML	0.997	0.13608	0.999	0.16204	1.000	0.21359
		Bayes1	0.908	0.06773	0.962	0.08073	0.990	**0.10598**
		Bayes2	0.858	-	0.936	-	0.990	0.10700
		Bayes3	0.907	**0.06709**	0.958	**0.07997**	0.983	-
		MCMC	0.908	0.06766	0.958	0.08065	0.983	-
	15	ML	1.000	0.11131	1.000	0.13263	1.000	0.17381
		Bayes1	0.894	0.05531	0.959	0.06591	0.984	0.08671
		Bayes2	0.886	0.05568	0.951	0.06635	0.977	-
		Bayes3	0.894	**0.05496**	0.954	**0.06550**	0.988	**0.08617**
		MCMC	0.895	0.05528	0.951	0.06588	0.985	0.08664
	20	ML	0.998	0.09601	1.000	0.11484	1.000	0.15072
		Bayes1	0.895	0.04800	0.950	0.05711	0.991	0.07510
		Bayes2	0.882	0.04824	0.943	0.05740	0.983	-
		Bayes3	0.895	**0.04777**	0.950	**0.05684**	0.990	**0.07474**
		MCMC	0.894	0.04796	0.949	0.05707	0.990	0.07509
50	10	ML	0.999	0.10551	1.000	0.12544	1.000	0.16524
		Bayes1	0.889	0.05250	0.946	0.06261	0.992	0.08222
		Bayes2	0.880	-	0.940	0.06298	0.992	0.08271
		Bayes3	0.898	**0.05220**	0.952	**0.06226**	0.988	**0.08175**
		MCMC	0.904	0.05248	0.944	0.06261	0.989	0.08217
	20	ML	1.000	0.07441	1.000	0.08872	1.000	0.11676
		Bayes1	0.904	0.03719	0.958	0.04431	0.991	0.05820
		Bayes2	0.889	0.03731	0.951	0.04444	0.988	0.05838
		Bayes3	0.904	**0.03709**	0.958	**0.04418**	0.989	**0.05803**
		MCMC	0.899	0.03719	0.955	0.04429	0.990	0.05818
	30	ML	0.999	0.06073	1.000	0.07244	1.000	0.09533
		Bayes1	0.902	0.03038	0.948	0.03619	0.994	0.04753

ตารางที่ 4.6 (ต่อ) ค่าสัมประสิทธิ์ความเชื่อมั่น (CC) และความกว้างเฉลี่ยของช่วงความเชื่อมั่น (AW) ที่ได้จากการประมาณ ที่ระดับความเชื่อมั่น 90% 95% และ 99% ตามลำดับ เมื่อ p เท่ากับ 0.8

n	r	วิธีประมาณ	ระดับความเชื่อมั่น					
			90%		95%		99%	
			CC	AW	CC	AW	CC	AW
50	30	Bayes2	0.895	0.03044	0.938	0.03626	0.996	0.04763
		Bayes3	0.903	**0.03032**	0.947	**0.03612**	0.994	**0.04744**
		MCMC	0.907	0.03038	0.948	0.03619	0.994	0.04752
70	15	ML	0.995	0.07270	1.000	0.08664	1.000	0.11379
		Bayes1	0.903	0.03628	0.944	0.04323	0.996	0.05683
		Bayes2	0.893	0.03639	0.941	0.04335	0.994	0.05699
		Bayes3	0.902	**0.03618**	0.951	**0.04311**	0.995	**0.05667**
		MCMC	0.897	0.03626	0.950	0.04322	0.995	0.05683
	30	ML	0.999	0.05141	1.000	0.06125	1.000	0.08039
		Bayes1	0.886	0.02566	0.960	0.03058	0.991	0.04021
		Bayes2	<u>0.879</u>	-	0.955	0.03063	0.992	0.04027
		Bayes3	0.891	**0.02563**	0.959	**0.03054**	0.989	**0.04016**
		MCMC	0.887	0.02566	0.961	0.03058	0.990	0.04020
	45	ML	0.999	0.04198	1.000	0.04995	1.000	0.06564
		Bayes1	0.917	0.02096	0.963	0.02499	0.994	0.03284
		Bayes2	0.912	0.02098	0.964	0.02501	0.993	0.03287
		Bayes3	0.920	**0.02094**	0.966	**0.02496**	0.993	**0.03281**
		MCMC	0.919	0.02096	0.965	0.02498	0.993	0.03283

<u>หมายเหตุ</u>: ขีดเส้นใต้ หมายถึง ค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการประมาณต่ำกว่าเกณฑ์ที่กำหนด

"-" หมายถึง ไม่สามารถคำนวณค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นได้ เนื่องจากให้ค่าสัมประสิทธิ์ความเชื่อมั่นต่ำกว่าเกณฑ์ที่กำหนด

ตัวหนา หมายถึง ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นแคบสุด ในแต่ละสถานการณ์และระดับความเชื่อมั่นที่กำหนด

จากตารางที่ 4.6 แสดงค่าสัมประสิทธิ์ความเชื่อมั่นและความกว้างเฉลี่ยของช่วงความเชื่อมั่นที่ได้จากการประมาณ ที่ระดับความเชื่อมั่น 90% 95% และ 99% ตามลำดับ ที่ขนาดตัวอย่าง 30 ค่าพารามิเตอร์ r เท่ากับ 10 15 และ 20 ขนาดตัวอย่าง 50 ค่าพารามิเตอร์ r เท่ากับ 10 20 และ 30 และขนาดตัวอย่าง 70 ค่าพารามิเตอร์ r เท่ากับ 15 30 และ 45 เมื่อ p เท่ากับ 0.8

พิจารณาค่าประมาณสัมประสิทธิ์ความเชื่อมั่น

- ที่ระดับความเชื่อมั่น 90% พบว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการประมาณ ส่วนใหญ่ไม่ต่ำกว่าเกณฑ์ที่กำหนด ยกเว้นค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการประมาณจากวิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (2,6) ที่ขนาดตัวอย่าง 30 50 ค่าพารามิเตอร์ r = 10 และขนาดตัวอย่าง 70 ค่าพารามิเตอร์ r = 30

- ที่ระดับความเชื่อมั่น 95% พบว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการประมาณ ส่วนใหญ่ไม่ต่ำกว่าเกณฑ์ที่กำหนด ยกเว้นค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการประมาณจากวิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (2,6) ที่ขนาดตัวอย่าง 30 ค่าพารามิเตอร์ r = 10

- ที่ระดับความเชื่อมั่น 99% พบว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการประมาณ ส่วนใหญ่ไม่ต่ำกว่าเกณฑ์ที่กำหนด ยกเว้นค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการประมาณจากวิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (2,6) ที่ขนาดตัวอย่าง 30 ค่าพารามิเตอร์ r = 15 และ 20 วิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (6,2) ที่ขนาดตัวอย่าง 30 ค่าพารามิเตอร์ r = 10 และวิธีมาร์คอฟ เชน มอนติคาร์โล ที่ขนาดตัวอย่าง 30 ค่าพารามิเตอร์ r = 10

พิจารณาค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่น

- ที่ระดับความเชื่อมั่น 90% และ 95% พบว่า วิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (6,2) ให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นแคบที่สุดทุกสถานการณ์

- ที่ระดับความเชื่อมั่น 99% พบว่า วิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (6,2) ส่วนใหญ่ให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นแคบที่สุด ยกเว้นที่ขนาดตัวอย่าง 30 ค่าพารามิเตอร์ r = 10

บทที่ 5
สรุปผลการวิจัยและข้อเสนอแนะ

การวิจัยนี้มีจุดมุ่งหมายเพื่อศึกษาวิธีการประมาณค่าแบบจุดและแบบช่วงสำหรับพารามิเตอร์ p ของการแจกแจงทวินามเชิงลบ ด้วยวิธีภาวะน่าจะเป็นสูงสุด วิธีของเบส์ ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา และวิธีมาร์คอฟ เชน มอนติคาร์โล โดยศึกษาและเปรียบเทียบในแต่ละสถานการณ์ดังนี้

1. กำหนดค่าพารามิเตอร์ของประชากร (p) เท่ากับ 0.2 0.5 และ 0.8 ตามลำดับ

2. กำหนดค่าพารามิเตอร์ r และขนาดตัวอย่าง (n) สำหรับการประมาณค่าแบบจุด แสดงดังตารางที่ 5.1

ตารางที่ 5.1 ค่าพารามิเตอร์สำหรับการประมาณค่าแบบจุด

n	r
10	3 5 7
30	10 15 20
50	10 20 30

3. กำหนดค่าพารามิเตอร์ r และขนาดตัวอย่าง (n) สำหรับการประมาณค่าแบบช่วง แสดงดังตารางที่ 5.2

ตารางที่ 5.2 ค่าพารามิเตอร์สำหรับการประมาณค่าแบบช่วง

n	r
30	10 15 20
50	10 20 30
70	15 30 45

4. กำหนดระดับความเชื่อมั่น 3 ระดับ คือ 90% 95% และ 99%

สำหรับการประมาณค่าแบบจุด ผู้วิจัยทำการทดสอบสมมติฐานการเท่ากันระหว่างค่าเฉลี่ยความน่าจะเป็นของความสำเร็จในการทดลองแต่ละครั้งของตัวอย่างกับค่าพารามิเตอร์ของประชากร โดยใช้การทดสอบที และสำหรับการประมาณค่าแบบช่วง ผู้วิจัยได้ทำการตรวจสอบค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการประมาณและเปรียบเทียบค่าเฉลี่ยความกว้างของช่วงความเชื่อมั่น เพื่อหาวิธีการประมาณที่ดีที่สุดในแต่ละสถานการณ์

5.1 สรุปผลการวิจัย

สัญลักษณ์แทนความหมายต่างๆ ดังนี้

- ML แทน วิธีภาวะน่าจะเป็นสูงสุด
- Bayes1 แทน วิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (3,3)
- Bayes2 แทน วิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (2,6)
- Bayes3 แทน วิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (6,2)
- MCMC แทน วิธีมาร์คอฟ เชน มอนติคาร์โล

ตารางที่ 5.3 วิธีประมาณค่าพารามิเตอร์แบบจุดที่ให้ค่าเฉลี่ยความน่าจะเป็นของความสำเร็จในการทดลองแต่ละครั้งของตัวอย่างเท่ากับค่าพารามิเตอร์ของประชากร จำแนกตามขนาดตัวอย่าง (n) ค่าพารามิเตอร์ r และค่าพารามิเตอร์ของประชากร (p)

p	n = 10			n = 30			n = 50		
	r = 3	r = 5	r = 7	r = 10	r = 15	r = 20	r = 10	r = 20	r = 30
0.2	ML	ML	ML	ML	ML	ML	ML Bayes2 MCMC	ML Bayes2	ML Bayes2
0.5	ML	ML	ML	ML Bayes1 MCMC	ML Bayes1 MCMC	ML Bayes1 Bayes2 MCMC	ML Bayes2	ML Bayes1 MCMC	ML Bayes1 Bayes2 MCMC
0.8	ML	ML Bayes3	ML Bayes3	Bayes3 MCMC	Bayes3 MCMC	Bayes3 MCMC	MCMC	Bayes1	ML Bayes1 Bayes3 MCMC

เมื่อนำค่าเฉลี่ยความน่าจะเป็นของความสำเร็จในการทดลองแต่ละครั้งของตัวอย่างที่ได้จากวิธีการประมาณทั้ง 5 วิธี มาทดสอบสมมติฐานการเท่ากันกับค่าพารามิเตอร์ของประชากร (p) โดยใช้การทดสอบที จะได้วิธีประมาณที่ให้ค่าเฉลี่ยความน่าจะเป็นของความสำเร็จในการทดลองแต่ละครั้งของตัวอย่างเท่ากับค่าพารามิเตอร์ของประชากร ดังตารางที่ 5.3 สรุปได้ดังนี้

ที่พารามิเตอร์ของประชากร (p) เท่ากับ 0.2

- วิธีภาวะน่าจะเป็นสูงสุด ให้ค่าเฉลี่ยความน่าจะเป็นของความสำเร็จในการทดลองแต่ละครั้งของตัวอย่างเท่ากับค่าพารามิเตอร์ของประชากร ในทุกสถานการณ์

- วิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (2,6) ให้ค่าเฉลี่ยความน่าจะเป็นของความสำเร็จในการทดลองแต่ละครั้งของตัวอย่างเท่ากับค่าพารามิเตอร์ของประชากร เมื่อขนาดตัวอย่างเท่ากับ 50 ที่ทุกค่าพารามิเตอร์ r

- วิธีมาร์คอฟ เชน มอนติคาร์โลให้ค่าเฉลี่ยความน่าจะเป็นของความสำเร็จในการทดลองแต่ละครั้งของตัวอย่างเท่ากับค่าพารามิเตอร์ของประชากร เมื่อขนาดตัวอย่างเท่ากับ 50 และค่าพารามิเตอร์ r เท่ากับ 10

ที่พารามิเตอร์ของประชากร (p) เท่ากับ 0.5

- วิธีภาวะน่าจะเป็นสูงสุด ให้ค่าเฉลี่ยความน่าจะเป็นของความสำเร็จในการทดลองแต่ละครั้งของตัวอย่างเท่ากับค่าพารามิเตอร์ของประชากร ในทุกสถานการณ์

- วิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (3,3) ให้ค่าเฉลี่ยความน่าจะเป็นของความสำเร็จในการทดลองแต่ละครั้งของตัวอย่างเท่ากับค่าพารามิเตอร์ของประชากร เมื่อขนาดตัวอย่างเท่ากับ 30 ที่ทุกค่าพารามิเตอร์ r และเมื่อขนาดตัวอย่างเท่ากับ 50 ค่าพารามิเตอร์ r เท่ากับ 20 และ 30

- วิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (2,6) ให้ค่าเฉลี่ยความน่าจะเป็นของความสำเร็จในการทดลองแต่ละครั้งของตัวอย่างเท่ากับค่าพารามิเตอร์ของประชากร เมื่อขนาดตัวอย่างเท่ากับ 30 ค่าพารามิเตอร์ r เท่ากับ 20 และขนาดตัวอย่างเท่ากับ 50 ค่าพารามิเตอร์ r เท่ากับ 10 และ 30

- วิธีมาร์คอฟ เชน มอนติคาร์โล ให้ค่าเฉลี่ยความน่าจะเป็นของความสำเร็จในการทดลองแต่ละครั้งของตัวอย่างเท่ากับค่าพารามิเตอร์ของประชากร เมื่อขนาดตัวอย่างเท่ากับ 30 ที่ทุกค่าพารามิเตอร์ r และขนาดตัวอย่างเท่ากับ 50 ค่าพารามิเตอร์ r เท่ากับ 20 และ 30

ที่พารามิเตอร์ของประชากร (p) เท่ากับ 0.8

- วิธีภาวะน่าจะเป็นสูงสุด ให้ค่าเฉลี่ยความน่าจะเป็นของความสำเร็จในการทดลองแต่ละครั้งของตัวอย่างเท่ากับค่าพารามิเตอร์ของประชากร เมื่อขนาดตัวอย่างเท่ากับ 10 ทุกค่าพารามิเตอร์ r และขนาดตัวอย่างเท่ากับ 50 ค่าพารามิเตอร์ r เท่ากับ 30

- วิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (3,3) ให้ค่าเฉลี่ยความน่าจะเป็นของความสำเร็จในการทดลองแต่ละครั้งของตัวอย่างเท่ากับค่าพารามิเตอร์ของประชากร เมื่อขนาดตัวอย่างเท่ากับ 50 ค่าพารามิเตอร์ r เท่ากับ 20 และ 30

- วิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (6,2) ให้ค่าเฉลี่ยความน่าจะเป็นของความสำเร็จในการทดลองแต่ละครั้งของตัวอย่างเท่ากับค่าพารามิเตอร์ของประชากร เมื่อขนาดตัวอย่างเท่ากับ 10 ค่าพารามิเตอร์ r เท่ากับ 5 และ 7 ขนาดตัวอย่างเท่ากับ 30 ที่ทุกค่าพารามิเตอร์ r และขนาดตัวอย่างเท่ากับ 50 ค่าพารามิเตอร์ r เท่ากับ 30

- วิธีมาร์คอฟ เชน มอนติคาร์โล ให้ค่าเฉลี่ยความน่าจะเป็นของความสำเร็จในการทดลองแต่ละครั้งของตัวอย่างเท่ากับค่าพารามิเตอร์ของประชากร เมื่อขนาดตัวอย่างเท่ากับ 30 ที่ทุกค่าพารามิเตอร์ r และขนาดตัวอย่างเท่ากับ 50 ค่าพารามิเตอร์ r เท่ากับ 10 และ 30

เมื่อนำสถานการณ์ที่ตัวประมาณให้ค่าสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่าเกณฑ์ที่กำหนดมาเปรียบเทียบค่าเฉลี่ยความกว้างของช่วงความเชื่อมั่น จะได้วิธีประมาณที่ให้ค่าเฉลี่ยความกว้างของช่วงความเชื่อมั่นที่แคบที่สุด ดังตารางที่ 5.4 ถึง 5.6 สรุปได้ดังนี้

ตารางที่ 5.4 วิธีประมาณค่าพารามิเตอร์แบบช่วงที่ให้ค่าเฉลี่ยความกว้างของช่วงความเชื่อมั่นแคบสุด จำแนกตามขนาดตัวอย่าง (n) ค่าพารามิเตอร์ r และระดับความเชื่อมั่น เมื่อ p เท่ากับ 0.2

n	r	**ระดับความเชื่อมั่น**		
		90%	**95%**	**99%**
30	10	Bayes2	Bayes2	Bayes2
	15	Bayes2	Bayes2	Bayes2
	20	Bayes2	Bayes2	Bayes2
50	10	Bayes1	Bayes2	Bayes2
	20	Bayes2	Bayes2	Bayes2
	30	Bayes2	Bayes1 Bayes2 MCMC	Bayes2
70	15	Bayes2	Bayes2	Bayes2
	30	Bayes2	Bayes2	Bayes2
	45	Bayes1 Bayes2 Bayes3 MCMC	Bayes1 Bayes2	Bayes1 Bayes2 Bayes3 MCMC

- วิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (3,3) ให้ค่าเฉลี่ยความกว้างของช่วงความเชื่อมั่นที่แคบที่สุด เมื่อขนาดตัวอย่างเท่ากับ 50 ที่ค่าพารามิเตอร์ r เท่ากับ 10 และขนาดตัวอย่างเท่ากับ 70 ที่ค่าพารามิเตอร์ r เท่ากับ 45 ที่ระดับความเชื่อมั่น 90% เมื่อขนาดตัวอย่างเท่ากับ 50 ที่ค่าพารามิเตอร์ r เท่ากับ 30 และขนาดตัวอย่างเท่ากับ 70 ที่ค่าพารามิเตอร์ r เท่ากับ 45 ที่ระดับความเชื่อมั่น 95% และเมื่อขนาดตัวอย่างเท่ากับ 70 ที่ค่าพารามิเตอร์ r เท่ากับ 45 ที่ระดับความเชื่อมั่น 99%
- วิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (2,6) ส่วนใหญ่ให้ค่าเฉลี่ยความกว้างของช่วงความเชื่อมั่นที่แคบที่สุด ในทุกระดับความเชื่อมั่น ยกเว้น เมื่อขนาดตัวอย่างเท่ากับ 50 ที่ค่าพารามิเตอร์ r เท่ากับ 10 ที่ระดับความเชื่อมั่น 90%

- วิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (6,2) ให้ค่าเฉลี่ยความกว้างของช่วงความเชื่อมั่นที่แคบที่สุด เมื่อขนาดตัวอย่างเท่ากับ 70 ที่ค่าพารามิเตอร์ r เท่ากับ 45 ที่ระดับความเชื่อมั่น 90% และ 99%

- วิธีมาร์คอฟ เชน มอนติคาร์โล ให้ค่าเฉลี่ยความกว้างของช่วงความเชื่อมั่นที่แคบที่สุด เมื่อขนาดตัวอย่างเท่ากับ 70 ที่ค่าพารามิเตอร์ r เท่ากับ 45 ที่ระดับความเชื่อมั่น 90% และ 99% และเมื่อขนาดตัวอย่างเท่ากับ 50 ที่ค่าพารามิเตอร์ r เท่ากับ 30 ที่ระดับความเชื่อมั่น 95%

ตารางที่ 5.5 วิธีประมาณค่าพารามิเตอร์แบบช่วงที่ให้ค่าเฉลี่ยความกว้างของช่วงความเชื่อมั่นแคบสุด จำแนกตามขนาดตัวอย่าง (n) ค่าพารามิเตอร์ r และระดับความเชื่อมั่น เมื่อ p เท่ากับ 0.5

n	r	ระดับความเชื่อมั่น		
		90%	**95%**	**99%**
30	10	Bayes2 Bayes3	Bayes3	Bayes2 Bayes3
	15	Bayes3	Bayes2	Bayes2
	20	Bayes2 Bayes3	Bayes2 Bayes3	Bayes2 Bayes3
50	10	Bayes2 Bayes3	Bayes2 Bayes3	Bayes2 Bayes3
	20	Bayes2 Bayes3	Bayes2 Bayes3	Bayes2 Bayes3
	30	Bayes2 Bayes3	Bayes2 Bayes3	Bayes2 Bayes3
70	15	Bayes2 Bayes3	Bayes2 Bayes3	Bayes2 Bayes3
	30	Bayes2 Bayes3 MCMC	Bayes2 Bayes3	Bayes1 Bayes2 Bayes3
	45	Bayes1 Bayes2 Bayes3	Bayes2 Bayes3	Bayes1 Bayes2 Bayes3

- วิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (3,3) ให้ค่าเฉลี่ยความกว้างของช่วงความเชื่อมั่นที่แคบที่สุด เมื่อขนาดตัวอย่างเท่ากับ 70 ที่ค่าพารามิเตอร์ r เท่ากับ 45 ที่ระดับความเชื่อมั่น 90% และ 99% และขนาดตัวอย่างเท่ากับ 70 ที่ค่าพารามิเตอร์ r เท่ากับ 30 ที่ระดับความเชื่อมั่น 99%

- วิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (2,6) ส่วนใหญ่ค่าเฉลี่ยความกว้างของช่วงความเชื่อมั่นที่แคบที่สุด ในทุกระดับความเชื่อมั่น ยกเว้น เมื่อขนาดตัวอย่างเท่ากับ 30 ที่ค่าพารามิเตอร์ r เท่ากับ 15 ที่ระดับความเชื่อมั่น 90% และเมื่อขนาดตัวอย่างเท่ากับ 30 ที่ค่าพารามิเตอร์ r เท่ากับ 10 ที่ระดับความเชื่อมั่น 95%

- วิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (6,2) ส่วนใหญ่ค่าเฉลี่ยความกว้างของช่วงความเชื่อมั่นที่แคบที่สุด ในทุกระดับความเชื่อมั่น ยกเว้น เมื่อขนาดตัวอย่างเท่ากับ 30 ที่ค่าพารามิเตอร์ r เท่ากับ 15 ที่ระดับความเชื่อมั่น 95% และ 99%

- วิธีมาร์คอฟ เชน มอนติคาร์โล ให้ค่าเฉลี่ยความกว้างของช่วงความเชื่อมั่นที่แคบที่สุด เมื่อขนาดตัวอย่างเท่ากับ 70 ที่ค่าพารามิเตอร์ r เท่ากับ 30 ที่ระดับความเชื่อมั่น 90%

ตารางที่ 5.6 วิธีประมาณค่าพารามิเตอร์แบบช่วงที่ให้ค่าเฉลี่ยความกว้างของช่วงความเชื่อมั่นแคบสุด จำแนกตามขนาดตัวอย่าง (n) ค่าพารามิเตอร์ r และระดับความเชื่อมั่น เมื่อ p เท่ากับ 0.8

n	r	ระดับความเชื่อมั่น		
		90%	**95%**	**99%**
30	10	Bayes3	Bayes3	Bayes1
	15	Bayes3	Bayes3	Bayes3
	20	Bayes3	Bayes3	Bayes3
50	10	Bayes3	Bayes3	Bayes3
	20	Bayes3	Bayes3	Bayes3
	30	Bayes3	Bayes3	Bayes3
70	15	Bayes3	Bayes3	Bayes3
	30	Bayes3	Bayes3	Bayes3
	45	Bayes3	Bayes3	Bayes3

- วิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (3,3) ให้ค่าเฉลี่ยความกว้างของช่วงความเชื่อมั่นที่แคบที่สุด เมื่อขนาดตัวอย่างเท่ากับ 30 ที่ค่าพารามิเตอร์ r เท่ากับ 10 ที่ระดับความเชื่อมั่น 99%

- วิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงบีตา (6,2) ส่วนใหญ่ให้ค่าเฉลี่ยความกว้างของช่วงความเชื่อมั่นที่แคบที่สุด ในทุกระดับความเชื่อมั่น ยกเว้น เมื่อขนาดตัวอย่างเท่ากับ 30 ที่ค่าพารามิเตอร์ r เท่ากับ 10 ที่ระดับความเชื่อมั่น 99%

5.2 ข้อเสนอแนะ

5.2.1 ด้านการนำไปใช้ประโยชน์

- เมื่อต้องการทำการประมาณค่าแบบจุด ขนาดตัวอย่างเล็กอาจใช้วิธีภาวะน่าจะเป็นสูงสุดเพื่อความสะดวกในการคำนวณ หากตัวอย่างมีขนาดปานกลางหรือใหญ่ควรใช้วิธีของเบส์ แต่หากไม่ทราบค่าพารามิเตอร์ของการแจกแจงก่อน อาจเลือกใช้วิธีมาร์คอฟ เชน มอนติคาร์โลแทนได้
- เมื่อต้องการทำการประมาณค่าแบบช่วง วิธีของเบส์ส่วนใหญ่จะให้ความกว้างเฉลี่ยของช่วงความเชื่อมั่นแคบที่สุด ซึ่งสอดคล้องกับ Araveeporn (2015) และ Araveeporn (2016)
- ขนาดตัวอย่างมีผลต่อการประมาณค่าแบบช่วง โดยความกว้างเฉลี่ยของช่วงความเชื่อมั่นที่ได้จากการประมาณจะมีแนวโน้มลดลง เมื่อขนาดตัวอย่างมีค่ามากขึ้น จึงควรใช้ขนาดตัวอย่างใหญ่

5.2.2 ด้านการศึกษาวิจัย

เพื่อเป็นแนวทางให้ผู้สนใจศึกษาได้ศึกษาเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ของการแจกแจงทวินามเชิงลบ โดยใช้วิธีการประมาณอื่นๆ เช่น วิธีบูทแสตรป วิธีสกอล วิธีวิลค์ เป็นต้น

บรรณานุกรม

นฤพล วงศ์เจริญสันติ และคณะ. 2556. "การเปรียบเทียบการประมาณค่าพารามิเตอร์ด้วยวิธีความควรจะเป็นสูงสุดและวิธีของเบส์ของการแจกแจงปัวซง". ปัญหาพิเศษระดับปริญญาตรี สาขาวิชาสถิติประยุกต์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง.

น้อมจิต กิตติโชติพาณิชย์. 2553. **เอกสารประกอบการสอนวิชาสถิติคณิตศาสตร์ 1**. คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง.

ประชุม สุวัตถี. 2553. **ทฤษฎีการอนุมานเชิงสถิติ**. พิมพ์ครั้งที่ 3. กรุงเทพฯ: สำนักงานกิจการโรงพิมพ์ องค์การสงเคราะห์ทหารผ่านศึก.

ปรางค์มาส ทรัพย์กระจ่าง และยงยุทธ ไชยพงศ์. 2555. "การประมาณค่าแบบช่วงสำหรับค่าสัดส่วนทวินามลบโดยวิธีบูทแสตรป". หน้า 85-89. ใน **การประชุมวิชาการระดับชาติ "วิทยาศาสตร์วิจัย" ครั้งที่ 4**. พิษณุโลก: คณะวิทยาศาสตร์ มหาวิทยาลัยนเรศวร.

ฝันจิต แต้มทอง. 2537. "การอนุมานค่าเฉลี่ยประชากรโดยวิธีเบย์เซียน". วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต สาขาวิชาสถิติประยุกต์ บัณฑิตวิทยาลัย, มหาวิทยาลัยเชียงใหม่.

สายชล สินสมบูรณ์ทอง. 2548. **ความน่าจะเป็น**. พิมพ์ครั้งที่ 3. กรุงเทพฯ: จามจุรี โปรดักส์.

สิรีธร น้อยเจริญ. 2549. "การเปรียบเทียบวิธีการประมาณค่าแบบช่วงสำหรับพารามิเตอร์ของการแจกแจงทวินามลบ". วิทยานิพนธ์สถิติศาสตรมหาบัณฑิต สาขาวิชาสถิติ ภาควิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย .

อนุชา กล่ำน้อย. 2548. "คุณลักษณะบางประการของตัวประมาณแบบเบส์ที่มีคุณสมบัติสอดคล้องกับ MLE และ MVUE". วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต สาขาวิชาสถิติประยุกต์ มหาวิทยาลัยนเศวร.

อัชฌา อระวีพร. 2555. การวิเคราะห์เบส์จากโปรแกรมวินบั๊กสู่โปรแกรมอาร์. *วารสารมหาวิทยาลัยนเรศวร: วิทยาศาสตร์และเทคโนโลยี*. 9(1), 30-44.

อาทิตย์ เทศขำ. 2559. "การเปรียบเทียบการประมาณค่าพารามิเตอร์ของการแจกแจงเลขชี้กำลังด้วยวิธีภาวะน่าจะเป็นสูงสุด วิธีของเบส์ และวิธีมาร์คอฟ เชน มอนติคาร์โล". วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต สาขาวิชาสถิติประยุกต์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง.

Araveeporn, A. 2014. "Parameter Estimation of Poisson Distribution by Using Maximum Likelihood, Markov Chain Monte Carlo, and Bayes methods". *Thammasat International Journal of Science and Technology*. 19(3): 1-14.

Araveeporn, A. 2015. "The Interval Estimation of Poisson Distribution by Using Maximum Likelihood, Markov Chain Monte Carlo, and Bayes methods". 36-42. in **International Conference on Applied Statistics 2015**. Pattaya, Thailand.

Araveeporn, A. 2016. "The Interval Estimation for the mean of Normal Distribution by Markov Chain Monte Carlo method". 1-6. in **International Conference on Applied Statistics 2016**. Phuket, Thailand.

Ganji, M., Eghbali, N. and Azimian, M. 2013. "Bayes and Empirical Bayes Estimation of Parameter k in Negative Binomial Distribution". *Journal of Hyperstructures*. 2(2): 185-200.

Gelfand, A.E. Hills, S.E. Racine-Poon, A. and Smith, A.F.M. 1990. "Illustration of Bayesian Inference in Normal Data Models Using Gibbs Sampling". *Journal American Statistical Association*. 85: 398-409.

Geman, S. and Geman, D. 1984. **Stochastic Relaxation, Gibbs Distribution and the Bayesian Restoration of Images**. IEEE Transactions on Pattern analysis and Machine Intelligence, 6(6): 721-741.

Krishnamoorthy, K. 2016. **Handbook of Statistical Distributions with Applications**. London: Chapman & Hall.

Kumar, J. and Tomer, S.K. 2016. "Reliability Estimation for Negative Binomial Distribution Under Type-II Censoring Scheme". Journal of Statistics Applications & Probability. 5(1): 89-98.

Mir, K.A. 2008. "Sized-Biased Generalized Negative Binomial Distrbution". *Journal of Modern Applied Statistics Methods*. 7(2): 446-453.

Thetkham, A. and Araveeporn, A. 2016. "A Comparison of Parameter Estimation of Exponential Distribution Using Maximum Likelihood, Bayes, and Markov Chain Monte Carlo methods". 55-61. in **International Conference on Applied Statistics 2015**. Phuket, Thailand.

ภาคผนวก

ภาคผนวก ก
ผลการวิจัย

ตารางที่ 1 ค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการประมาณ ที่ระดับความเชื่อมั่น 90% 95% และ 99% ตามลำดับ เมื่อ p เท่ากับ 0.2

n	r	วิธีประมาณ	ระดับความเชื่อมั่น		
			90%	95%	99%
30	10	ML	**0.563**	**0.698**	**0.812**
		Bayes1	0.885	0.949	0.993
		Bayes2	0.889	0.946	0.994
		Bayes3	**0.878**	0.947	0.989
		MCMC	0.885	0.945	0.994
	15	ML	**0.553**	**0.653**	**0.779**
		Bayes1	0.895	0.944	0.989
		Bayes2	0.895	0.949	0.988
		Bayes3	0.887	**0.933**	0.987
		MCMC	0.896	0.949	0.987
	20	ML	**0.607**	**0.652**	**0.820**
		Bayes1	0.899	0.958	0.996
		Bayes2	0.900	0.960	0.995
		Bayes3	0.889	0.948	0.995
		MCMC	0.898	0.961	0.995
50	10	ML	**0.608**	**0.688**	**0.806**
		Bayes1	0.885	0.965	0.991
		Bayes2	**0.880**	0.962	0.991
		Bayes3	0.882	0.961	0.991
		MCMC	**0.881**	0.962	0.991
	20	ML	**0.561**	**0.673**	**0.802**
		Bayes1	0.886	0.961	0.993
		Bayes2	0.888	0.960	0.991
		Bayes3	0.884	0.961	0.988
		MCMC	0.890	0.959	0.991
	30	ML	**0.600**	**0.651**	**0.815**
		Bayes1	0.906	0.944	0.988
		Bayes2	0.905	0.943	0.988

ตารางที่ 1 (ต่อ) ค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการประมาณ ที่ระดับความเชื่อมั่น 90% 95% และ 99% ตามลำดับ เมื่อ p เท่ากับ 0.2

n	r	วิธีประมาณ	ระดับความเชื่อมั่น		
			90%	**95%**	**99%**
50	30	Bayes3	0.904	0.940	0.988
		MCMC	0.905	0.944	0.987
70	15	ML	**0.570**	**0.677**	**0.826**
		Bayes1	**0.876**	0.951	0.991
		Bayes2	0.884	0.948	0.991
		Bayes3	**0.877**	0.944	0.987
		MCMC	0.883	0.948	0.991
	30	ML	**0.601**	**0.686**	**0.816**
		Bayes1	0.896	0.962	0.988
		Bayes2	0.894	0.961	0.987
		Bayes3	0.898	0.962	0.989
		MCMC	0.895	0.961	0.988
	45	ML	**0.584**	**0.666**	**0.807**
		Bayes1	0.905	0.950	0.987
		Bayes2	0.905	0.947	0.987
		Bayes3	0.904	0.945	0.987
		MCMC	0.906	0.950	0.987

หมายเหตุ: ตัวหนา หมายถึง ค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการประมาณต่ำกว่าเกณฑ์ที่กำหนด

ตารางที่ 2 ค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการประมาณ ที่ระดับความเชื่อมั่น 90% 95% และ 99% ตามลำดับ เมื่อ p เท่ากับ 0.5

n	r	**วิธีประมาณ**	**ระดับความเชื่อมั่น**		
			90%	**95%**	**99%**
30	10	ML	0.913	0.956	0.990
		Bayes1	0.915	0.962	0.993
		Bayes2	0.914	0.952	0.990
		Bayes3	0.897	0.958	0.993
		MCMC	0.912	0.959	0.993
	15	ML	0.909	**0.687**	0.985
		Bayes1	0.910	0.954	0.984
		Bayes2	0.909	0.951	0.984
		Bayes3	0.897	0.954	**0.982**
		MCMC	0.910	0.951	0.985
	20	ML	0.886	0.953	0.991
		Bayes1	0.891	0.952	0.989
		Bayes2	0.887	0.954	0.991
		Bayes3	0.888	0.947	0.988
		MCMC	0.889	0.954	0.989
50	10	ML	0.896	0.956	0.989
		Bayes1	0.898	0.960	0.990
		Bayes2	0.902	0.955	0.989
		Bayes3	0.903	0.952	0.990
		MCMC	0.897	0.954	0.989
	20	ML	0.917	0.949	0.987
		Bayes1	0.920	0.949	0.987
		Bayes2	0.922	0.949	0.987
		Bayes3	0.916	0.948	0.989
		MCMC	0.919	0.950	0.987
	30	ML	0.907	0.946	0.991
		Bayes1	0.907	0.947	0.990
		Bayes2	0.905	0.945	0.991

ตารางที่ 2 (ต่อ) ค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการประมาณ ที่ระดับความเชื่อมั่น 90% 95% และ 99% ตามลำดับ เมื่อ p เท่ากับ 0.5

n	r	วิธีประมาณ	ระดับความเชื่อมั่น		
			90%	**95%**	**99%**
50	30	Bayes3	0.910	0.940	0.991
		MCMC	0.905	0.944	0.990
70	15	ML	0.898	0.939	0.992
		Bayes1	0.903	0.939	0.991
		Bayes2	0.898	0.939	0.992
		Bayes3	0.900	0.940	0.993
		MCMC	0.904	0.939	0.991
	30	ML	0.895	0.949	0.990
		Bayes1	0.895	0.948	0.988
		Bayes2	0.894	0.949	0.989
		Bayes3	0.895	0.950	0.988
		MCMC	0.896	0.949	0.988
	45	ML	0.892	0.947	0.989
		Bayes1	0.892	0.946	0.990
		Bayes2	0.890	0.945	0.989
		Bayes3	0.895	0.942	0.989
		MCMC	0.893	0.945	0.989

หมายเหตุ: ตัวหนา หมายถึง ค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการประมาณต่ำกว่าเกณฑ์ที่กำหนด

ตารางที่ 3 ค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการประมาณ ที่ระดับความเชื่อมั่น 90% 95% และ 99% ตามลำดับ เมื่อ p เท่ากับ 0.8

n	r	**วิธีประมาณ**	**ระดับความเชื่อมั่น**		
			90%	**95%**	**99%**
30	10	ML	0.997	0.999	1.000
		Bayes1	0.908	0.962	0.990
		Bayes2	**0.858**	**0.936**	0.990
		Bayes3	0.907	0.958	**0.983**
		MCMC	0.908	0.958	**0.983**
	15	ML	1.000	1.000	1.000
		Bayes1	0.894	0.959	0.984
		Bayes2	0.886	0.951	**0.977**
		Bayes3	0.894	0.954	0.988
		MCMC	0.895	0.951	0.985
	20	ML	0.998	1.000	1.000
		Bayes1	0.895	0.950	0.991
		Bayes2	0.882	0.943	**0.983**
		Bayes3	0.895	0.950	0.990
		MCMC	0.894	0.949	0.990
50	10	ML	0.999	1.000	1.000
		Bayes1	0.889	0.946	0.992
		Bayes2	**0.880**	0.940	0.992
		Bayes3	0.898	0.952	0.988
		MCMC	0.904	0.944	0.989
	20	ML	1.000	1.000	1.000
		Bayes1	0.904	0.958	0.991
		Bayes2	0.889	0.951	0.988
		Bayes3	0.904	0.958	0.989
		MCMC	0.899	0.955	0.990
	30	ML	0.999	1.000	1.000
		Bayes1	0.902	0.948	0.994
		Bayes2	0.895	0.938	0.996

ตารางที่ 3 (ต่อ) ค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการประมาณ ที่ระดับความเชื่อมั่น 90% 95% และ 99% ตามลำดับ เมื่อ p เท่ากับ 0.8

n	r	วิธีประมาณ	ระดับความเชื่อมั่น		
			90%	**95%**	**99%**
50	30	Bayes3	0.903	0.947	0.994
		MCMC	0.907	0.948	0.994
70	15	ML	0.995	1.000	1.000
		Bayes1	0.903	0.944	0.996
		Bayes2	0.893	0.941	0.994
		Bayes3	0.902	0.951	0.995
		MCMC	0.897	0.950	0.995
	30	ML	0.999	1.000	1.000
		Bayes1	0.886	0.960	0.991
		Bayes2	**0.879**	0.955	0.992
		Bayes3	0.891	0.959	0.989
		MCMC	0.887	0.961	0.990
	45	ML	0.999	1.000	1.000
		Bayes1	0.917	0.963	0.994
		Bayes2	0.912	0.964	0.993
		Bayes3	0.920	0.966	0.993
		MCMC	0.919	0.965	0.993

หมายเหตุ: ตัวหนา หมายถึง ค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการประมาณต่ำกว่าเกณฑ์ที่กำหนด

ตารางที่ 4 ค่าความกว้างเฉลี่ยที่ได้จากการประมาณ ที่ระดับความเชื่อมั่น 90% 95% และ 99% ตามลำดับ เมื่อ p เท่ากับ 0.2

n	r	วิธีประมาณ	ระดับความเชื่อมั่น		
			90%	**95%**	**99%**
30	10	ML	-	-	-
		Bayes1	0.03405	0.04045	0.05343
		Bayes2	**0.03397**	**0.04035**	**0.05331**
		Bayes3	-	0.04055	0.05357
		MCMC	0.03404	0.04045	0.05341
	15	ML	-	-	-
		Bayes1	0.02773	0.03320	0.04348
		Bayes2	**0.02769**	**0.03315**	**0.04341**
		Bayes3	0.02778	-	0.04355
		MCMC	0.02774	0.03321	0.04346
	20	ML	-	-	-
		Bayes1	0.02401	0.02866	0.03758
		Bayes2	**0.02398**	**0.02862**	**0.03753**
		Bayes3	0.02404	0.02869	0.03763
		MCMC	0.02401	0.02866	0.03757
50	10	ML	-	-	-
		Bayes1	**0.02637**	0.03142	0.04118
		Bayes2	-	**0.03137**	**0.04112**
		Bayes3	0.02641	0.03147	0.04124
		MCMC	-	0.03141	0.04118
	20	ML	-	-	-
		Bayes1	0.01863	0.02219	0.02915
		Bayes2	**0.01861**	**0.02217**	**0.02913**
		Bayes3	0.01864	0.02221	0.02917
		MCMC	0.01863	0.02218	0.02914
	30	ML	-	-	-
		Bayes1	0.01519	**0.01809**	0.02379
		Bayes2	**0.01518**	**0.01809**	**0.02378**

ตารางที่ 4 (ต่อ) ค่าความกว้างเฉลี่ยที่ได้จากการประมาณ ที่ระดับความเชื่อมั่น 90% 95% และ 99% ตามลำดับ เมื่อ p เท่ากับ 0.2

n	r	วิธีประมาณ	ระดับความเชื่อมั่น		
			90%	**95%**	**99%**
50	30	Bayes3	0.01520	0.01810	0.02380
		MCMC	0.01519	**0.01809**	0.02379
70	15	ML	-	-	-
		Bayes1	-	0.02168	0.02850
		Bayes2	**0.01816**	**0.02166**	**0.02848**
		Bayes3	-	0.02169	0.02852
		MCMC	0.01817	0.02167	0.02851
	30	ML	-	-	-
		Bayes1	0.01285	0.01531	0.02011
		Bayes2	**0.01284**	**0.01530**	**0.02010**
		Bayes3	0.01285	0.01531	0.02012
		MCMC	0.01285	0.01531	0.02011
	45	ML	-	-	-
		Bayes1	**0.01049**	**0.01249**	**0.01643**
		Bayes2	**0.01049**	**0.01249**	**0.01643**
		Bayes3	**0.01049**	0.01250	**0.01643**
		MCMC	**0.01049**	0.01250	**0.01643**

หมายเหตุ: ตัวหนา หมายถึง ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นแคบสุด ในแต่ละสถานการณ์และระดับความเชื่อมั่นที่กำหนด

"-" หมายถึง ไม่สามารถคำนวณค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นได้ เนื่องจากให้ค่าสัมประสิทธิ์ความเชื่อมั่นต่ำกว่าเกณฑ์ที่กำหนด

ตารางที่ 5 ค่าความกว้างเฉลี่ยที่ได้จากการประมาณ ที่ระดับความเชื่อมั่น 90% 95% และ 99% ตามลำดับ เมื่อ p เท่ากับ 0.5

n	r	**วิธีประมาณ**	**ระดับความเชื่อมั่น**		
			90%	**95%**	**99%**
30	10	ML	0.06729	0.08007	0.10527
		Bayes1	0.06678	0.07952	0.10453
		Bayes2	**0.06667**	0.07939	**0.10435**
		Bayes3	**0.06667**	**0.07938**	**0.10435**
		MCMC	0.06698	0.07976	0.10486
	15	ML	0.05505	-	0.08596
		Bayes1	0.05469	0.03306	0.08552
		Bayes2	0.05464	**0.03300**	**0.08543**
		Bayes3	**0.05463**	0.03311	-
		MCMC	0.05478	0.03306	0.08574
	20	ML	0.04749	0.05656	0.07454
		Bayes1	0.04733	0.05639	0.07420
		Bayes2	**0.04729**	**0.05634**	**0.07414**
		Bayes3	**0.04729**	**0.05634**	**0.07414**
		MCMC	0.04740	0.05647	0.07432
50	10	ML	0.05194	0.06194	0.08144
		Bayes1	0.05177	0.06172	0.08112
		Bayes2	**0.05172**	**0.06165**	**0.08104**
		Bayes3	**0.05172**	**0.06165**	**0.08104**
		MCMC	0.05189	0.06184	0.08130
	20	ML	0.03676	0.04379	0.05766
		Bayes1	0.03670	0.04372	0.05751
		Bayes2	**0.03668**	**0.04370**	**0.05748**
		Bayes3	**0.03668**	**0.04370**	**0.05748**
		MCMC	0.03672	0.04377	0.05755
	30	ML	0.03005	0.03577	0.04704
		Bayes1	0.03000	0.03573	0.04697
		Bayes2	**0.02999**	**0.03572**	**0.04695**

ตารางที่ 5 (ต่อ) ค่าความกว้างเฉลี่ยที่ได้จากการประมาณ ที่ระดับความเชื่อมั่น 90% 95% และ 99% ตามลำดับ เมื่อ p เท่ากับ 0.5

n	r	วิธีประมาณ	ระดับความเชื่อมั่น		
			90%	**95%**	**99%**
50	30	Bayes3	**0.02999**	**0.03572**	**0.04695**
		MCMC	0.03001	0.03576	0.04699
70	15	ML	0.03586	0.04279	0.05616
		Bayes1	0.03581	0.04270	0.05608
		Bayes2	**0.03579**	**0.04268**	**0.05605**
		Bayes3	**0.03579**	**0.04268**	**0.05605**
		MCMC	0.03583	0.04275	0.05613
	30	ML	0.02540	0.03024	0.03976
		Bayes1	0.02537	0.03021	**0.03971**
		Bayes2	**0.02536**	**0.03020**	**0.03971**
		Bayes3	**0.02536**	**0.03020**	**0.03971**
		MCMC	**0.02536**	0.03022	0.03974
	45	ML	0.02073	0.02469	0.03247
		Bayes1	**0.02071**	0.02468	**0.03244**
		Bayes2	**0.02071**	**0.02467**	**0.03244**
		Bayes3	**0.02071**	**0.02467**	**0.03244**
		MCMC	0.02072	0.02469	0.03245

หมายเหตุ: ตัวหนา หมายถึง ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นแคบสุด ในแต่ละสถานการณ์และระดับความเชื่อมั่นที่กำหนด

"-" หมายถึง ไม่สามารถคำนวณค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นได้ เนื่องจากให้ค่าสัมประสิทธิ์ความเชื่อมั่นต่ำกว่าเกณฑ์ที่กำหนด

ตารางที่ 6 ค่าความกว้างเฉลี่ยที่ได้จากการประมาณ ที่ระดับความเชื่อมั่น 90% 95% และ 99% ตามลำดับ เมื่อ p เท่ากับ 0.8

n	r	**วิธีประมาณ**	**ระดับความเชื่อมั่น**		
			90%	**95%**	**99%**
30	10	ML	0.13608	0.16204	0.21359
		Bayes1	0.06773	0.08073	**0.10598**
		Bayes2	-	-	0.10700
		Bayes3	**0.06709**	**0.07997**	-
		MCMC	0.06766	0.08065	-
	15	ML	0.11131	0.13263	0.17381
		Bayes1	0.05531	0.06591	0.08671
		Bayes2	0.05568	0.06635	-
		Bayes3	**0.05496**	**0.06550**	**0.08617**
		MCMC	0.05528	0.06588	0.08664
	20	ML	0.09601	0.11484	0.15072
		Bayes1	0.04800	0.05711	0.07510
		Bayes2	0.04824	0.05740	-
		Bayes3	**0.04777**	**0.05684**	**0.07474**
		MCMC	0.04796	0.05707	0.07509
50	10	ML	0.10551	0.12544	0.16524
		Bayes1	0.05250	0.06261	0.08222
		Bayes2	-	0.06298	0.08271
		Bayes3	**0.05220**	**0.06226**	**0.08175**
		MCMC	0.05248	0.06261	0.08217
	20	ML	0.07441	0.08872	0.11676
		Bayes1	0.03719	0.04431	0.05820
		Bayes2	0.03731	0.04444	0.05838
		Bayes3	**0.03709**	**0.04418**	**0.05803**
		MCMC	0.03719	0.04429	0.05818
	30	ML	0.06073	0.07244	0.09533
		Bayes1	0.03038	0.03619	0.04753
		Bayes2	0.03044	0.03626	0.04763

ตารางที่ 6 (ต่อ) ค่าความกว้างเฉลี่ยที่ได้จากการประมาณ ที่ระดับความเชื่อมั่น 90% 95% และ 99% ตามลำดับ เมื่อ p เท่ากับ 0.8

n	r	วิธีประมาณ	ระดับความเชื่อมั่น		
			90%	95%	99%
50	30	Bayes3	**0.03032**	**0.03612**	**0.04744**
		MCMC	0.03038	0.03619	0.04752
70	15	ML	0.07270	0.08664	0.11379
		Bayes1	0.03628	0.04323	0.05683
		Bayes2	0.03639	0.04335	0.05699
		Bayes3	**0.03618**	**0.04311**	**0.05667**
		MCMC	0.03626	0.04322	0.05683
	30	ML	0.05141	0.06125	0.08039
		Bayes1	0.02566	0.03058	0.04021
		Bayes2	-	0.03063	0.04027
		Bayes3	**0.02563**	**0.03054**	**0.04016**
		MCMC	0.02566	0.03058	0.04020
	45	ML	0.04198	0.04995	0.06564
		Bayes1	0.02096	0.02499	0.03284
		Bayes2	0.02098	0.02501	0.03287
		Bayes3	**0.02094**	**0.02496**	**0.03281**
		MCMC	0.02096	0.02498	0.03283

หมายเหตุ: ตัวหนา หมายถึง ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นแคบสุด ในแต่ละสถานการณ์และระดับความเชื่อมั่นที่กำหนด

"-" หมายถึง ไม่สามารถคำนวณค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นได้ เนื่องจากให้ค่าสัมประสิทธิ์ความเชื่อมั่นต่ำกว่าเกณฑ์ที่กำหนด

ภาคผนวก ข
คำสั่งโปรแกรมอาร์ที่ใช้ในงานวิจัย

คำสั่งโปรแกรมอาร์ที่ใช้ในการประมาณค่าแบบจุด

```
m=1000
n=10
r=3
p=0.2
a1=3 ; b1=3
a2=2 ; b2=6
a3=6 ; b3=2
set.seed(5)
####################################################################
mle=c()
baye1=c()
baye2=c()
baye3=c()
a=c()
b=c()
baye.mcmc=c()
mcmc=c()
####################################################################
for(j in 1:m){
x=rnbinom(n,r,p)
mle[j]=((r*n)-1)/((r*n)+sum(x)-1)
baye1[j]=((r*n)+a1)/((r*n)+a1+sum(x)+b1)
baye2[j]=((r*n)+a2)/((r*n)+a2+sum(x)+b2)
baye3[j]=((r*n)+a3)/((r*n)+a3+sum(x)+b3)
####################################################################
library(rjags)
dataset=list(n=n,x=x,r=r)
inits=list(p=0.2,a=1,b=1)
####################################################################
jagmod<-jags.model('aa.txt',data=dataset,inits=inits,n.chains=1,n.adapt=1000)
update(jagmod,n.iter=10000,progress.bar="text")
posterior=coda.samples(jagmod,c("p","a","b"),n.iter=10000,progress.bar="text",thin=1)
####################################################################
post=as.data.frame(as.matrix(posterior))
mcmc[j]=mean(post$p)
a[j] = mean(post$a)
```

```
b[j] = mean(post$b)
baye.mcmc[j]=((r*n)+a[j])/((r*n)+a[j]+sum(x)+b[j])
######################################################################
cat(c("loop:",j),fill=T)
}
######################################################################
t.test(mle, mu=p)
sd1=sd(mle)
sd1
mu1 = mean(mle)
######################################################################
t.test(baye1, mu=p)
sd2=sd(baye1)
sd2
mu2 = mean(baye1)
######################################################################
t.test(baye2, mu=p)
sd3=sd(baye2)
sd3
mu3 = mean(baye2)
######################################################################
t.test(baye3, mu=p)
sd4=sd(baye3)
sd4
mu4 = mean(baye3)
######################################################################
t.test(mcmc, mu=p)
sd5=sd(mcmc)
sd5
mu5 = mean(mcmc)
######################################################################
t.test(baye.mcmc, mu=p)
sd6=sd(baye.mcmc)
sd6
mu6 = mean(baye.mcmc)
######################################################################
cat('\tn = ' , n, '\tr = ', r, '\tM = ', m,'\tp = ', p,'\n',
```

```
'\tmle.mean = ' , mu1, '\tmle.sd = ', sd1, '\tStatistic = ' , t.test(mle, mu=p)$statistic,
'\tPvalue = ', t.test(mle, mu=p)$p.value, '\n',
'\tbaye1.mean = ' , mu2, '\tbaye1.sd = ', sd2, '\tStatistic = ' , t.test(baye1,
mu=p)$statistic, '\tPvalue = ', t.test(baye1, mu=p)$p.value, '\n',
'\tbaye2.mean = ' , mu3, '\tbaye2.sd = ', sd3, '\tStatistic = ' , t.test(baye2,
mu=p)$statistic, '\tPvalue = ', t.test(baye2, mu=p)$p.value, '\n',
'\tbaye3.mean = ' , mu4, '\tbaye3.sd = ', sd4, '\tStatistic = ' , t.test(baye3,
mu=p)$statistic, '\tPvalue = ', t.test(baye3, mu=p)$p.value, '\n',
'\tmcmc.mean = ' , mu5, '\tmcmc.sd= ', sd5, '\tStatistic = ' , t.test(mcmc,
mu=p)$statistic, '\tPvalue = ', t.test(mcmc, mu=p)$p.value, '\n',
'\tbaye.mcmc.mean = ' , mu6, '\tbaye.mcmc.sd = ', sd6, '\tStatistic = ' ,
t.test(baye.mcmc, mu=p)$statistic, '\tPvalue = ', t.test(baye.mcmc, mu=p)$p.value,
'\n')
######################################################################
output_MLE=data.frame(mle,baye1, baye2, baye3, mcmc, baye.mcmc)
write.table(output_MLE,file="NB_n10_r3_p(0.2).txt",sep="\t",row.name=FALSE)
```

คำสั่งโปรแกรมอาร์ที่ใช้ในการประมาณค่าแบบช่วง

```
m <- 1000
p <- 0.2
n <- 30
r <- 10
a1 <- 3 ; b1 <- 3
a2 <- 2 ; b2 <- 6
a3 <- 6 ; b3 <- 2
alpha <- 0.1
conf.cv <- 0.8814
#######################################################################
mle = rep(0,m); var.mle = rep(0,m)
bayes1 = rep(0,m) ;var.bayes1 = rep(0,m)
bayes2 = rep(0,m) ;var.bayes2 = rep(0,m)
bayes3 = rep(0,m) ;var.bayes3 = rep(0,m)
mcmc = rep(0,m); var.mcmc = rep(0,m)
lower.mle = rep(0,m); upper.mle = rep(0,m)
lower.bayes1 = rep(0,m); upper.bayes1 = rep(0,m)
lower.bayes2 = rep(0,m); upper.bayes2 = rep(0,m)
lower.bayes3 = rep(0,m); upper.bayes3 = rep(0,m)
lower.mcmc = rep(0,m); upper.mcmc = rep(0,m)
temp.mle = rep(0,m)
temp.bayes1 = rep(0,m)
temp.bayes2 = rep(0,m)
temp.bayes3 = rep(0,m)
temp.mcmc = rep(0,m)
length.mle = rep(0,m)
length.bayes1 = rep(0,m)
length.bayes2 = rep(0,m)
length.bayes3 = rep(0,m)
length.mcmc = rep(0,m)
#######################################################################
 for ( j in 1:m)
{x <- rnbinom(n,r,p)
mle[j] <- ((r*n)-1)/((r*n)+sum(x)-1)
var.mle[j] <- ((((p^2)*(1-p))/2)*((2*sum(x))+2-p)/(sum(x)*(sum(x)-p+2)))
bayes1[j] <- ((r*n)+a1)/((r*n)+a1+sum(x)+b1)
```

```
var.bayes1[j] <-
(((r*n)+a1)*(sum(x)+b1))/((((r*n)+a1+sum(x)+b1)^2)*((r*n)+a1+sum(x)+b1+1))
bayes2[j] <- ((r*n)+a2)/((r*n)+a2+sum(x)+b2)
var.bayes2[j] <-
(((r*n)+a2)*(sum(x)+b2))/((((r*n)+a2+sum(x)+b2)^2)*((r*n)+a2+sum(x)+b2+1))
bayes3[j] <- ((r*n)+a3)/((r*n)+a3+sum(x)+b3)
var.bayes3[j] <-
(((r*n)+a3)*(sum(x)+b3))/((((r*n)+a3+sum(x)+b3)^2)*((r*n)+a3+sum(x)+b3+1))
r1=r
######################################################################
library(rjags)
dataset=list(n=n,x=x,r=r1)
inits=list(p=p,a=1,b=1)
######################################################################
jagmod<-jags.model('aa.txt',data=dataset,inits=inits,n.chains=1,n.adapt=1000)
update(jagmod,n.iter=10000,progress.bar="text")
posterior=coda.samples(jagmod,c("p","a","b"),n.iter=10000,progress.bar="text",thin=1)
######################################################################
post=as.data.frame(as.matrix(posterior))
mcmc[j] = mean(post$p)
a[j] = mean(post$a)
b[j] = mean(post$b)
var.mcmc[j] = var(post$p)
######################################################################
lower.mle[j] <- mle[j]-qnorm(1- alpha/2)*sqrt(var.mle[j])
upper.mle[j] <- mle[j]+qnorm(1- alpha/2)*sqrt(var.mle[j])
lower.bayes1[j] <- bayes1[j]-qnorm(1- alpha/2)*sqrt(var.bayes1[j])
upper.bayes1[j] <- bayes1[j]+qnorm(1- alpha/2)*sqrt(var.bayes1[j])
lower.bayes2[j] <- bayes2[j]-qnorm(1- alpha/2)*sqrt(var.bayes2[j])
upper.bayes2[j] <- bayes2[j]+qnorm(1- alpha/2)*sqrt(var.bayes2[j])
lower.bayes3[j] <- bayes3[j]-qnorm(1- alpha/2)*sqrt(var.bayes3[j])
upper.bayes3[j] <- bayes3[j]+qnorm(1- alpha/2)*sqrt(var.bayes3[j])
lower.mcmc[j] <- mcmc[j]-qnorm(1- alpha/2)*sqrt(var.mcmc[j])
upper.mcmc[j] <- mcmc[j]+qnorm(1- alpha/2)*sqrt(var.mcmc[j])
######################################################################
if ((p >=lower.mle[j]) & (p <=upper.mle[j]))  {temp.mle[j] <-1}
if ((p >=lower.bayes1[j]) & (p <=upper.bayes1[j])) {temp.bayes1[j] <-1}
if ((p >=lower.bayes2[j]) & (p <=upper.bayes2[j])) {temp.bayes2[j] <-1}
```

```
if ((p >=lower.bayes3[j]) & (p <=upper.bayes3[j])) {temp.bayes3[j] <-1}
if ((p >=lower.mcmc[j]) & (p <=upper.mcmc[j]))  {temp.mcmc[j] <-1}
####################################################################
length.mle[j] <-upper.mle[j]-lower.mle[j]
length.bayes1[j] <-upper.bayes1[j]-lower.bayes1[j]
length.bayes2[j] <-upper.bayes2[j]-lower.bayes2[j]
length.bayes3[j] <-upper.bayes3[j]-lower.bayes3[j]
length.mcmc[j] <-upper.mcmc[j]-lower.mcmc[j]
}
####################################################################
conf.mle = mean(temp.mle)
conf.bayes1 = mean(temp.bayes1)
conf.bayes2 = mean(temp.bayes2)
conf.bayes3 = mean(temp.bayes3)
conf.mcmc = mean(temp.mcmc)
####################################################################
if(conf.mle>=conf.cv)
{ width.mle <-mean(length.mle)}
if(conf.mle < conf.cv)
{width.mle <-c("-")}
####################################################################
if(conf.bayes1>=conf.cv)
{width.bayes1 <-mean(length.bayes1)}
if(conf.bayes1 < conf.cv)
{width.bayes1 <-c("-")}
####################################################################
if(conf.bayes2>=conf.cv)
{width.bayes2 <-mean(length.bayes2)}
if(conf.bayes2 < conf.cv)
{width.bayes2 <-c("-")}
####################################################################
if(conf.bayes3>=conf.cv)
{width.bayes3 <-mean(length.bayes3)}
if(conf.bayes3 < conf.cv)
{width.bayes3 <-c("-")}
####################################################################
if(conf.mcmc>=conf.cv)
{ width.mcmc <-mean(length.mcmc)}
```

```
if(conf.mcmc < conf.cv)
{ width.mcmc <-c("-")}
####################################################################
cat('\tn = ',n,'\tr = ',r,'\tp = ',p,'\talpha = ',alpha,'\tconfidence coef = ',conf.cv,'\n',
'\tconf mle = ',conf.mle, '\tconf bayes1 = ', conf.bayes1,'\tconf bayes2 = ',
conf.bayes2,
'\tconf bayes3 = ', conf.bayes3, '\tconf mcmc = ', conf.mcmc,'\n',
'\twidth mle = ',width.mle, '\twidth bayes1 = ', width.bayes1,'\twidth bayes2 = ',
width.bayes2,
'\twidth bayes3 = ', width.bayes3, '\twidth mcmc = ', width.mcmc,'\n')
####################################################################
output_interval=data.frame(n,r,p,conf.mle,width.mle,conf.bayes1,width.bayes1,conf.b
ayes2,width.bayes2,conf.bayes3,width.bayes3,conf.mcmc,width.mcmc)
write.table(output_interval,"Interval1.txt",sep="\t",row.name=FALSE)
```

ประวัติผู้เขียน

ชื่อ นามสกุล นางสาวอัญมณี กุมมาระกะ

วัน เดือน ปีเกิด 23 เมษายน 2535

ที่อยู่ปัจจุบัน 120 หมู่ 3 ตำบลสนามแจง อำเภอบ้านหมี่ จังหวัดลพบุรี 15110

ประวัติการศึกษา

ระดับมัธยมศึกษา โรงเรียนบ้านหมี่วิทยา จังหวัดลพบุรี

ระดับอุดมศึกษา วิทยาศาสตรบัณฑิต สาขาสถิติประยุกต์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

วิทยาศาสตรมหาบัณฑิต สาขาสถิติประยุกต์
สถาบันเทคโนโลยีพระจอมกเกล้าเจ้าคุณทหารลาดกระบัง

ประวัติการทำงาน

เจ้าหน้าที่ส่วนงาน Budgeting, Actuarial Department
บริษัท ไทยสมุทรประกันชีวิต จำกัด (มหาชน)

ผลงานทางวิชาการ

อัญมณี กุมมาระกะ และ อัชฌา อระวีพร. **การเปรียบเทียบการประมาณค่าพารามิเตอร์ของการแจกแจงทวินามเชิงลบ ด้วยวิธีภาวะน่าจะเป็นสูงสุด วิธีของเบส์และวิธีมาร์คอฟ เชน มอนติคาร์โล**. วารสารวิทยาศาสตร์และเทคโนโลยี. 26(1).

Kummaraka, U. and Araveeporn A. 2017. A Comparison of Confidence Intervals of Negative Binomial Parameter p by Maximum Likelihood, Bayesian and Markov Chain Monte Carlo Methods. PROCEEDINGS International Conference on Applied Statistics 2017.