



วิทยานิพนธ์

การศึกษาเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์
สำหรับการแจกแจงทวินาม

A COMPARATIVE STUDY OF PARAMETER ESTIMATION
METHODS FOR BINOMIAL DISTRIBUTION

นางสาวมัลลิกา ทานูลิทธิ

บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์

พ.ศ. 2551



ใบรับรองวิทยานิพนธ์
บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์

วิทยาศาสตร์มหาบัณฑิต (สถิติ)

ปริญญา

สถิติ

สาขา

สถิติ

ภาควิชา

เรื่อง การศึกษาเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์สำหรับการแจกแจงทวินาม

A Comparative Study of Parameter Estimation Methods for Binomial Distribution

นามผู้วิจัย นางสาวมัลลิกา ทานุสิทธิ์

ได้พิจารณาเห็นชอบโดย

อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

(รองศาสตราจารย์ประสิทธิ์ พยัคฆพงษ์, วท.ม.)

อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม

(อาจารย์อัมไพ ทองธีรภาพ, Ph.D.)

หัวหน้าภาควิชา

(รองศาสตราจารย์สายสุดา สมจิต, วท.ม.)

บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์รับรองแล้ว

(รองศาสตราจารย์วินัย อาจคงหาญ, M.A.)

คณบดีบัณฑิตวิทยาลัย

วันที่ 29 เดือน พฤษภาคม พ.ศ. 2551

วิทยานิพนธ์

เรื่อง

การศึกษาเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์สำหรับการแจกแจงทวินาม

A Comparative Study of Parameter Estimation Methods for Binomial Distribution

โดย

นางสาวมัลลิกา ทานุสิทธิ์

เสนอ

บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์
เพื่อความสมบูรณ์แห่งปริญญาวิทยาศาสตรมหาบัณฑิต (สถิติ)

พ.ศ. 2551

มัลลิกา ทานูลิทธิ 2551: การศึกษาเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์สำหรับการแจกแจงทวินาม ปริญาวิทยาศาสตร์มหาบัณฑิต (สถิติ) สาขาสถิติ ภาควิชาสถิติ อาจารย์ที่ปรึกษาวิทยานิพนธ์
หลัก: รองศาสตราจารย์ประสิทธิ์ พยัคฆพงษ์, วท.ม. 97 หน้า

การศึกษาวิจัยครั้งนี้ มีวัตถุประสงค์เพื่อศึกษาเปรียบเทียบวิธีการประมาณค่าแบบจุดและแบบช่วง สำหรับการประมาณค่าพารามิเตอร์ของการแจกแจงทวินาม พิจารณาวิธีประมาณค่าแบบจุด 3 วิธี คือ วิธีภาวะน่าจะเป็นสูงสุด วิธีเบส์เซียนและวิธีมินิแมกซ์ โดยใช้ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยต่ำสุดเป็นเกณฑ์ในการเลือกวิธีประมาณ และพิจารณาวิธีประมาณค่าแบบช่วง 3 วิธีคือ วิธีปกติ วิธีโลจิสต์ และวิธีสคอร์แบบปรับความต่อเนื่อง โดยใช้ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นและค่าความกว้างเฉลี่ยต่ำสุดของช่วงความเชื่อมั่นเป็นเกณฑ์ในการเลือกวิธีประมาณ ขอบเขตการศึกษาประกอบด้วย กำหนดขนาดตัวอย่าง คือ ขนาดเล็ก (n เท่ากับ 5 และ 10) ขนาดปานกลาง (n เท่ากับ 30 และ 70) และขนาดใหญ่ (n เท่ากับ 100 และ 200) ค่าพารามิเตอร์ p เท่ากับ 0.05, 0.07, 0.09, 0.10, 0.30 และ 0.50 และระดับความเชื่อมั่นที่ใช้ในการประมาณค่าแบบช่วงคือ 90%, 95% และ 99% ในการวิจัยนี้ดำเนินการโดยการจำลองข้อมูลจากโปรแกรม Microsoft Visual Basic 6 กระทำซ้ำ 5,000 ครั้ง ผลการวิจัยสรุปได้ดังนี้

การประมาณค่าแบบจุด เมื่อตัวอย่างขนาดเล็ก ทุกค่าพารามิเตอร์ p ที่ทำการศึกษา ควรเลือกใช้วิธีเบส์เซียน เนื่องจากให้ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยต่ำสุด เมื่อตัวอย่างขนาดปานกลางและใหญ่ กรณีค่า p เท่ากับ 0.05 ถึง 0.10 พบว่าวิธีเบส์เซียนและวิธีภาวะน่าจะเป็นสูงสุด ให้ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยใกล้เคียงกัน กรณีค่า p เท่ากับ 0.30 ถึง 0.50 พบว่าทั้ง 3 วิธี ให้ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยใกล้เคียงกัน ดังนั้นควรเลือกใช้วิธีภาวะน่าจะเป็นสูงสุด เพราะเป็นวิธีที่ง่ายในการวิเคราะห์และสะดวกกว่า

การประมาณค่าแบบช่วง เมื่อตัวอย่างขนาดเล็ก ควรเลือกใช้วิธีสคอร์แบบปรับความต่อเนื่อง สำหรับค่า p เท่ากับ 0.05 ในขณะที่วิธีโลจิสต์ ควรเลือกใช้สำหรับค่า p เท่ากับ 0.07 ถึง 0.50 เมื่อตัวอย่างขนาดปานกลางและใหญ่ ควรเลือกใช้วิธีโลจิสต์และวิธีปกติลำดับ สำหรับค่า p เท่ากับ 0.05 ถึง 0.10 และวิธีสคอร์แบบปรับความต่อเนื่อง ควรเลือกใช้สำหรับค่า p เท่ากับ 0.30 ถึง 0.50

มัลลิกา ทานูลิทธิ
ลายมือชื่อนิติ์

ประสิทธิ์ พยัคฆพงษ์
ลายมือชื่ออาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

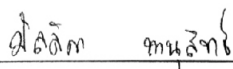
7 / มิ.ค. / 2551

Manlika Tanusit 2008: A Comparative Study of Parameter Estimation Methods for Binomial Distribution. Master of Science (Statistics) , Major Field: Statistics, Department of Statistics.
Thesis Advisor: Associate Professor Prasit Payakkapong, M.S. 97 pages.

The objective of this study was to compare point estimation methods and interval estimation methods for parameter of binomial distribution . Three methods of point estimation were considered : Maximum likelihood method, Bayesian method and Minimax method. The criteria for a selection from these methods were their performance on the lowest mean absolute error and three methods of interval estimation were considered: Normal method , Logit method and Score continuity corrected method. The criteria for a selection from these methods are their performance on the approximate confidence coefficients and the lowest average width of the confidence. The scope of this study consisted of three sample sizes : small (n were 5 and 10), medium (n were 30 and 70), and large (n were 100 and 200) , the parameter p equal to 0.05,0.07,0.09,0.10,0.30 and 0.50 and three confidence levels of 90%,95% and 99%. Data were generated by using the Microsoft visual basic 6 software that replicated 5,000 times. The results of this research are as follow.

For point estimation , we recommended that for all parameter p, Bayesian method should be used for small sample size . For medium and large sample sizes with value of p from 0.05 to 0.10 , the Bayesian and the maximum likelihood methods gave similar mean absolute error and parameter p between 0.30 to 0.50 all three methods gave similar mean absolute error , anyway, the maximum likelihood method should be considered because it is easier and more convenient than the others.

In the case of interval estimation for small sample size, score continuity corrected method is recommended for p equal to 0.05, whereas logit method should be used for value of p between 0.07 to 0.50. The logit and the normal method are suitable for medium and large sample sizes respectively with values of p ranging from 0.05 to 0.10 and score continuity corrected method should be considered with values of p from 0.30 to 0.50.


Student's signature

 7 / 2A / 2551
Thesis Advisor's signature

กิตติกรรมประกาศ

ข้าพเจ้าขอกราบขอบพระคุณ รองศาสตราจารย์ประสิทธิ์ พัทธพงษ์ อาจารย์ที่ปรึกษา
วิทยานิพนธ์หลัก ที่ได้ให้ความช่วยเหลือในการวางแผนการวิจัยในวิทยานิพนธ์ฉบับนี้ ตลอดจนให้
คำปรึกษา แนะนำและตรวจแก้ไขข้อบกพร่องต่างๆ ขอขอบพระคุณ ดร.อำไพ ทองธีรภาพ
อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม รองศาสตราจารย์เปรมใจ ศรีสรานุวัฒนา ประธานกรรมการสอบ
วิทยานิพนธ์ ผู้ช่วยศาสตราจารย์ ดร. วินัย โพธิ์สุวรรณ ผู้ทรงคุณวุฒิภายนอก ที่กรุณาให้
คำปรึกษา แนะนำและช่วยเหลือในการทำวิทยานิพนธ์ให้สำเร็จลุล่วงไปด้วยดี

ผู้วิจัยขอน้อมรำลึกถึงพระคุณของบิดา มารดา และครอบครัวที่ส่งเสริมสนับสนุน
การศึกษาของผู้วิจัยตลอดมา ขอขอบพระคุณครูบาอาจารย์ทุกท่านที่ได้ประสิทธิ์ประสาทวิชา
ความรู้แก่ผู้วิจัย ขอขอบคุณ คุณ เอกชัย อุดมะ ที่ช่วยเหลือและให้คำแนะนำในการเขียน
โปรแกรมในการวิจัยนี้ ขอขอบคุณพี่และเพื่อนๆทุกคนที่เป็นกำลังใจ และช่วยเหลือมาโดยตลอด
ไว้ ณ โอกาสนี้ด้วย

มัลลิกา ทานุสิทธิ์

พฤศจิกายน 2550

สารบัญ

หน้า

สารบัญ	(1)
สารบัญตาราง	(2)
สารบัญภาพ	(3)
คำนำ	1
วัตถุประสงค์	4
การตรวจเอกสาร	6
อุปกรณ์และวิธีการ	40
อุปกรณ์	40
วิธีการ	40
ผลและวิจารณ์	47
สรุปและข้อเสนอแนะ	68
สรุป	68
ข้อเสนอแนะ	71
เอกสารและสิ่งอ้างอิง	73
ภาคผนวก	76
ภาคผนวก ก โปรแกรมที่ใช้จำลองชุดข้อมูลในการวิจัย	77
ภาคผนวก ข แผนผังการคำนวณแต่ละวิธีสำหรับการประมาณค่าแบบจุด และแบบช่วงที่ใช้ในงานวิจัย	90
ประวัติการศึกษา และการทำงาน	97

สารบัญตาราง

ตารางที่		หน้า
1	ตัวประมาณค่าแบบจุดสำหรับการแจกแจงทวินาม	29
2	ตัวประมาณแบบช่วงสำหรับการแจกแจงทวินาม	39
3	ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยของการประมาณทั้ง 3 วิธี	48
4	ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นของช่วงความเชื่อมั่นของพารามิเตอร์ p ที่ระดับความเชื่อมั่น 90%	54
5	ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นของช่วงความเชื่อมั่นของพารามิเตอร์ p ที่ระดับความเชื่อมั่น 95%	56
6	ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นของช่วงความเชื่อมั่นของพารามิเตอร์ p ที่ระดับความเชื่อมั่น 99%	58
7	ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นจากวิธีการประมาณทั้ง 3 วิธี ที่ระดับความเชื่อมั่น 90%	61
8	ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นจากวิธีการประมาณทั้ง 3 วิธี ที่ระดับความเชื่อมั่น 95%	63
9	ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นจากวิธีการประมาณทั้ง 3 วิธี ที่ระดับความเชื่อมั่น 99%	65
10	วิธีที่เหมาะสมในการประมาณค่าแบบจุด จำแนกตามขนาดตัวอย่าง (n) และค่าพารามิเตอร์ (p)	68
11	กรณีที่แต่ละวิธีให้ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่าค่าสัมประสิทธิ์ ความเชื่อมั่นที่กำหนดจำแนกตามขนาดตัวอย่าง (n) และค่าพารามิเตอร์ (p)	70
12	วิธีที่เหมาะสมในการประมาณค่าแบบช่วง จำแนกตามขนาดตัวอย่าง (n) และค่าพารามิเตอร์ (p)	70

สารบัญภาพ

ภาพที่		หน้า
1	การแจกแจงความหนาแน่นน่าจะเป็นของ Beta (1/2,1/2)	15
2	การแจกแจงความหนาแน่นน่าจะเป็นของ Beta (3,3)	16
3	ขั้นตอนการประมาณค่าแบบจุด	44
4	ขั้นตอนการประมาณค่าแบบช่วง	45
5	ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยสำหรับค่าพารามิเตอร์ p เท่ากับ 0.05	50
6	ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยสำหรับค่าพารามิเตอร์ p เท่ากับ 0.07	50
7	ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยสำหรับค่าพารามิเตอร์ p เท่ากับ 0.09	51
8	ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยสำหรับค่าพารามิเตอร์ p เท่ากับ 0.10	51
9	ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยสำหรับค่าพารามิเตอร์ p เท่ากับ 0.30	52
10	ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยสำหรับค่าพารามิเตอร์ p เท่ากับ 0.50	52
ภาพผนวกที่		
ข1	ขั้นตอนการประมาณค่าแบบจุดโดยวิธีภาวะน่าจะเป็นสูงสุด	91
ข2	ขั้นตอนการประมาณค่าแบบจุดโดยวิธีเบย์เซียน	92
ข3	ขั้นตอนการประมาณค่าแบบจุดโดยวิธีมินิแมกซ์	93
ข4	ขั้นตอนการประมาณค่าแบบจุดโดยวิธีปกติ	94
ข5	ขั้นตอนการประมาณค่าแบบจุดโดยวิธีโลจิส	95
ข6	ขั้นตอนการประมาณค่าแบบจุดโดยวิธีสคอร์แบบปรับความต่อเนื่อง	96

คำอธิบายสัญลักษณ์และคำย่อ

$\hat{p}$	=	ค่าสัดส่วนตัวอย่าง
n	=	ขนาดตัวอย่าง
n	=	ขนาดตัวอย่าง
σ^2	=	ค่าความแปรปรวนของประชากร
μ	=	ค่าเฉลี่ยของประชากร
$L(t; \theta)$	=	ฟังก์ชันความสูญเสีย
$L(\theta)$	=	ฟังก์ชันภาวะน่าจะเป็น
$g(\theta)$	=	ความน่าจะเป็นก่อน
$\xrightarrow{d}$	=	การลู่เข้าในเชิงการแจกแจง
$\xrightarrow{p}$	=	การลู่เข้าในเชิงความน่าจะเป็น

การศึกษาเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์สำหรับการแจกแจงทวินาม

A Comparative Study of Parameter Estimation Methods for Binomial Distribution

คำนำ

สถิติทวินาม เป็นศาสตร์ที่อาศัยทฤษฎีความน่าจะเป็นซึ่งเป็นองค์ประกอบสำคัญที่ใช้ในการวิเคราะห์ โดยเกิดขึ้นเมื่อมีการนำข้อมูลที่รวบรวมจากตัวอย่างสุ่มไปอนุมาน เพื่อนำไปสู่ข้อสรุปของผลการทดลองซึ่งจะนำไปสู่การตัดสินใจ การสรุปดังกล่าวอาจมีโอกาสผิดพลาดได้ ทั้งนี้ขึ้นอยู่กับ การประมาณว่าจะถูกต้องใกล้เคียงกับพารามิเตอร์เพียงใด โดยที่การประมาณนั้นขึ้นกับผลของข้อมูลที่รวบรวมได้จากตัวอย่างที่สุ่มมาจากประชากรนั้นๆ

การอนุมานสามารถทำได้สองลักษณะคือ การประมาณค่า (Estimation) และการทดสอบสมมติฐาน (Test of hypothesis) ซึ่งในที่นี้ผู้วิจัยพิจารณาเกี่ยวกับการประมาณค่าเท่านั้น โดยการประมาณค่าเป็นวิธีการวิเคราะห์ทางสถิติที่มีความสำคัญมาก จะพบว่าในปัจจุบันนี้องค์กรต่างๆได้นำการประมาณค่ามาใช้กันอย่างแพร่หลาย ซึ่งการประมาณค่าเป็นการนำข้อมูลจากตัวอย่างสุ่มที่มีอยู่มาประมาณค่าหรือสรุปผลเกี่ยวกับค่าพารามิเตอร์ที่สนใจว่ามีค่าเป็นเท่าไร เช่น การประมาณรายได้เฉลี่ยรายปีของห้างสรรพสินค้าแห่งหนึ่งในปีหน้าเพื่อวางแผนการตลาด การประมาณสัดส่วนของจำนวนคนที่จบการศึกษาระดับปริญญาตรีในเขตชุมชนหนึ่ง เป็นต้น การประมาณค่าดังกล่าว คือการประมาณค่าเฉลี่ย (μ) และค่าสัดส่วนประชากร (p) หรือที่เรียกว่า การประมาณพารามิเตอร์ที่สนใจนั่นเอง ในการประมาณค่าพารามิเตอร์แบ่งออกเป็นสองแบบคือ การประมาณค่าแบบจุด (Point estimation) และการประมาณค่าแบบช่วง (Interval estimation) ซึ่งการประมาณค่าแบบจุด คือ การประมาณค่าพารามิเตอร์ โดยประมาณเป็นค่าใดค่าหนึ่งเพียงค่าเดียวเท่านั้น เช่น การประมาณค่าเฉลี่ย การประมาณค่าความแปรปรวน และการประมาณค่าสัดส่วนของประชากร เป็นต้น การประมาณค่าแบบช่วง คือ การประมาณค่าพารามิเตอร์ของประชากรว่าจะอยู่ในช่วงใดช่วงหนึ่ง เช่น การประมาณค่าเฉลี่ยของประชากรอยู่ในช่วง a และ b ณ ระดับความเชื่อมั่น 95% เป็นต้น ซึ่งในการประมาณค่าแบบช่วงนี้ ขอบเขตล่างและขอบเขตบนของช่วงความเชื่อมั่นจะขึ้นอยู่กับระดับความเชื่อมั่น (Level of confidence) ในการประมาณค่าด้วย

สำหรับการวิจัยครั้งนี้ ผู้วิจัยสนใจศึกษาการประมาณค่าพารามิเตอร์ p ของการแจกแจงทวินาม ซึ่งการแจกแจงทวินามเป็นการทดลองแบบเบอร์นูลลีซ้ำๆ กัน n ครั้ง โดยทดลองแต่ละครั้งเป็นอิสระกัน ในการทดลองแต่ละครั้งมีความน่าจะเป็นที่จะเกิดลักษณะที่สนใจ p คงที่ทุกครั้ง นั่นคือกำหนดให้ $X_1, X_2, \dots, X_n$ คือ ตัวแปรสุ่มที่มีการแจกแจงแบบเบอร์นูลลีที่เป็นอิสระต่อกัน โดยมีพารามิเตอร์คือ p จะได้ว่า $Y = \sum_{i=1}^n X_i$ คือ ตัวแปรสุ่มที่มีการแจกแจงแบบทวินาม โดยมีพารามิเตอร์ n และ p ซึ่ง Y เป็นจำนวนครั้งที่เกิดลักษณะที่สนใจในตัวอย่างสุ่มขนาด n การอนุมานที่เกี่ยวกับพารามิเตอร์ p อาจทำได้ทั้งการประมาณค่าแบบจุดและการประมาณค่าแบบช่วง ในการประมาณค่าแบบจุดจะได้ว่า ตัวประมาณภาวน่าจะเป็นสูงสุด (maximum likelihood estimators) คือ $\hat{p} = Y/n$ โดย $\hat{p}$ เป็นสัดส่วนการเกิดลักษณะที่สนใจในตัวอย่างสุ่ม เป็นตัวประมาณที่มีคุณสมบัติไม่เอนเอียง และมีความแปรปรวนต่ำสุดอย่างเอกรูป (Uniformly Minimum Variance Unbiased Estimator : UMVUE) สำหรับการประมาณค่าแบบช่วงสามารถประมาณได้โดยวิธีปกติ (Normal method) ซึ่งจากทฤษฎีบทลิมิตเข้าสู่ส่วนกลาง (Central limit theorem) ได้ว่า ถ้าจำนวนครั้งของการทดลอง n มีขนาดใหญ่ ค่าสัดส่วนการเกิดลักษณะที่สนใจในตัวอย่าง ($\hat{p}$) จะมีแนวโน้มเข้าสู่การแจกแจงแบบปกติที่มีค่าเฉลี่ย p และความแปรปรวน $p(1-p)/n$ และจะสามารถหาค่าประมาณแบบช่วงที่ระดับความเชื่อมั่น $100(1-\alpha)\%$ คือ

$$(\hat{p} - Z_{1-\alpha/2} \sqrt{\hat{p}(1-\hat{p})/n}, \hat{p} + Z_{1-\alpha/2} \sqrt{\hat{p}(1-\hat{p})/n})$$

$Z_{1-\alpha/2}$ คือ ควอนไทล์ที่ $1-\alpha/2$ ของการแจกแจงปกติมาตรฐาน

วิธีปกติเป็นวิธีที่คำนวณง่าย แต่พบว่ามีข้อเสียก็คือ มีแนวโน้มที่จะให้ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นจริงต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด เมื่อ n เล็กและ p เข้าใกล้ 0 หรือ 1 จึงมีข้อแนะนำจากนักวิชาการหลายท่าน แนะนำว่าควรใช้วิธีปกติประมาณการแจกแจงทวินามโดย Aczel (1993); Anderson *et al.*, (1994) และ Creighton (1994) แนะนำว่าควรใช้วิธีปกติประมาณการแจกแจงทวินามเมื่อ np และ $n(1-p)$ มากกว่า 5 Larson (1995) แนะนำว่าควรใช้วิธีปกติประมาณการแจกแจงทวินามเมื่อ $n > 9p^*$ เมื่อ $p^* = \max\{(1-p)/p, p/(1-p)\}$ Ross (1994) แนะนำว่าควรใช้วิธีปกติประมาณการแจกแจงทวินามเมื่อ $np(1-p) \geq 10$

ในงานวิจัยครั้งนี้ ผู้วิจัยสนใจศึกษาเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ p สำหรับการแจกแจงทวินาม โดยพิจารณาวิธีการประมาณค่าแบบจุด ได้แก่ วิธีภาวน่าจะเป็นสูงสุด วิธีเบส์เซียน วิธีมินิแมกซ์ และวิธีการประมาณค่าแบบช่วง ได้แก่ วิธีปกติ วิธีโลจิส วิธีสคอร์แบบปรับความต่อเนื่อง เพื่อหาวิธีที่เหมาะสมภายใต้ขอบเขตของการศึกษาเดียวกัน

วัตถุประสงค์

1. ศึกษาเปรียบเทียบวิธีการประมาณค่าแบบจุด สำหรับพารามิเตอร์ p ได้แก่
 - 1.1 วิธีภาวน่าจะเป็นสูงสุด
 - 1.2 วิธีเบส์เซียน
 - 1.3 วิธีมินิแมกซ์
2. ศึกษาเปรียบเทียบวิธีการประมาณค่าแบบช่วง สำหรับพารามิเตอร์ p ได้แก่
 - 2.1 วิธีปกติ
 - 2.2 วิธีโลจิส
 - 2.3 วิธีสคอร์แบบปรับความต่อเนื่อง
3. หาข้อสรุปในการเลือกใช้วิธีการประมาณที่เหมาะสม ในแต่ละสถานการณ์ของค่า n และ p

ขอบเขตการวิจัย

การศึกษาในงานวิจัยนี้ ทำการศึกษาภายใต้ขอบเขต ดังนี้

1. กำหนดขนาดตัวอย่างที่ใช้ในการศึกษา ได้แก่ 5,10,30,70,100 และ 200
2. กำหนดค่าพารามิเตอร์ p ได้แก่ 0.05,0.07,0.09,0.10,0.30 และ 0.50
3. กำหนดค่าระดับความเชื่อมั่น 3 ระดับ คือ 90%,95% และ 99%
4. สำหรับวิธีเบส์เซียน กำหนดความน่าจะเป็นก่อน คือ $\text{Beta}(1/2,1/2)$ เมื่อ p มีค่าเท่ากับ 0.05, 0.07, 0.09, 0.10 และ $\text{Beta}(3,3)$ เมื่อ p มีค่าเท่ากับ 0.30 และ 0.50

5. การวิจัยครั้งนี้ข้อมูลที่ใช้ในการวิจัยได้มาจากการจำลองโดยใช้โปรแกรม Microsoft Visual Basic6

6. เกณฑ์ที่ใช้ในการประมาณค่าแบบจุด ทำการเปรียบเทียบค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean absolute error) โดยวิธีการใดที่ให้ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยต่ำสุด จะเป็นวิธีที่เหมาะสมที่สุดสำหรับสถานการณ์นั้นๆ

7. เกณฑ์ที่ใช้ในการประมาณค่าแบบช่วง ทำการเปรียบเทียบค่าประมาณสัมประสิทธิ์ความเชื่อมั่นและค่าความกว้างเฉลี่ยของค่าประมาณแบบช่วง ที่คำนวณได้ในแต่ละสถานการณ์การทดลอง โดยวิธีการประมาณใดสามารถให้ค่าความกว้างเฉลี่ยของค่าประมาณแบบช่วงต่ำสุดที่ระดับความเชื่อมั่นเดียวกัน จะถือว่าวิธีประมาณนั้นให้ช่วงประมาณที่เหมาะสมที่สุดสำหรับสถานการณ์นั้นๆ

ประโยชน์ที่คาดว่าจะได้รับ

1. ผลที่ได้จากการวิจัยจะเป็นแนวทางในการเลือกวิธีที่เหมาะสมในแต่ละสถานการณ์ n และ p ในการประมาณค่าแบบจุดและแบบช่วงของการแจกแจงแบบทวินาม

2. เป็นแนวทางในการศึกษาเปรียบเทียบวิธีการประมาณค่าแบบจุดและแบบช่วง สำหรับการประมาณแบบอื่นหรือการแจกแจงอื่นอีกต่อไป

การตรวจเอกสาร

การตรวจเอกสารสำหรับการศึกษาวิจัยครั้งนี้แบ่งออกเป็น 2 ส่วน คือ งานวิจัยที่เกี่ยวข้อง และวิธีการทางสถิติ

งานวิจัยที่เกี่ยวข้อง

จอมขวัญ(2548) ศึกษาเปรียบเทียบวิธีประมาณค่าแบบช่วงของสัดส่วนประชากรที่มีการแจกแจงแบบทวินาม โดยใช้ 5 วิธี คือ การประมาณการแจกแจงทวินามด้วยการแจกแจงปกติ วิธีปรับค่าของวาลด์ วิธีปรับค่าของวาลด์โดยใช้วิธีแจกแจงไบนารีในการประมาณค่าความแปรปรวน วิธีเอ็กแซ็ก และวิธีบูตสเตรปเปอร์เซ็นต์ไทล์ กำหนดขนาดตัวอย่าง n มีค่าตั้งแต่ 5 ถึง 50 ค่าสัดส่วนประชากร p มีทั้งหมด 10 ค่าคือ 0.01,0.03,0.05,0.07,0.09,0.1,0.2,0.3,0.4 และ 0.5 กำหนดระดับความเชื่อมั่น 3 ระดับ ได้แก่ 90%,95% และ 99% โดยทำการทดลองซ้ำในแต่ละสถานการณ์ 5,000 ครั้ง หลักเกณฑ์ในการเปรียบเทียบจะพิจารณาจากค่าสัมประสิทธิ์ความเชื่อมั่นว่าไม่น้อยกว่าค่าที่กำหนด จากนั้นจึงตรวจสอบค่าความยาวเฉลี่ยของช่วงที่สั้นที่สุดจากวิธีประมาณทั้ง 5 วิธีผลการวิจัยสรุปว่า ที่ระดับความเชื่อมั่น 90% การประมาณค่าแบบช่วงโดยวิธีเอ็กแซ็กจะเป็นวิธีที่เหมาะสมที่สุดในทุกสถานการณ์ ที่ระดับความเชื่อมั่น 95% การประมาณค่าแบบช่วงโดยวิธีเอ็กแซ็กจะเป็นวิธีที่เหมาะสมที่สุด เมื่อมีขนาดตัวอย่าง $0.01 \leq p \leq 0.07$ และการประมาณค่าแบบช่วงโดยวิธีปรับค่าของวาลด์จะเป็นวิธีที่เหมาะสมที่สุดในทุกขนาดตัวอย่างเมื่อ $0.09 \leq p \leq 0.5$ ที่ระดับความเชื่อมั่น 99% การประมาณค่าแบบช่วงโดยวิธีปรับค่าของวาลด์จะเป็นวิธีที่เหมาะสมที่สุดในทุกสถานการณ์

พลารุช(2534) ศึกษาและสร้างตารางสำเร็จรูป แสดงช่วงความเชื่อมั่นสำหรับค่าพารามิเตอร์ p ของการแจกแจงทวินาม โดยคำนวณจากสูตรที่แท้จริงซึ่งสัมพันธ์กับการแจกแจงเอฟ และพิจารณาความเหมาะสมของวิธีการประมาณโดยใช้การแจกแจงปกติ สำหรับขนาดตัวอย่าง $n=1$ ถึง 30 ตารางสำเร็จรูปที่สร้างขึ้น ได้กำหนดสัมประสิทธิ์ความเชื่อมั่นไว้ 8 ระดับ คือ 0.05 ,0.06, 0.07, 0.08,0.09,0.95,0.98 และ 0.99 ข้อมูลที่ใช้ในการวิจัยจำลองโดยใช้ภาษาฟอร์แทรน 77 ใช้เทคนิคมอนติคาร์โลกระทำซ้ำ 1,000 ครั้ง ผลที่ได้จากการวิจัยสรุปได้ดังนี้ ช่วงความเชื่อมั่นที่สร้างจากสูตรที่แท้จริงซึ่งสัมพันธ์กับการแจกแจงเอฟ ให้ค่าระดับความเชื่อมั่นที่คำนวณได้ไม่ต่ำกว่าค่า

สัมประสิทธิ์ที่กำหนด เมื่อพิจารณาที่ค่าพารามิเตอร์ (n, p) ใดๆ ภายใต้ค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนดค่าหนึ่งๆพบว่า ค่าระดับความเชื่อมั่นของช่วงความเชื่อมั่นที่สร้างขึ้นจากสูตรที่แท้จริงซึ่งสัมพันธ์กับการแจกแจงเอฟมีค่ามากกว่าหรือเท่ากับ ค่าระดับความเชื่อมั่นของช่วงความเชื่อมั่นที่สร้างขึ้นจากสูตรการประมาณโดยใช้การแจกแจงปกติเสมอ ภายใต้เงื่อนไขตัวอย่างขนาดเล็ก ($n=1$ ถึง 30) การใช้สูตรการประมาณโดยใช้การแจกแจงปกติ ในการคำนวณช่วงความเชื่อมั่นสำหรับพารามิเตอร์ p ของการแจกแจงทวินาม ไม่มีความเหมาะสมไม่ว่าค่าพารามิเตอร์ p จะมีค่าเข้าใกล้ 0.05 หรือผลคูณ np และ $n(1-p)$ จะมากกว่า 5 ทั้งคู่ก็ตาม

สงขลา(2542) ศึกษาเปรียบเทียบวิธีประมาณค่าแบบช่วงสำหรับค่าสัดส่วนประชากรของการแจกแจงแบบทวินาม โดยทำการเปรียบเทียบค่าระดับความเชื่อมั่นและค่าความยาวเฉลี่ยของค่าประมาณแบบช่วง วิธีประมาณที่ใช้ คือ วิธีปกติ วิธีแปลงแบบอาร์คไซน์ วิธีสคอร์ วิธีปัวส์ซอง และวิธีเอฟ การเปรียบเทียบกระทำภายใต้สถานการณ์ของขนาดตัวอย่างมีค่า 2 ถึง 50 ค่าสัดส่วนประชากรมีค่า 0.01 ถึง 0.09 โดยค่าเพิ่มขึ้นทีละ 0.01 และ 0.1 ถึง 0.5 โดยค่าเพิ่มขึ้นครั้งละ 0.05 กำหนดค่าระดับความเชื่อมั่นมีค่า 90%, 95% และ 99% ข้อมูลที่ใช้ในการวิจัยได้จากการจำลองด้วยเทคนิคมอนติคาร์โลและทำการทดลองซ้ำๆกัน 2,000 ครั้ง ในแต่ละสถานการณ์ที่กำหนด ผลการวิจัยสรุปได้ว่า ค่าประมาณแบบช่วงของวิธีปกติ จะให้ค่าระดับความเชื่อมั่นไม่ต่ำกว่าระดับความเชื่อมั่นที่กำหนดและให้ค่าความยาวเฉลี่ยของค่าประมาณแบบช่วงต่ำที่สุด เมื่อขนาดตัวอย่างมีค่ามากกว่า 45 และค่าสัดส่วนประชากรมีค่าอยู่ในช่วง 0.35 ถึง 0.50 ค่าประมาณแบบช่วงของวิธีแปลงแบบอาร์คไซน์ จะให้ค่าระดับความเชื่อมั่นไม่ต่ำกว่าระดับความเชื่อมั่นที่กำหนดและให้ค่าความยาวเฉลี่ยของค่าประมาณแบบช่วงต่ำที่สุด เมื่อขนาดตัวอย่างมีค่ามากกว่า 30 และค่าสัดส่วนประชากรมีค่าอยู่ในช่วง 0.30 ถึง 0.50 ค่าประมาณแบบช่วงของวิธีสคอร์ จะให้ค่าระดับความเชื่อมั่นไม่ต่ำกว่าระดับความเชื่อมั่นที่กำหนดและให้ค่าความยาวเฉลี่ยของค่าประมาณแบบช่วงต่ำที่สุดในทุกระดับของขนาดตัวอย่าง โดยที่สามารถคลุมค่าสัดส่วนประชากรได้คือ ถ้าขนาดตัวอย่างมีค่าน้อยกว่า 30 ค่าสัดส่วนประชากรมีค่าอยู่ในช่วง 0.15 ถึง 0.50 และ ถ้าขนาดตัวอย่างมีค่าตั้งแต่ 30 ขึ้นไป ค่าสัดส่วนประชากรมีค่าอยู่ในช่วง 0.05 ถึง 0.50 ค่าประมาณแบบช่วงของวิธีปัวส์ซอง จะให้ค่าระดับความเชื่อมั่นไม่ต่ำกว่าระดับความเชื่อมั่นที่กำหนดในทุกระดับของขนาดตัวอย่างและค่าสัดส่วนประชากร แต่ไม่สามารถให้ค่าความยาวเฉลี่ยของค่าประมาณแบบช่วงต่ำที่สุดได้ ค่าประมาณแบบช่วงของวิธีเอฟ จะให้ค่าระดับความเชื่อมั่นไม่ต่ำกว่าระดับความเชื่อมั่นที่กำหนดและให้ค่าความยาวเฉลี่ยของค่าประมาณแบบช่วงต่ำที่สุดในทุกระดับของขนาดตัวอย่าง โดยที่สามารถคลุมค่าสัดส่วนประชากรได้ คือ ถ้าขนาดตัวอย่างมีค่าน้อยกว่า 30 ค่าสัดส่วนประชากรมีค่า

อยู่ในช่วง 0.01 ถึง 0.10 และ ถ้าขนาดตัวอย่างมีค่าตั้งแต่ 30 ขึ้นไป ค่าสัดส่วนประชากรมีค่าอยู่ในช่วง 0.01 ถึง 0.04

สาริณีษ(2546) ศึกษาเปรียบเทียบวิธีประมาณค่าแบบช่วงสำหรับพารามิเตอร์ในการแจกแจงทวินาม โดยการเปรียบเทียบค่าสัมประสิทธิ์ความเชื่อมั่นและค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นที่คำนวณโดยวิธีการแตกต่างกัน ซึ่งมี 5 วิธี คือ วิธีปกติ วิธีค่าปรับแก้เพื่อความต่อเนื่องของเยจส์ วิธีแปลงแบบอาร์คไซน์ วิธีสคอร์ และ วิธีเพิ่มจำนวนลักษณะที่สนใจและจำนวนลักษณะที่ไม่สนใจอีก 2 ค่า กำหนดขนาดตัวอย่างที่ใช้ 3 กลุ่ม คือ กลุ่มตัวอย่างขนาดเล็ก ($n < 30$) กลุ่มตัวอย่างขนาดปานกลาง ($50 \leq n \leq 100$) กลุ่มตัวอย่างขนาดใหญ่ ($n > 100$) กำหนดให้ค่าพารามิเตอร์ p มีค่าตั้งแต่ 0.05 ถึง 0.09 โดยเพิ่มค่าทีละ 0.02 และเพิ่มค่าทีละ 0.05 เมื่อ p มีค่าตั้งแต่ 0.10 ถึง 0.50 ประมาณช่วงความเชื่อมั่นโดยใช้ระดับความเชื่อมั่น 90%, 95% และ 99% ข้อมูลที่ใช้ในการวิจัยได้จากการจำลองด้วยเทคนิคมอนติคาร์โล และทำการทดลองซ้ำๆกัน 2,000 ครั้ง ในแต่ละสถานการณ์ผลการวิจัย สรุปได้ดังนี้ เมื่อกลุ่มตัวอย่างมีขนาดเล็กและขนาดปานกลาง การประมาณช่วงความเชื่อมั่นด้วยวิธีสคอร์ จะให้ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด ทั้งที่ระดับ 90%, 95% และ 99% และวิธีสคอร์มีค่าความกว้างเฉลี่ยต่ำสุดหรือใกล้ค่าต่ำสุด ในทุกค่าพารามิเตอร์ p ที่กำหนด เมื่อกลุ่มตัวอย่างมีขนาดใหญ่ ช่วงความเชื่อมั่นที่คำนวณได้จากวิธีแปลงแบบอาร์คไซน์ จะให้ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด และมีค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นต่ำสุด เมื่อ $p \leq 0.15$ และ เมื่อ $p \geq 0.20$ วิธีสคอร์มีค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นต่ำสุด สำหรับวิธีเพิ่มลักษณะที่สนใจและไม่สนใจอีก 2 ค่า เป็นวิธีที่ให้ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด แม้จะไม่สามารถให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นที่ต่ำสุดเท่ากับวิธีสคอร์ แต่สามารถให้ค่าใกล้เคียงกับค่าต่ำสุดได้ในทุกขนาดตัวอย่างและเกือบทุกค่าพารามิเตอร์ เนื่องจากวิธีนี้คำนวณง่ายและสะดวกกว่าวิธีสคอร์มาก จึงเป็นวิธีการที่เหมาะสมจะนำมาใช้ในการประมาณค่าแบบช่วงสำหรับพารามิเตอร์ในการแจกแจงทวินามได้ในทุกขนาดตัวอย่าง และทุกค่า p ที่กำหนด

Agresti and Coull(1998) ได้ทำการศึกษาเปรียบเทียบวิธีการประมาณค่าแบบช่วงสำหรับพารามิเตอร์ในการแจกแจงทวินาม ระหว่างวิธีสคอร์ และวิธีปกติ ที่ระดับความเชื่อมั่น 95% ใช้ขนาดตัวอย่าง 5, 10, 20, 30, 50, 100 ผลจากการวิจัยสรุปได้ว่า วิธีสคอร์ ให้ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นของช่วงความเชื่อมั่นไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด และมีค่าสูงกว่าค่า

สัมประสิทธิ์ความเชื่อมั่นโดยวิธีปกติ ในทุกๆขนาดตัวอย่างที่ทำการศึกษา แม้แต่ตัวอย่างขนาดเล็กที่สุดเท่ากับ 5

Agresti and Caffo(2000) ได้ทำการศึกษาเปรียบเทียบวิธีการประมาณค่าแบบช่วงสำหรับพารามิเตอร์ p ในการแจกแจงทวินาม โดยวิธีเพิ่มจำนวนลักษณะที่สนใจและจำนวนลักษณะที่ไม่สนใจอีก 2 ค่า กับวิธีปกติ ที่ระดับความเชื่อมั่น 95% และ 99% ใช้ขนาดตัวอย่าง 5,10,20 จากการศึกษาพบว่า วิธีปกติให้ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด ในทางตรงกันข้ามวิธีเพิ่มจำนวนลักษณะที่สนใจและจำนวนลักษณะที่ไม่สนใจอีก 2 ค่า มีค่าประมาณสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด ในทุกๆขนาดตัวอย่าง และทุกๆระดับความเชื่อมั่นที่กำหนด

Brown *et al.* (2002, 2003) ศึกษาการประมาณค่าแบบช่วงสำหรับค่าสัดส่วนทวินาม โดยเพิ่มวิธีทางเลือกอีก 4 วิธี คือ วิธีสคอร์ วิธีอัตราส่วนภาวะน่าจะเป็น วิธีเบส์เซียน (โดยกำหนดความน่าจะเป็นก่อนเป็นแบบ Jeffreys) และวิธีเพิ่มจำนวนลักษณะที่สนใจและจำนวนลักษณะที่ไม่สนใจอีก 2 ค่า สาเหตุที่เพิ่มวิธีทางเลือกอื่นเพราะ วิธีปกติที่ใช้กันอยู่ทั่วไปนั้นมีคุณสมบัติการครอบคลุมพารามิเตอร์ที่ต่ำ โดยใช้คุณสมบัติการครอบคลุมและค่าความกว้างเฉลี่ยแบบช่วงเป็นเกณฑ์การเปรียบเทียบระหว่างปกติกับวิธีทางเลือกทั้ง 4 วิธี ผลการทดลองสรุปว่า สำหรับคุณสมบัติการครอบคลุมค่าพารามิเตอร์นั้น วิธีสคอร์ วิธีอัตราส่วนภาวะน่าจะเป็น วิธีเบส์เซียนให้ความน่าจะเป็นในช่วงความเชื่อมั่นจะครอบคลุมค่าพารามิเตอร์รอบๆค่าระดับความเชื่อมั่นที่กำหนด ส่วนวิธีเพิ่มจำนวนลักษณะที่สนใจและจำนวนลักษณะที่ไม่สนใจอีก 2 ค่า ให้การครอบคลุมที่ใหญ่กว่าทางด้านค่าความกว้างเฉลี่ย พบว่าวิธีเพิ่มจำนวนลักษณะที่สนใจและจำนวนลักษณะที่ไม่สนใจอีก 2 ค่า ให้ค่าความกว้างเฉลี่ยกว้างที่สุดเมื่อเทียบกับวิธีทางเลือกที่เหลืออีก 3 วิธี วิธีเบส์เซียน ให้ค่าความกว้างเฉลี่ยสั้นที่สุด แต่วิธีทางเลือกทั้ง 4 วิธี ต่างก็ให้ค่าความกว้างเฉลี่ยต่ำกว่าวิธีปกติ ซึ่งวิธีเพิ่มจำนวนลักษณะที่สนใจและจำนวนลักษณะที่ไม่สนใจอีก 2 ค่า เป็นวิธีที่ผู้วิจัยแนะนำให้ใช้ เพราะถึงแม้ว่า วิธีนี้จะให้ความน่าจะเป็นในช่วงความเชื่อมั่นจะครอบคลุมค่าพารามิเตอร์ แม้ให้ค่าความกว้างเฉลี่ยกว้างกว่าวิธีอื่น แต่ก็กว้างกว่าไปไม่มากนัก อีกทั้งวิธีนี้ยังง่ายต่อการคำนวณ จึงเป็นวิธีที่ผู้วิจัยแนะนำขึ้น อีกทั้งยังได้พิจารณาการประมาณค่าแบบช่วงสำหรับค่าเฉลี่ยในวงศักรูปไปเนนเซียลด้วยฟังก์ชันความแปรปรวนกำลังสอง การแจกแจงที่นำมาพิจารณา ได้แก่ การแจกแจงแบบทวินาม การแจกแจงแบบปัวส์ซอง การแจกแจงแบบเรขาคณิต การแจกแจงแบบปกติ การแจกแจงแบบแกมมา และ six distribution และวิธีที่นำมาใช้ในการประมาณคือ วิธีของ วาลด์

วิธีสคอร์ วิธีเบสส์เซียน (โดยกำหนดความน่าจะเป็นก่อนเป็นแบบเจฟเฟอรี่) และ วิธีอัตราส่วน ภาวะน่าจะเป็น โดยพบว่า วิธีเบสส์เซียน และ วิธีอัตราส่วนภาวะน่าจะเป็น ให้การประมาณค่าแบบ ช่วงที่ให้ความน่าจะเป็นในช่วงความเชื่อมั่นจะครอบคลุมค่าเฉลี่ยที่ดีกว่าวิธีของ Wald และ วิธีสคอร์

วิธีการทางสถิติ

1. ทฤษฎีบทลิมิตเข้าสู่ส่วนกลาง

ทฤษฎีบทลิมิตเข้าสู่ส่วนกลาง ถูกกล่าวไว้ใน Mood *et al.*(1974) ว่าเมื่อสุ่มตัวอย่างขนาดใหญ่จากประชากรที่มีการแจกแจงแบบใดก็ตามที่มีความแปรปรวนเป็นค่าจำกัด ค่าเฉลี่ยตัวอย่าง จะมีการแจกแจงเข้าสู่การแจกแจงแบบปกติ นั่นคือ

ถ้า $\bar{Y}$ คือ ค่าเฉลี่ยของตัวอย่างสุ่ม $Y_1, Y_2, \dots, Y_n$ จากการแจกแจงที่มีค่าเฉลี่ย μ และมีค่า ความแปรปรวน $\sigma^2 < \infty$ แล้ว เมื่อ $n \rightarrow \infty$ ได้ว่า

$\bar{Y}$ มีการแจกแจงใกล้เคียงการแจกแจงปกติ มีค่าเฉลี่ย μ และความแปรปรวน $\sigma_{\bar{Y}}^2$
หรือ

$$Z = \frac{\bar{Y} - \mu}{\sigma / \sqrt{n}} \text{ มีการแจกแจงใกล้เคียงการแจกแจงปกติ มีค่าเฉลี่ย 1 และความแปรปรวน 0}$$

2. การแจกแจงของประชากร (Distribution of Population)

การแจกแจงของประชากร ถูกกล่าวไว้ใน Freund (1992) โดยการแจกแจงของประชากร หมายถึง การแจกแจงของค่าที่สนใจศึกษาจากข้อมูลทุกหน่วยของประชากร การประมาณค่าของ ประชากรหรือการทดสอบสมมติฐานทางสถิติ ส่วนมากหรือเกือบทั้งหมดนั้นมักมีข้อกำหนด เบื้องต้นว่า กลุ่มตัวอย่างประชากรที่สุ่มได้นั้น มาจากประชากรที่ทราบล่วงหน้าว่ามีลักษณะการ แจกแจงแบบใดแบบหนึ่ง เช่น การแจกแจงของประชากรมีลักษณะแบบปกติ การแจกแจง แบบปัวส์ซอง การแจกแจงแบบเอกซ์โปเนนเชียล การแจกแจงแบบแกมมา การแจกแจงแบบทวินาม เป็นต้น

3. การแจกแจงเบอร์นูลลี

การทดลองสุ่มที่มีผลลัพธ์เป็นไปได้อีก 2 อย่าง คือ ลักษณะที่สนใจและลักษณะที่ไม่สนใจ ซึ่งเกิดขึ้นด้วยความน่าจะเป็น p และ $1-p$ ตามลำดับ การทดลองนี้เรียกว่า การทดลองเบอร์นูลลี (Bernoulli experiment) ถ้าให้ $X = 1$ เมื่อเกิดลักษณะที่สนใจ และ $X = 0$ เมื่อเกิดลักษณะที่ไม่สนใจ ตัวแปรสุ่ม X จะมีการแจกแจงความน่าจะเป็นแบบเบอร์นูลลีที่มีพารามิเตอร์ p และมีฟังก์ชันความน่าจะเป็นในรูป

$$f(x) = p^x (1-p)^{1-x} \quad \text{เมื่อ } x = 0, 1; 0 \leq p \leq 1$$

มีค่าคาดหวังของ X เป็น $E(X) = \mu = p$ และค่าความแปรปรวนของ X เป็น $V(X) = p(1-p)$

4. การแจกแจงทวินาม

การแจกแจงทวินามเป็นการทดลองแบบเบอร์นูลลีซ้ำๆ กัน n ครั้ง โดยที่การทดลองแต่ละครั้งเป็นอิสระต่อกัน และความน่าจะเป็นของการเกิดความสำเร็จในการทดลองแต่ละครั้งคงที่ ถ้ากำหนดให้ตัวแปรสุ่ม Y แทน จำนวนครั้งของความสำเร็จจากการเกิดลักษณะที่สนใจแบบเบอร์นูลลี n ครั้ง เรียก Y ว่าเป็นตัวแปรสุ่มทวินาม ซึ่งมีฟังก์ชันความน่าจะเป็น คือ

$$f(y) = \binom{n}{y} p^y (1-p)^{n-y} \quad ; y = 0, 1, \dots, n; 0 \leq p \leq 1$$

โดยมีพารามิเตอร์คือ n และ p มีค่าคาดหวังคือ $E(Y) = np$ และค่าความแปรปรวนคือ $V(Y) = np(1-p)$

5. การประมาณค่าสัดส่วนประชากร

เมื่อทำการทดลองแบบเบอร์นูลลีซ้ำกัน n ครั้ง โดยมีค่าความน่าจะเป็นของการเกิดลักษณะที่สนใจเท่ากับ p ถ้าจำนวนรวมของการเกิดลักษณะที่สนใจคือ Y เมื่อต้องการหาค่าประมาณแบบจุดของค่าสัดส่วนประชากร จะได้ว่าค่าสัดส่วนตัวอย่าง ได้แก่ $\hat{p} = \frac{Y}{n}$

สำหรับคุณสมบัติของตัวประมาณ เนื่องจาก Y มีการแจกแจงทวินามมีพารามิเตอร์คือ n และ p มีค่าเฉลี่ย np และค่าความแปรปรวน $np(1-p)$ ดังนั้นเมื่อพิจารณาค่าเฉลี่ยของตัวประมาณ จะได้ว่า

$$E(\hat{p}) = \frac{1}{n} E(Y) = \frac{np}{n} = p$$

แสดงให้เห็นว่า $\hat{p}$ เป็นตัวประมาณที่ไม่เอนเอียง (Unbiased estimator) ของ p และเมื่อพิจารณาค่าความแปรปรวนของตัวประมาณ $\hat{p}$ จะได้ว่า

$$V(\hat{p}) = \frac{1}{n^2} V(Y) = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}$$

และตัวประมาณของความแปรปรวนจะได้จากค่าประมาณตัวอย่าง $\hat{p}$ ของ p นั่นคือ

$$V(\hat{p}) = \frac{\hat{p}(1-\hat{p})}{n}$$

ถ้า n มีขนาดใหญ่ จากทฤษฎีบทลิมิตเข้าสู่ส่วนกลาง $\hat{p} = \frac{Y}{n}$ สามารถประมาณได้ด้วยการแจกแจงแบบปกติที่มีค่าเฉลี่ย p และความแปรปรวน $\frac{p(1-p)}{n}$

$$\text{หรือ } \hat{p} \sim N\left(p, \frac{p(1-p)}{n}\right) \quad \text{เมื่อ } n \rightarrow \infty$$

ดังนั้น

$$Z = \frac{p - \hat{p}}{\sqrt{\frac{p(1-p)}{n}}} \sim N(0,1)$$

6. วิธีการประมาณค่าแบบจุด

6.1 วิธีภาวจะน่าจะเป็นสูงสุด

6.1.1 แนวคิดภาวจะน่าจะเป็น (likelihood approach)

วิธีภาวจะน่าจะเป็นสูงสุดถูกกล่าวไว้ใน Rohatgi and Saleh (2001) ว่าแนวคิดวิธีภาวจะน่าจะเป็นสูงสุดจะนำเอาการแจกแจงของตัวอย่างสุ่มมาพิจารณา ถ้าทราบค่าของตัวแปรสุ่มแล้ว ค่าเหล่านี้จะบอกว่าค่าประมาณของพารามิเตอร์จะเป็นค่านี้ด้วยความน่าจะเป็นสูงสุดเท่าไร กล่าวคือ ค่าประมาณนี้จะทำให้ฟังก์ชันการแจกแจงของตัวอย่างสุ่มมีค่าสูงสุด แนวคิดในการอนุมานทางสถิติแบบภาวจะน่าจะเป็นนี้จะประมาณค่า โดยพิจารณาเลือกค่า θ ที่ทำให้ฟังก์ชันภาวจะน่าจะเป็น $L(\theta; x)$ มีค่าสูงสุด นั่นคือ

$$L(\theta_k; x_1, x_2, \dots, x_n) > L(\theta_j; x_1, x_2, \dots, x_n) \quad \text{โดยที่ } j=1, 2, \dots$$

ถ้า θ_k มีคุณสมบัตินี้ก็จะได้ θ_k เป็นตัวประมาณแบบภาวจะน่าจะเป็นสูงสุดของ θ

นิยาม 1 ให้ $X_1, \dots, X_n$ เป็นตัวอย่างสุ่มจากประชากรที่มีฟังก์ชันความหนาแน่น $f(x; \theta)$ ฟังก์ชันภาวจะน่าจะเป็น (likelihood function) ของตัวอย่างสุ่ม คือ ฟังก์ชันความหนาแน่นร่วมของ $X_1, \dots, X_n$ โดยที่เป็นฟังก์ชันของพารามิเตอร์ θ แทนฟังก์ชันภาวจะน่าจะเป็นด้วยสัญลักษณ์ L หรือ $L(\theta)$ หรือ $L(\theta; x_1, \dots, x_n)$ นั่นคือ

$$L(\theta) = L(\theta; x_1, \dots, x_n)$$

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

นิยาม 2 ตัวประมาณแบบภาวะน่าจะเป็นสูงสุด (Maximum likelihood estimator) ของพารามิเตอร์ θ คือ $\hat{\theta} = \hat{\theta}(X_1, \dots, X_n)$ ที่ทำให้ฟังก์ชันภาวะน่าจะเป็นมีค่าสูงที่สุด

6.1.2 การหาตัวประมาณแบบภาวะน่าจะเป็นสูงสุด

ให้ $X_1, X_2, \dots, X_n$ เป็นตัวแปรสุ่มจากการแจกแจงเบอร์นูลลี ดังนั้น $Y = \sum_{i=1}^n X_i$ จะมีการแจกแจงแบบทวินาม (n, p)

$$f(y; p) = \binom{n}{y} p^y (1-p)^{n-y} \quad ; y = 0, 1, \dots, n$$

$$L(p) = \binom{n}{y} p^y (1-p)^{n-y}$$

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

$$\frac{d \ln L(p)}{dp} = \frac{y}{p} - \frac{n-y}{1-p} = 0$$

$$\frac{y}{p} = \frac{n-y}{1-p}$$

$$p = \frac{y}{n}$$

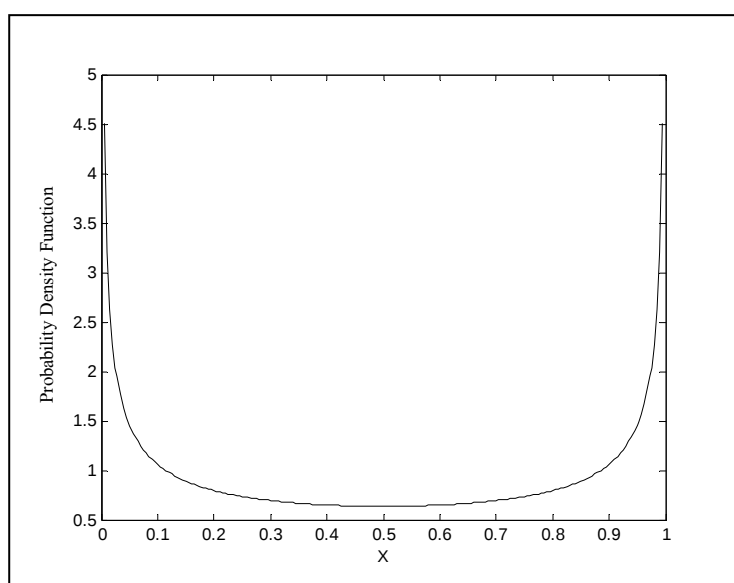
นั่นคือ ตัวประมาณภาวะน่าจะเป็นสูงสุดของ p คือ $\hat{p} = \frac{y}{n}$

6.2 วิธีเบส์เซียน

6.2.1 แนวคิดแบบเบส์เซียน (Bayesian approach)

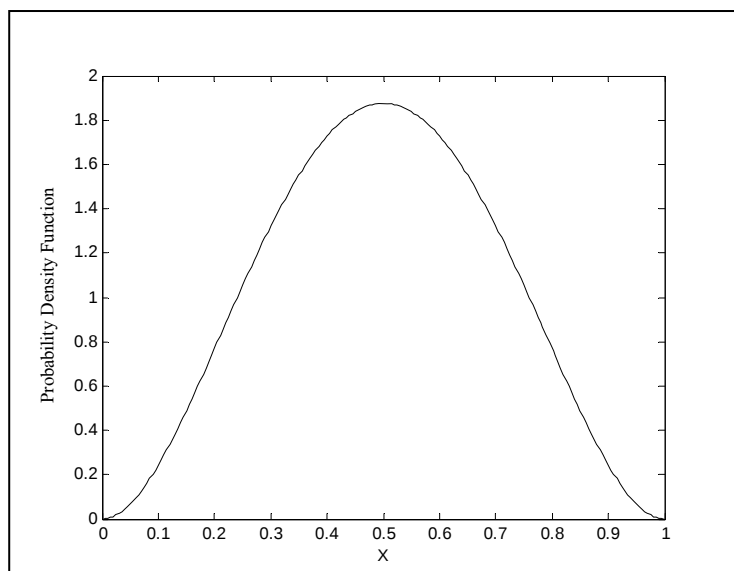
วิธีเบส์เซียน ถูกกล่าวไว้ใน Roger and Elliot (2006) ว่าสถิติตามแนวของเบส์แตกต่างจากสถิติตามแนวเดิม (Classical approach) นั่นคือ การประมาณพารามิเตอร์ θ ของสถิติแนวเดิมมักถือว่าเริ่มจากการสุ่มตัวอย่างจากประชากรที่มีฟังก์ชันความหนาแน่น $f(x; \theta)$ และถือว่าพารามิเตอร์ θ เป็นค่าคงที่แต่ไม่ทราบค่า แต่ในแนวของเบส์จะมีความพยายามใช้ความรู้เดิมเกี่ยวกับ θ ให้เป็นประโยชน์ในการประมาณ θ ให้ได้ดียิ่งขึ้น ดังนั้นจึงถือว่า θ เป็นค่าของตัวแปรสุ่ม Θ มีการแจกแจงความน่าจะเป็นรูปใดรูปหนึ่ง

เกณฑ์ในการเลือกตัวประมาณแบบเบส์จึงสนใจที่จะนำการแจกแจงพารามิเตอร์มาใช้ โดยการหาฟังก์ชันการเสี่ยงเทียบกับความน่าจะเป็นก่อน แล้วหาว่าตัวประมาณใดจะให้ฟังก์ชันการเสี่ยงแบบเบส์น้อยที่สุด เนื่องจากในการวิจัยนี้ สนใจประมาณค่าพารามิเตอร์ ซึ่งมีค่าอยู่ในช่วง 0 ถึง 1 ดังนั้นจึงเลือกความน่าจะเป็นก่อนเป็นบีตา เพราะค่าของตัวแปรสุ่มบีตา มีค่าอยู่ในช่วง 0 ถึง 1 และทำการเลือกค่า $\alpha = \beta$ เพราะกราฟความหนาแน่นน่าจะเป็นมีลักษณะสมมาตรกัน ดังภาพที่ 1



ภาพที่ 1 การแจกแจงความหนาแน่นน่าจะเป็นของ Beta (1/2,1/2)

จากภาพที่ 1 จะเห็นว่า ค่าความหนาแน่นน่าจะเป็นของ Beta (1/2,1/2) สูง เมื่อ p เข้าใกล้ 0 หรือ 1



ภาพที่ 2 การแจกแจงความหนาแน่นน่าจะเป็นของ Beta (3,3)

จากภาพที่ 2 จะเห็นว่า ค่าความหนาแน่นน่าจะเป็นของ Beta (3,3) สูง เมื่อ p เข้าใกล้ 0.5

ดังนั้นในการวิจัยครั้งนี้ จะเลือกความน่าจะเป็นก่อน คือ Beta (1/2,1/2) เมื่อ p เข้าใกล้ 0 หรือ 1 และ Beta (3,3) เมื่อ p เข้าใกล้ 0.5

ให้ $X_1, X_2, \dots, X_n$ เป็นตัวอย่างสุ่มจากประชากรที่มีฟังก์ชันความหนาแน่น $f(x; \theta) = f(x|\theta)$ โดยที่ θ เป็นค่าของตัวแปรสุ่ม Θ ที่เป็นไปได้ทั้งหมด Θ มีความน่าจะเป็นก่อนเป็น $g(\theta)$ และ $f(x|\theta)$ เป็นฟังก์ชันความหนาแน่นแบบมีเงื่อนไข (Conditional density function) ของตัวแปรสุ่ม X เมื่อกำหนดให้ $\Theta = \theta$

นิยาม 3 ฟังก์ชัน $g(\theta)$ เรียกว่า ความน่าจะเป็นก่อน (Prior probability) ของ Θ และ $h(\theta|X_1, \dots, X_n)$ เรียกว่า ฟังก์ชันความหนาแน่นภายหลัง (Posterior probability density function) หรือ ฟังก์ชันความหนาแน่นปรับแก้ (Revised probability density function) ของ Θ

ฟังก์ชันความหนาแน่นร่วมของ $X_1, \dots, X_n$ เมื่อกำหนด $\Theta = \theta$ ให้ คือ

$$f(x_1, \dots, x_n | \theta) = f(x_1 | \theta) \dots f(x_n | \theta)$$

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

ฟังก์ชันความหนาแน่นร่วมของ $X_1, \dots, X_n$ และ θ คือ

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

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

ดังนั้น ฟังก์ชันการแจกแจงภายหลังหรือฟังก์ชันโพสทีเรียของ Θ คือ

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

หรือ

$$h(\theta | x_1, \dots, x_n) = \frac{\prod_{i=1}^n f(x_i | \theta) g(\theta)}{\int \prod_{i=1}^n f(x_i | \theta) g(\theta) d\theta}$$

นิยาม 4 ให้ $X_1, \dots, X_n$ เป็นตัวอย่างสุ่มจากประชากรที่มีฟังก์ชันความหนาแน่น $f(x | \theta)$ โดยที่ θ เป็นค่าของตัวแปรสุ่ม Θ ที่มีความน่าจะเป็นก่อน $g(\theta)$ ตัวประมาณแบบเบส์ภายหลัง (Posterior Bayes estimator) ของฟังก์ชัน $\tau(\theta)$ ของ θ เทียบกับความน่าจะเป็นก่อน $g(\theta)$ คือ

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

$$\hat{\tau}(\theta) = \int \tau(\theta) h(\theta | x_1, \dots, x_n) d\theta$$

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

นิยาม 5 เมื่อ t เป็นตัวประมาณของ $\tau(\theta)$ ฟังก์ชันความสูญเสีย (Loss function) $L(t; \theta)$ คือ ฟังก์ชันค่าจริง (Real valued function) ที่มีคุณสมบัติดังนี้

1. $L(t; \theta) \geq 0$ ทุกตัวสถิติ t และ $\theta \in \Theta$
2. $L(t; \theta) = 0$ เมื่อ $t = \tau(\theta)$

ถ้า $L(t; \theta) = [t - \tau(\theta)]^2$ แล้วเรียก L ว่าเป็น ฟังก์ชันความสูญเสียคลาดเคลื่อนกำลังสอง (Square error loss function)

ถ้า $L(t; \theta) = |t - \tau(\theta)|$ แล้วเรียก L ว่าเป็น ฟังก์ชันความสูญเสียคลาดเคลื่อนสัมบูรณ์ (Absolute error loss function)

นิยาม 6 ฟังก์ชันเสี่ยง (Risk function) ของตัวประมาณ $t(X_1, \dots, X_n)$ ของ $\tau(\theta)$ คือ ฟังก์ชัน $R(\theta; t)$ หรือ $R_t(\theta)$ ซึ่งเป็นค่าคาดหวังของฟังก์ชันความสูญเสีย นั่นคือ

$$R(\theta; t) = E[L(T; \theta)]$$

$$= \int \dots \int L(t(x_1, \dots, x_n); \theta) f(x_1, \dots, x_n | \theta) dx_1 \dots dx_n$$

นิยาม 7 การเสี่ยงแบบเบย์ (Bayes risk) ของตัวประมาณ $t(X_1, \dots, X_n)$ เทียบกับฟังก์ชันสูญเสีย $L(t; \theta)$ และความน่าจะเป็นก่อน $g(\theta)$ คือ $r(t)$ ซึ่งเป็นค่าคาดหวังของฟังก์ชันการเสี่ยง นั่นคือ

$$r(t) = E[R(\theta; t)]$$

$$= \int R(\theta; t) g(\theta) d\theta$$

นิยาม 8 ตัวประมาณแบบเบย์ (Bayes estimator) ของ $\tau(\theta)$ เทียบกับฟังก์ชันความสูญเสีย $L(t; \theta)$ และความน่าจะเป็นก่อน $g(\theta)$ คือ ตัวประมาณ t^* ที่มีการเสี่ยงแบบเบย์ต่ำสุด นั่นคือ

$$r(t^*) \leq r(t)$$

ทุกค่าสถิติ t ที่เป็นตัวประมาณค่าของ $\tau(\theta)$

ทฤษฎีบท 1 ให้ $X_1, \dots, X_n$ เป็นตัวอย่างสุ่มจากประชากรที่มีฟังก์ชันความหนาแน่น $f(x|\theta)$ โดยที่ θ เป็นค่าของตัวแปรเชิงสุ่ม Θ ซึ่งมีความน่าจะเป็นก่อน $g(\theta)$ ให้ $L(t; \theta)$ เป็นฟังก์ชันสูญเสียจากการใช้ $t(X_1, X_2, \dots, X_n)$ ประมาณ $\tau(\theta)$ แล้วตัวประมาณแบบเบย์ของ $\tau(\theta)$ ได้แก่ ตัวประมาณที่ทำให้การเสี่ยงแบบเบย์ภายหลังมีค่าต่ำสุด

พิสูจน์ จากนิยาม 7 จะได้ว่า

$$\begin{aligned} \text{การเสี่ยงแบบเบย์} = r(t) &= \int R(\theta; t) g(\theta) d\theta \\ &= \int \left[\int \dots \int L[t(x_1, \dots, x_n); \theta] \prod_{i=1}^n f(x_i | \theta) dx_1 \dots dx_n \right] g(\theta) d\theta \\ &= \int \dots \int \left[\int L[t(x_1, \dots, x_n); \theta] h(\theta | x_1, \dots, x_n) d\theta \right] f(x_1, \dots, x_n) dx_1, \dots, dx_n \end{aligned}$$

ดังนั้น $r(t)$ ต่ำสุดเมื่อ

$$r'(t) = \int L[t(x_1, \dots, x_n); \theta] h(\theta | x_1, \dots, x_n) d\theta \quad \text{มีค่าต่ำสุด}$$

ทฤษฎีบท 2 ให้ $X_1, \dots, X_n$ เป็นตัวอย่างสุ่มจากประชากรที่มีฟังก์ชันความหนาแน่น $f(x|\theta)$ โดยที่ θ เป็นค่าของตัวแปรเชิงสุ่ม Θ ซึ่งมีความน่าจะเป็นก่อน $g(\theta)$ ให้ $L(t; \theta) = [t - \tau(\theta)]^2$ เป็นฟังก์ชันสูญเสียคลาดเคลื่อนกำลังสองจากการใช้ $t(X_1, \dots, X_n)$ ประมาณ $\tau(\theta)$ แล้วตัวประมาณแบบเบย์ของ $\tau(\theta)$ ได้แก่ $E[\tau(\Theta) | X_1, \dots, X_n]$ ซึ่งจะเหมือนกับตัวประมาณแบบโพสทีเรียเบย์ของ $t(\theta)$

พิสูจน์ จากทฤษฎีบท 1 จะได้ว่า

$$\begin{aligned} r'(t) &= \int L[t(x_1, \dots, x_n); \theta] h(\theta | x_1, \dots, x_n) d\theta \\ &= \int [t(x_1, \dots, x_n) - \tau(\theta)]^2 h(\theta | x_1, \dots, x_n) d\theta \\ &= \int [\tau(\theta) - t(x_1, \dots, x_n)]^2 h(\theta | x_1, \dots, x_n) d\theta \end{aligned}$$

ซึ่งมีค่าต่ำสุดเมื่อ

$$t(x_1, \dots, x_n) = E[\tau(\Theta) | x_1, \dots, x_n]$$

6.2.2 การหาตัวประมาณแบบเบสส์เซียน

กำหนด ความน่าจะเป็นก่อน $g(p) = \text{Beta}(1/2, 1/2)$ ซึ่งจะให้ค่าประมาณที่เหมาะสม เมื่อค่า p เข้าใกล้ 0 และ $\text{Beta}(3, 3)$ ซึ่งจะให้ค่าประมาณที่เหมาะสม เมื่อค่า p เข้าใกล้ 0.5 ฟังก์ชันความสูญเสียอยู่ในรูปกำลังสองของความแตกต่างของตัวประมาณและพารามิเตอร์ คือ $L[p; d] = [p - d(y)]^2$ โดยที่ $d(y)$ เป็นตัวประมาณแบบเบสส์ที่ทำให้ค่าเฉลี่ยการสูญเสีย เมื่อ $Y = y$ มีค่าต่ำสุด ให้ $X_1, \dots, X_n$ เป็นตัวอย่างสุ่มที่มีการแจกแจงเบอร์นูลลีและกำหนด $L[p; d] = [p - d(y)]^2$ ดังนั้น $Y = \sum_{i=1}^n X_i$ จะมีการแจกแจงแบบทวินาม (n, p) มีความน่าจะเป็นก่อน คือ

$$g(p) = \text{Beta}(\alpha, \beta)$$

$$= \frac{1}{B(\alpha, \beta)} p^{\alpha-1} (1-p)^{\beta-1}; 0 \leq p \leq 1$$

ดังนั้น ฟังก์ชันความหนาแน่นร่วมของ p และ y คือ

$$k(p, y) = f(y|p)g(p)$$

$$= \binom{n}{y} \frac{1}{B(\alpha, \beta)} p^{y+\alpha-1} (1-p)^{n+\beta-y-1}$$

และมีฟังก์ชันความหนาแน่นส่วนริม (marginal p.d.f) ของ y เป็น

$$\begin{aligned} k(y) &= \int_0^1 f(y|p)g(p)dp \\ &= \binom{n}{y} \frac{1}{B(\alpha, \beta)} \int_0^1 p^{y+\alpha-1} (1-p)^{n+\beta-y-1} dp \\ &= \binom{n}{y} \frac{B(y+\alpha, n+\beta-y)}{B(\alpha, \beta)} \end{aligned}$$

จะได้ $k(p|y)$ เป็นฟังก์ชันความหนาแน่นอย่างมีเงื่อนไขของ p เมื่อ กำหนด $Y=y$ เราเรียก $k(p|y)$ ว่า Posterior density function ของ p ดังนี้

$$\begin{aligned} k(p|y) &= \frac{k(p, y)}{k(y)} \\ &= \frac{f(y|p)g(p)}{k(y)} \\ &= \frac{\binom{n}{y} \frac{1}{B(\alpha, \beta)} p^{(y+\alpha)-1} (1-p)^{(n+\beta-y)-1}}{\binom{n}{y} \frac{B(y+\alpha, n+\beta-y)}{B(\alpha, \beta)}} \\ &= \frac{p^{(y+\alpha)-1} (1-p)^{(n+\beta-y)-1}}{B(y+\alpha, n+\beta-y)} \end{aligned}$$

ดังนั้น ตัวประมาณแบบเบส์เชียน คือ $\hat{p}$ ที่ทำให้

$$E[L(p, k(y)) | Y = y] = \int_0^1 [p - d]^2 k(p|y) dp \quad \text{มีค่าต่ำที่สุด}$$

ซึ่งหา $\hat{p}$ ได้โดยการหาอนุพันธ์ของฟังก์ชันนี้เทียบกับ d แล้วกำหนดให้เท่ากับ 0 และเนื่องจากฟังก์ชันความสูญเสีย $L[p, d(y)]$ อยู่ในรูปกำลังสองของความคลาดเคลื่อน ดังนั้นตัวประมาณแบบเบส์ คือ ตัวประมาณแบบเบส์ภายหลัง

$$\begin{aligned} \hat{p} &= \int_0^1 p \cdot \frac{p^{(y+\alpha)-1} (1-p)^{(n+\beta-y)-1}}{B(y+\alpha, n+\beta-y)} dp \\ &= \frac{1}{B(y+\alpha, n+\beta-y)} \int_0^1 p^{(y+\alpha+1)-1} (1-p)^{(n+\beta-y)-1} dp \\ &= \frac{B(y+\alpha+1, n+\beta-y)}{B(y+\alpha, n+\beta-y)} \\ \hat{p} &= \frac{y+\alpha}{\alpha+\beta+n} \end{aligned}$$

กำหนดความน่าจะเป็นก่อนเป็น Beta (1/2, 1/2) เมื่อ p เข้าใกล้ 0 หรือ 1

ดังนั้น ตัวประมาณเบส์เชียนของ p คือ $\hat{p} = \frac{y+0.5}{n+1}$

กำหนดความน่าจะเป็นก่อนเป็น Beta (3, 3) เมื่อ p มีค่าเข้าใกล้ 0.5

ดังนั้น ตัวประมาณเบส์เชียนของ p คือ $\hat{p} = \frac{y+3}{n+6}$

6.3 วิธีมินิแมกซ์

6.3.1 แนวคิดแบบมินิแมกซ์ (Minimax approach)

วิธีมินิแมกซ์ได้ถูกกล่าวอ้างไว้ใน William (2004) ว่า ในการประมาณค่าพารามิเตอร์ โดยการพิจารณาฟังก์ชันสูญเสีย มักต้องการตัวประมาณที่มีความสูญเสียต่ำ สิ่งที่ต้องพิจารณาจึงได้แก่ ค่าคาดหวังของความสูญเสีย หรือการเสี่ยง นั่นเอง วิธีการแบบมินิแมกซ์ซึ่งใช้ฟังก์ชันการเสี่ยงในการเปรียบเทียบ โดยการหาตัวประมาณซึ่งทำให้ฟังก์ชันการเสี่ยงที่จะเกิดมากที่สุดมีค่าน้อยที่สุด เพราะว่าในปัญหาชีวิตประจำวันบางประการ เราสนใจที่จะหลีกเลี่ยงการเสี่ยงที่จะเกิดขึ้นมากที่สุดซึ่งอาจจะเป็นอันตรายถึงชีวิต เช่น การรักษาโรคมะเร็งอาจจะมียา 2 ชนิด คือ ชนิด ก. และ ข. ซึ่งในระดับยาระดับหนึ่งยาชนิด ก.อาจจะให้ผลดีในการรักษามากกว่าชนิด ข. แต่ยาชนิด ก. อาจมีคุณสมบัติที่เป็นอันตรายต่อชีวิต ในขณะที่ยาชนิด ข. ไม่มี เมื่อเพิ่มระดับยาตัวนี้หรือต้องใช้ติดต่อกันเป็นเวลานาน ดังนั้น แพทย์ก็ควรใช้ยาชนิด ข. มากกว่ายาชนิด ก.

นิยาม 9 ให้ $X_1, \dots, X_n$ เป็นตัวอย่างสุ่มขนาด n จากประชากรที่มีฟังก์ชันการแจกแจง $f(\underline{x}|\theta)$ โดยที่ θ เป็นตัวแปรสุ่มซึ่งมีความน่าจะเป็นก่อน $g(\theta)$ และมีฟังก์ชันการแจกแจงโพสทีเรีย $h(\theta|\underline{x})$ ให้ t_1 และ t_2 เป็นตัวประมาณของ $\tau(\theta)$ ตัวประมาณ t_1 ดีกว่าตัวประมาณ t_2 ของ $\tau(\theta)$ ก็ต่อเมื่อ $R(\theta; t_1) \leq R(\theta; t_2)$ สำหรับทุกค่า $\theta \in \Omega$ และ $R(\theta; t_1) < R(\theta; t_2)$ อย่างน้อยค่าหนึ่งของ $\theta \in \Omega$

นิยาม 10 t_1 เป็นตัวประมาณแบบมินิแมกซ์ (Minimax estimator) ของ $\tau(\theta)$ ก็ต่อเมื่อ

$$\sup R(\theta; t_1) \leq \sup R(R; t)$$

สำหรับทุกตัวประมาณ t ของ $\tau(\theta)$

กล่าวคือตัวประมาณแบบมินิแมกซ์ t_1 เป็นตัวประมาณที่มีการเสี่ยงสูงสุด (Maximum risk) น้อยกว่าหรือเท่ากับการเสี่ยงสูงสุดของตัวประมาณ t ใดๆ

ทฤษฎีบท 3 ถ้า t_1 เป็นตัวประมาณเบย์เซียน (Bayesian estimator) ของ $\tau(\theta)$ ที่มีการเสี่ยงคงที่ (Constant risk) กล่าวคือ $R(\theta; t_1) = K$ ซึ่งมีค่ามากกว่า 0 ทุกๆค่า θ แล้ว t_1 จะเป็นตัวประมาณแบบมินิแมกซ์ของ $\tau(\theta)$

$$\begin{aligned}\text{พิสูจน์ } \sup R(\theta; t_1) &= R(\theta; t_1) \text{ คงที่} \\ &= \int R(\theta; t_1) g(\theta) d\theta = r(t_1) \\ &\leq r(t) = \int R(\theta; t) g(\theta) d\theta \\ &\leq \sup R(\theta; t) \int g(\theta) d\theta = \sup R(\theta; t)\end{aligned}$$

ไม่ว่า t จะเป็นตัวประมาณใดๆของ $\tau(\theta)$

นั่นคือ t_1 เป็นตัวประมาณแบบมินิแมกซ์ของ $\tau(\theta)$

6.3.2 การหาตัวประมาณมินิแมกซ์

จาก ตัวประมาณเบย์เซียนของ p คือ $\hat{p} = \frac{y + \alpha}{\alpha + \beta + n}$ และ กำหนด ฟังก์ชันการสูญเสียเป็นแบบคลาดเคลื่อนกำลังสอง ตัวประมาณแบบมินิแมกซ์ของ p คือ ตัวประมาณแบบเบย์ของ p ที่มีการเสี่ยงคงที่

พิจารณาตัวประมาณซึ่งอยู่ในรูปของ $aY + b$ โดยที่ $a = \frac{1}{\alpha + \beta + n}$ และ $b = \frac{\alpha}{\alpha + \beta + n}$ จะได้ว่า

$$\begin{aligned}R(p, aY + b) &= E[(aY + b - p)^2] \\ &= E[\{a(Y - np) + anp - p + b\}^2]\end{aligned}$$

$$= a^2 \text{Var}(Y) + \{(an-1)p + b\}^2$$

$$= a^2 np(1-p) + (b - p + nap)^2$$

$$= p^2 [(na-1)^2 - na^2] + p[na^2 + 2(na-1)b] + b^2$$

ต้องการเลือกค่า a และ b ซึ่งทำให้สัมประสิทธิ์ของ p และ p^2 มีค่าเท่ากับ 0 กล่าวคือ $(na-1)^2 - na^2 = 0$ และ $na^2 + 2(na-1)b = 0$

พิจารณา $(na-1)^2 - na^2 = 0$

$$n(n-1)a^2 - 2na + 1 = 0$$

$$a = \frac{2n \pm \sqrt{4n^2 - 4n(n-1)}}{2n(-1)}$$

$$a = \frac{1}{n \pm \sqrt{n}}$$

กรณี 1 $a = \frac{1}{n + \sqrt{n}}$ หรือ $a = \frac{1}{\sqrt{n}(\sqrt{n} + 1)}$, $b = \frac{1}{2(\sqrt{n} + 1)}$

กรณี 2 $a = \frac{1}{n - \sqrt{n}}$ หรือ $a = \frac{1}{\sqrt{n}(\sqrt{n} - 1)}$, $b = \frac{-1}{2(\sqrt{n} - 1)}$

พิจารณากรณี 1 จากที่กำหนดให้ $a = \frac{1}{\alpha + \beta + n}$, $b = \frac{\alpha}{\alpha + \beta + n}$ แทนค่า a, b เพื่อหาค่า α, β

จาก $a = \frac{1}{\alpha + \beta + n}$

$$\frac{1}{\sqrt{n}(\sqrt{n} + 1)} = \frac{1}{\alpha + \beta + n}$$

$$\frac{\alpha + \beta + n}{\sqrt{n}(\sqrt{n} + 1)} = 1 \quad \dots (1)$$

และ
$$b = \frac{\alpha}{\alpha + \beta + n}$$

$$\frac{1}{2(\sqrt{n} + 1)} = \frac{\alpha}{\alpha + \beta + n}$$

$$\frac{\alpha + \beta + n}{2\alpha(\sqrt{n} + 1)} = 1 \quad \dots (2)$$

(1) = (2) จะได้

$$\sqrt{n}(\sqrt{n} + 1) = 2\alpha(\sqrt{n} + 1)$$

$$\alpha = \frac{\sqrt{n}}{2}$$

แทนค่า $\alpha = \frac{\sqrt{n}}{2}$ จะได้ค่า $\beta = \frac{\sqrt{n}}{2}$

พิจารณากรณี 2 ทำเช่นเดียวกับกรณีแรกจะได้ $\alpha = \beta = -\frac{\sqrt{n}}{2}$ ซึ่งเป็นไปไม่ได้
เนื่องจาก $\alpha > 0, \beta > 0$ ดังนั้นจึงไม่พิจารณากรณีนี้

ดังนั้นจาก ตัวประมาณเบสส์เซียนของ p คือ
$$\hat{p} = \frac{y + \alpha}{\alpha + \beta + n}$$

จะได้ ตัวประมาณมินิแมกซ์ของ p คือ
$$\hat{p} = \frac{y + \frac{\sqrt{n}}{2}}{n + \sqrt{n}}$$

7. คุณสมบัติของตัวประมาณ

คุณสมบัติของตัวประมาณ ได้ถูกกล่าวอ้างไว้ใน อโนทัย (2539) และ สายชล (2544) โดยคุณสมบัติที่ใช้เป็นเกณฑ์ในการเลือกตัวประมาณ มีดังนี้

7.1 ตัวประมาณที่ไม่เอนเอียง (Unbiased ness estimator)

นิยาม 11 ให้ $\hat{\theta} = t(X_1, \dots, X_n)$ เป็นตัวประมาณของพารามิเตอร์ θ ถ้า $\hat{\theta}$ เป็นตัวประมาณที่ไม่เอนเอียง (Unbiased ness estimator) ของ θ แล้ว จะได้ว่า

$$E(\hat{\theta}) = \theta$$

7.2 ตัวประมาณที่มีความคงเส้นคงวา (Consistent estimator)

นิยาม 12 ให้ $X_1, \dots, X_n$ เป็นตัวอย่างสุ่มจากประชากรที่มีฟังก์ชันความหนาแน่น $f(x; \theta)$ ตัวประมาณ $\hat{\theta}$ ที่ได้จากตัวอย่างสุ่มขนาด n จะเป็นตัวประมาณที่คงเส้นคงวาของ θ ถ้า $\hat{\theta}$ ลู่เข้าในเชิงความน่าจะเป็นไปสู่ θ นั่นคือ เมื่อ ε เป็นจำนวนจริงบวกที่มีค่าน้อยมาก จะได้ว่า

$$\lim_{n \rightarrow \infty} P(|\hat{\theta} - \theta| > \varepsilon) = 0$$

$$\lim_{n \rightarrow \infty} P(|\hat{\theta} - \theta| < \varepsilon) = 1$$

ทฤษฎีบท 4 ให้ $X_1, \dots, X_n$ เป็นตัวอย่างสุ่มจากประชากรที่มีฟังก์ชันความหนาแน่น $f(x; \theta)$ ตัวประมาณ $\hat{\theta}$ จะเป็นตัวประมาณที่คงเส้นคงวาของ θ ก็ต่อเมื่อ

$$\lim_{n \rightarrow \infty} E(\hat{\theta}) = \theta$$

$$\lim_{n \rightarrow \infty} V(\hat{\theta}) = 0$$

7.3 ตัวประมาณที่มีความพอเพียง (Sufficient estimator)

ทฤษฎีบท 5 ให้ $X_1, \dots, X_n$ เป็นตัวอย่างสุ่มจากประชากรที่มีฟังก์ชันความหนาแน่น $f(x; \theta)$ สถิติ $t(X_1, \dots, X_n)$ เป็นสถิติที่มีความพอเพียงของ θ ก็ต่อเมื่อมีฟังก์ชันที่ไม่เป็นลบสองฟังก์ชัน คือ $g(t(x_1, \dots, x_n); \theta)$ และ $H(x_1, \dots, x_n)$ ที่ทำให้

$$\begin{aligned} f(x_1, \dots, x_n; \theta) &= f(x_1; \theta) \dots f(x_n; \theta) \\ &= g(t(x_1, \dots, x_n); \theta) H(x_1, \dots, x_n) \end{aligned}$$

7.4 ตัวประมาณที่มีความแปรปรวนต่ำสุด (Minimum variance estimator)

นิยาม 13 $\hat{\theta}$ เป็นตัวประมาณที่มีความแปรปรวนต่ำสุดของ θ ถ้า $\hat{\theta}$ มีความแปรปรวนน้อยกว่าหรือเท่ากับความแปรปรวนของตัวประมาณใดๆ ของ θ

7.5 ตัวประมาณ UMVUE (Uniformly minimum variance unbiased estimator)

นิยาม 14 ให้ $X_1, \dots, X_n$ เป็นตัวอย่างสุ่มจากประชากรที่มีฟังก์ชันความหนาแน่น $f(x; \theta)$ ให้ $t(X_1, \dots, X_n)$ เป็นตัวประมาณที่ไม่เอนเอียงของ θ จะเรียก t ว่าเป็นตัวประมาณ UMVUE ถ้าความแปรปรวนของ t มีค่าเท่ากับ ขอบเขตล่างของคราเมอร์ราว

8. เกณฑ์ที่ใช้เปรียบเทียบการประมาณค่าแบบจุด

ใช้การเปรียบเทียบค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ย โดยสนใจในค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยของวิธีที่มีค่าต่ำกว่า

ตารางที่ 1 ตัวประมาณค่าแบบจุดสำหรับการแจกแจงทวินาม

วิธี	ตัวประมาณ
วิธีภาวน่าจะเป็นสูงสุด	$\hat{p} = \frac{y}{n}$
วิธีเบส์เซียน	$\hat{p} = \frac{y+0.5}{n+1}$ เมื่อ p มีค่าเข้าใกล้ 0 หรือ 1 $\hat{p} = \frac{y+3}{n+6}$ เมื่อ p มีค่าเข้าใกล้ 0.5
วิธีมินิแมกซ์	$\hat{p} = \frac{y + \frac{\sqrt{n}}{2}}{n + \sqrt{n}}$

9. วิธีการประมาณค่าแบบช่วง

9.1 การประมาณค่าพารามิเตอร์แบบช่วงโดยวิธีปกติ

วิธีปกติ ถูกกล่าวอ้างไว้ใน Brown *et al.*, (2001) นั่นคือ จากทฤษฎีบทลิมิตเข้าสู่ส่วนกลางกล่าวว่า ถ้าตัวอย่างมีขนาดใหญ่ ค่าสัดส่วนตัวอย่าง $\hat{p}$ จะมีการแจกแจงโดยประมาณแบบปกติที่มีค่าเฉลี่ย p และความแปรปรวน $\frac{p(1-p)}{n}$ นั่นคือ เมื่อ $n \rightarrow \infty$ $\hat{p}$ มีการแจกแจงใกล้เคียงการแจกแจงปกติ มีค่าเฉลี่ย p และความแปรปรวน $\frac{p(1-p)}{n}$ และได้ว่า

$$Z = \frac{\hat{p} - p}{\sqrt{p(1-p)/n}}$$

มีการแจกแจงใกล้เคียงการแจกแจงปกติ มีค่าเฉลี่ย 0 และความแปรปรวน 1

ดังนั้นในการประมาณค่าพารามิเตอร์ (p) หรือการประมาณสัดส่วนประชากรแบบช่วง จึงสามารถวิเคราะห์ได้จากความสัมพันธ์

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

$$P(-Z_{1-\alpha/2} < \frac{\hat{p}-p}{\sqrt{p(1-p)/n}} < Z_{1-\alpha/2}) = 1-\alpha$$

$$P\left[\hat{p}-Z_{1-\alpha/2}\sqrt{\frac{p(1-p)}{n}} < p < \hat{p}+Z_{1-\alpha/2}\sqrt{\frac{p(1-p)}{n}}\right] = 1-\alpha$$

แต่เนื่องจากไม่ทราบค่า p จึงไม่ทราบค่าความแปรปรวน $\frac{p(1-p)}{n}$ ดังนั้น จะประมาณด้วยค่าความแปรปรวนตัวอย่าง คือ $\frac{\hat{p}(1-\hat{p})}{n}$ และได้ค่าประมาณแบบช่วงของ p ที่ระดับความเชื่อมั่น $(1-\alpha)100\%$ คือ

$$P_L = \hat{p} - Z_{1-\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$P_U = \hat{p} + Z_{1-\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

9.2 วิธี โลจิส

วิธีโลจิส ถูกกล่าวอ้างไว้ใน Brown *et al.*, (2001) ว่า สำหรับวิธีปกติ สมมติค่า $\hat{p} = 0.1$, $n = 20$, $\alpha = 0.05$ จะได้ $(1-\alpha)100\% = 1.96$ ดังนั้น ช่วงความเชื่อมั่น คือ $\hat{p} \pm 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.1 \pm 0.131 = (-0.031, 0.231)$ ซึ่งค่าขอบเขตล่างอยู่นอกปฏิกิริยาพารามิเตอร์ $p \in [0,1]$ วิธีการหนึ่งที่จะช่วยแก้ปัญหานี้ คือ การแปลงแบบโลจิสติก

Cox and Snell (1989) ได้นำหลักการแปลงแบบโลจิสติกมาใช้ในการประมาณค่าพารามิเตอร์แบบช่วง โดยกำหนดให้ $\text{logit } p = \log\left(\frac{p}{1-p}\right)$ เรียกว่า log odd ratio (OR) ซึ่งเป็นอัตราส่วนของความน่าจะเป็นที่จะเกิดเหตุการณ์ที่สนใจกับความน่าจะเป็นที่จะไม่เกิดเหตุการณ์ที่สนใจ การแปลงดังกล่าวจะทำให้ปฏิกิริยาพารามิเตอร์ของ p เป็น $(-\infty, \infty)$

ทฤษฎีบท 4 วิธี เดลต้า (Delta Method)

วิธีเดลต้า ถูกกล่าวอ้างไว้ใน Casella and Berger (2002) นั่นคือ ให้ X_n เป็นลำดับของตัวแปรสุ่ม และ g เป็นฟังก์ชันของตัวแปรสุ่ม โดย g มีอนุพันธ์ที่ต่อเนื่อง ดังนั้น

$$\sqrt{n}(X_n - \theta) \xrightarrow{d} N(0, \sigma^2)$$

จะได้ว่า

$$\sqrt{n}[g(X_n) - g(\theta)] \xrightarrow{d} N(0, \sigma^2 [g'(\theta)]^2)$$

เมื่อ θ , σ^2 เป็นค่าคงที่ที่จำกัด และ $\xrightarrow{d}$ แทน การลู่เข้าสู่การแจกแจง

พิสูจน์ พิจารณาการกระจายเทย์เลอร์ ของ $g(X_n)$ รอบ $X_n = \theta$ จะได้

$$g(X_n) = g(\theta) + g'(\theta)(X_n - \theta) + \tau$$

เมื่อ τ คือ เทอมของเศษเหลือ (Remainder)

ทฤษฎีบท 5 Slutsky's Theorem

วิธีสลัทสกีถูกกล่าวอ้างไว้ใน Casella and Berger (2002) นั่นคือ ให้ $\{X_n\}, \{Y_n\}$ เป็นลำดับของตัวแปรสุ่ม และ c เป็นค่าคงที่

1. ถ้า $X_n \xrightarrow{d} X$ และ $Y_n \xrightarrow{p} 0, n \rightarrow \infty$ จะได้ว่า $X_n Y_n \xrightarrow{p} 0, n \rightarrow \infty$

2. ถ้า $X_n \xrightarrow{d} X$ และ $Y_n \xrightarrow{p} c, n \rightarrow \infty$ จะได้ว่า

ก. $X_n Y_n \xrightarrow{d} cX$

ข. $X_n + Y_n \xrightarrow{d} X + c$

ค. $\frac{X_n}{Y_n} \xrightarrow{d} \frac{X}{c}, c \neq 0$

3. ถ้า $X_n \xrightarrow{d} X$ และ $Y_n \xrightarrow{p} c, n \rightarrow \infty$ จะได้ว่า (X_n, Y_n) จะมีฟังก์ชันการแจกแจงร่วมเมื่อใกล้อนันต์ (asymptotic joint distribution function)

$$G(x, y) = \begin{cases} G(x), & y > c \\ 0 & , y < c \end{cases}$$

จาก Slutsky's Theorem จะได้ว่า เทอมเศษเหลือ $\rightarrow 0$ เมื่อ $X_n \rightarrow \theta$ ในเชิงความน่าจะเป็น ดังนั้น เทอมเศษเหลือ $\rightarrow 0$ ในเชิงความน่าจะเป็นด้วย

โดยการประยุกต์ Slutsky's Theorem และการจัดเทอมที่ไม่ใช่เทอมของเศษเหลือใหม่ โดยคูณด้วย $\sqrt{n}$ จะได้

$$\sqrt{n}[g(X_n) - g(\theta)] = g'(\theta)\sqrt{n}[X_n - \theta]$$

จาก

$$\sqrt{n}(X_n - \theta) \xrightarrow{d} N(0, \sigma^2)$$

ดังนั้น

$$\sqrt{n}[g(X_n) - g(\theta)] \xrightarrow{d} N(0, \sigma^2 [g'(\theta)]^2)$$

$$\text{พิจารณา } \lambda = g(p) = \log\left(\frac{p}{1-p}\right)$$

$$\text{ดังนั้น } \lambda = g(\hat{p}) = \log\left(\frac{\hat{p}}{1-\hat{p}}\right) = \log\left(\frac{y}{n-y}\right) \text{ เมื่อ } \hat{p} = \frac{y}{n}$$

$$\text{โดย วิธีเดลต้า แสดงว่า } \frac{(\hat{\lambda} - \lambda)}{\sqrt{V(\hat{\lambda})}} \rightarrow N(0, \sigma^2)$$

$$\text{เมื่อ } V(\hat{\lambda}) = [g'(\hat{p})]^2 V(\hat{p}) = \left[\frac{1}{\hat{p}(1-\hat{p})}\right]^2 \frac{\hat{p}(1-\hat{p})}{n} = \frac{1}{n\hat{p}(1-\hat{p})}$$

นั่นคือ ขอบเขตล่างและขอบเขตบน สำหรับการประมาณช่วงความเชื่อมั่น $100(1-\alpha)\%$ ของ λ คือ

$$\lambda_L = \hat{\lambda} - Z_{\alpha/2} V(\hat{\lambda})^{1/2} = \log\left(\frac{\hat{p}}{1-\hat{p}}\right) - Z_{\alpha/2} \left[\frac{1}{n\hat{p}(1-\hat{p})} \right]^{1/2}$$

$$\lambda_U = \hat{\lambda} + Z_{\alpha/2} V(\hat{\lambda})^{1/2} = \log\left(\frac{\hat{p}}{1-\hat{p}}\right) + Z_{\alpha/2} \left[\frac{1}{n\hat{p}(1-\hat{p})} \right]^{1/2}$$

ค่าดังกล่าวเป็นค่าขอบเขตล่างและค่าขอบเขตบนของตัวแบบโลจิสติก ผู้ใช้จึงต้องแปลงค่า กลับไปเป็น ขอบเขตล่างและขอบเขตบนของ p จะได้

$$[p_L, p_U] = [g^{-1}(\lambda_L), g^{-1}(\lambda_U)] = \left[\frac{e^{\lambda_L}}{1+e^{\lambda_L}}, \frac{e^{\lambda_U}}{1+e^{\lambda_U}} \right]$$

โดย Cox and Snell (1989) แนะนำว่าควรใช้

$$\tilde{\lambda} = \log\left(\frac{y+0.5}{n-y+0.5}\right)$$

อันเป็นผลมาจากการกระจายอนุกรมเทย์เลอร์และการปรับค่าสำหรับความเอนเอียงของ $\hat{\lambda}$ เมื่อเทียบกับตัวประมาณ λ และ Gart and Zweifel (1967) ได้มีการปรับปรุงค่าความแปรปรวนของ $\hat{\lambda}$ จะได้

$$V(\tilde{\lambda}) = \frac{(n+1)(n+2)}{n(y+1)(n-y+1)}$$

นั่นคือ ขอบเขตล่างและขอบเขตบน สำหรับการประมาณช่วงความเชื่อมั่น $100(1-\alpha)\%$ ของ λ คือ

$$\lambda_L = \hat{\lambda} - Z_{\alpha/2} V(\hat{\lambda})^{1/2} = \log\left(\frac{y+0.5}{n-y+0.5}\right) - Z_{\alpha/2} \left[\frac{(n+1)(n+2)}{n(y+1)(n-y+1)} \right]^{1/2}$$

$$\lambda_U = \hat{\lambda} + Z_{\alpha/2} V(\hat{\lambda})^{1/2} = \log\left(\frac{y+0.5}{n-y+0.5}\right) + Z_{\alpha/2} \left[\frac{(n+1)(n+2)}{n(y+1)(n-y+1)} \right]^{1/2}$$

ดังนั้น ขอบเขตล่างและขอบเขตบนสำหรับการประมาณช่วงความเชื่อมั่น $100(1-\alpha)\%$ สำหรับ p คือ

$$[p_L, p_U] = [g^{-1}(\lambda_L), g^{-1}(\lambda_U)] = \left[\frac{e^{\lambda_L}}{1+e^{\lambda_L}}, \frac{e^{\lambda_U}}{1+e^{\lambda_U}} \right]$$

9.3 วิธีสคอร์แบบปรับความต่อเนื่อง

วิธีสคอร์แบบปรับความต่อเนื่อง ถูกกล่าวอ้างไว้ใน Fleiss (2003) ว่า Ghosh (1979) และ Chen (1990) ได้เคยศึกษาวิธีประมาณค่าสัดส่วนประชากรแบบช่วง โดยใช้วิธีการของเบส์ (Bayes method) และพบว่าวิธีประมาณดังกล่าวสามารถสร้างได้จากสถิติที่ใช้สำหรับการทดสอบสมมติฐาน ซึ่งต่อมา Agresti and Coull (1998) ได้ทำการศึกษาวิธีประมาณนี้และเรียกว่าวิธีสคอร์ โดยมีหลักในการสร้างตัวประมาณ ดังนี้

สำหรับการทดสอบสมมติฐาน $H_0: p = p_0$ และ $H_1: p \neq p_0$ ที่ระดับนัยสำคัญ α พบว่าเราจะยอมรับสมมติฐานหลักก็ต่อเมื่อ

$$\frac{|\hat{p} - p_0|}{\sqrt{p_0(1-p_0)/n}} < Z_{1-\alpha/2}$$

การหาค่าขอบเขตล่างและขอบเขตบน จะหาได้จากคำตอบของสมการกำลังสอง ซึ่งถูกกำหนดโดย

$$\frac{n\{\hat{p} - p_0\}^2}{p_0(1-p_0)} = Z_{1-\alpha/2}^2$$

เมื่อ n ใหญ่ สามารถประมาณการแจกแจงแบบทวินามได้ด้วยการแจกแจงแบบปกติ แต่เนื่องจากการแจกแจงแบบทวินามเป็นการแจกแจงของตัวแปรสุ่มชนิดไม่ต่อเนื่อง (Discrete Random Variable) ในขณะที่ การแจกแจงแบบปกติเป็นการแจกแจงของตัวแปรสุ่มชนิดต่อเนื่อง (Continuous Random Variable) ดังนั้น ในการประมาณค่าแบบช่วงสำหรับพารามิเตอร์ในการแจกแจงแบบทวินามด้วยการแจกแจงแบบปกติ จึงควรปรับค่าความต่อเนื่องของข้อมูลด้วย ค่าที่ใช้ปรับ

เรียกว่า ค่าปรับความต่อเนื่อง (continuity correction) ค่าปรับความต่อเนื่องที่เยทส์เสนอไว้ (Yates's correction for continuity) มีค่าเท่ากับ $\frac{1}{2n}$ นั่นคือ สำหรับการทดสอบสมมติฐาน $H_0 : p = p_0$ และ $H_1 : p \neq p_0$ ที่ระดับนัยสำคัญ α พบว่าเราจะยอมรับสมมติฐานหลักก็ต่อเมื่อ

$$\frac{|\hat{p} - p_0| - (1/2n)}{\sqrt{p_0(1-p_0)/n}} < Z_{1-\alpha/2}$$

การหาค่าขอบเขตล่างและขอบเขตบน จะหาได้จากคำตอบของสมการกำลังสอง ซึ่งถูกกำหนดโดย

$$\frac{n \{ |\hat{p} - p_0| - (1/2n) \}^2}{p_0(1-p_0)} = Z_{1-\alpha/2}^2$$

พิจารณาการหาขอบเขตล่าง หาได้จากคำตอบของสมการกำลังสองซึ่งถูกกำหนดโดย

$$\frac{n \{ (\hat{p} - p_0) - (1/2n) \}^2}{p_0(1-p_0)} = Z_{1-\alpha/2}^2$$

เมื่อกระจายเทอมกำลังสอง สามารถเขียนได้ในรูป

$$(\hat{p} - p_0)^2 - 2(\hat{p} - p_0) \cdot \frac{1}{2n} + \frac{1}{4n^2} = Z_{1-\alpha/2}^2 [p_0(1-p_0)/n]$$

$$\hat{p}^2 - 2\hat{p}p_0 + p_0^2 - \frac{\hat{p}}{n} + \frac{p_0}{n} + \frac{1}{4n^2} - Z_{1-\alpha/2}^2 \frac{p_0}{n} + Z_{1-\alpha/2}^2 \frac{p_0^2}{n} = 0$$

$$\hat{p}^2 - p_0 \left(2\hat{p} - \frac{1}{n} + \frac{Z_{1-\alpha/2}^2}{n} \right) + p_0^2 \left(1 + \frac{Z_{1-\alpha/2}^2}{n} \right) + \frac{1}{4n^2} - \frac{\hat{p}}{n} = 0$$

$$p_0^2 \left(1 + \frac{Z_{1-\alpha/2}^2}{n} \right) - p_0 \left(2\hat{p} - \frac{1}{n} + \frac{Z_{1-\alpha/2}^2}{n} \right) + \left(\hat{p}^2 + \frac{1}{4n^2} - \frac{\hat{p}}{n} \right) = 0$$

อยู่ในรูปของสมการกำลังสอง (quadratic equation) คือ $ax^2 + bx + c = 0$ เมื่อ

$$a = \left(1 + \frac{Z_{1-\alpha/2}^2}{n}\right), b = -\left(2\hat{p} - \frac{1}{n} + \frac{Z_{1-\alpha/2}^2}{n}\right), c = \hat{p}^2 + \frac{1}{4n^2} - \frac{\hat{p}}{n}$$

แก้สมการกำลังสองและแทนค่า $p = p_0$ จะได้

$$p = \frac{2\hat{p} - \frac{1}{n} + \frac{Z_{1-\alpha/2}^2}{n} \pm \sqrt{\left(2\hat{p} - \frac{1}{n} + \frac{Z_{1-\alpha/2}^2}{n}\right)^2 - 4\left(1 + \frac{Z_{1-\alpha/2}^2}{n}\right)\left(\hat{p}^2 + \frac{1}{4n^2} - \frac{\hat{p}}{n}\right)}}{2\left(1 + \frac{Z_{1-\alpha/2}^2}{n}\right)}$$

$$\begin{aligned} \text{พิจารณา } & \sqrt{\left(2\hat{p} - \frac{1}{n} + \frac{Z_{1-\alpha/2}^2}{n}\right)^2 - 4\left(1 + \frac{Z_{1-\alpha/2}^2}{n}\right)\left(\hat{p}^2 + \frac{1}{4n^2} - \frac{\hat{p}}{n}\right)} \\ &= \sqrt{Z_{1-\alpha/2}^2 \left\{ Z_{1-\alpha/2}^2 - 2 - \frac{1}{n} + 4\hat{p} + 4n\hat{p} - 4n\hat{p}^2 \right\}} \\ &= Z_{1-\alpha/2} \sqrt{Z_{1-\alpha/2}^2 - \left(2 + \frac{1}{n}\right) + 4\hat{p}(1 + n - n\hat{p})} \end{aligned}$$

แสดงว่า

$$p = \frac{(2n\hat{p} + Z_{1-\alpha/2}^2 - 1) \pm Z_{1-\alpha/2} \sqrt{Z_{1-\alpha/2}^2 - \left(2 + \frac{1}{n}\right) + 4\hat{p}(1 + n\hat{q})}}{2(n + Z_{1-\alpha/2}^2)}$$

ดังนั้น ค่าขอบเขตล่าง คือ

$$P_L = \frac{(2n\hat{p} + Z_{1-\alpha/2}^2 - 1) - Z_{1-\alpha/2} \sqrt{Z_{1-\alpha/2}^2 - \left(2 + \frac{1}{n}\right) + 4\hat{p}(1 + n\hat{q})}}{2(n + Z_{1-\alpha/2}^2)}$$

พิจารณาการหาขอบเขตบน หาได้จากคำตอบของสมการกำลังสองซึ่งถูกกำหนดโดย

$$\frac{n\{(\hat{p}-p_0)+(1/2n)\}^2}{p_0(1-p_0)} = Z_{1-\alpha/2}^2$$

เมื่อกระจายเทอมกำลังสอง สามารถเขียนได้ในรูป

$$(\hat{p}-p_0)^2 + 2(\hat{p}-p_0) \cdot \frac{1}{2n} + \frac{1}{4n^2} = Z_{1-\alpha/2}^2 [p_0(1-p_0)/n]$$

$$\hat{p}^2 - 2\hat{p}p_0 + p_0^2 + \frac{\hat{p}}{n} - \frac{p_0}{n} + \frac{1}{4n^2} - Z_{1-\alpha/2}^2 \frac{p_0}{n} + Z_{1-\alpha/2}^2 \frac{p_0^2}{n} = 0$$

$$\hat{p}^2 - p_0 \left(2\hat{p} + \frac{1}{n} + \frac{Z_{1-\alpha/2}^2}{n} \right) + p_0^2 \left(1 + \frac{Z_{1-\alpha/2}^2}{n} \right) + \frac{1}{4n^2} + \frac{\hat{p}}{n} = 0$$

$$p_0^2 \left(1 + \frac{Z_{1-\alpha/2}^2}{n} \right) - p_0 \left(2\hat{p} + \frac{1}{n} + \frac{Z_{1-\alpha/2}^2}{n} \right) + \left(\hat{p}^2 + \frac{1}{4n^2} + \frac{\hat{p}}{n} \right) = 0$$

อยู่ในรูปของสมการกำลังสอง (Quadratic equation) คือ $ax^2 + bx + c = 0$ เมื่อ

$$a = \left(1 + \frac{Z_{1-\alpha/2}^2}{n} \right), b = - \left(2\hat{p} + \frac{1}{n} + \frac{Z_{1-\alpha/2}^2}{n} \right), c = \hat{p}^2 + \frac{1}{4n^2} + \frac{\hat{p}}{n}$$

แก้สมการกำลังสองและแทนค่า $p = p_0$ จะได้

$$p = \frac{2\hat{p} + \frac{1}{n} + \frac{Z_{1-\alpha/2}^2}{n} \pm \sqrt{\left(2\hat{p} + \frac{1}{n} + \frac{Z_{1-\alpha/2}^2}{n} \right)^2 - 4 \left(1 + \frac{Z_{1-\alpha/2}^2}{n} \right) \left(\hat{p}^2 + \frac{1}{4n^2} + \frac{\hat{p}}{n} \right)}}{2 \left(1 + \frac{Z_{1-\alpha/2}^2}{n} \right)}$$

พิจารณา
$$\sqrt{\left(2\hat{p} + \frac{1}{n} + \frac{Z_{1-\alpha/2}^2}{n}\right)^2 - 4\left(1 + \frac{Z_{1-\alpha/2}^2}{n}\right)\left(\hat{p}^2 + \frac{1}{4n^2} + \frac{\hat{p}}{n}\right)}$$

$$= \sqrt{Z_{1-\alpha/2}^2 \left\{ Z_{1-\alpha/2}^2 + 2 - \frac{1}{n} - 4\hat{p} + 4n\hat{p} - 4n\hat{p}^2 \right\}}$$

$$= Z_{1-\alpha/2} \sqrt{Z_{1-\alpha/2}^2 + \left(2 - \frac{1}{n}\right) + 4\hat{p}(n - n\hat{p} - 1)}$$

แสดงว่า

$$p = \frac{(2n\hat{p} + Z_{1-\alpha/2}^2 + 1) \pm Z_{1-\alpha/2} \sqrt{Z_{1-\alpha/2}^2 + \left(2 - \frac{1}{n}\right) + 4\hat{p}(n\hat{q} - 1)}}{2(n + Z_{1-\alpha/2}^2)}$$

ดังนั้น ค่าขอบเขตบน คือ

$$P_U = \frac{(2n\hat{p} + Z_{1-\alpha/2}^2 + 1) + Z_{1-\alpha/2} \sqrt{Z_{1-\alpha/2}^2 + \left(2 - \frac{1}{n}\right) + 4\hat{p}(n\hat{q} - 1)}}{2(n + Z_{1-\alpha/2}^2)}$$

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

$$P_L = \frac{(2n\hat{p} + Z_{1-\alpha/2}^2 - 1) - Z_{1-\alpha/2} \sqrt{Z_{1-\alpha/2}^2 - \left(2 + \frac{1}{n}\right) + 4\hat{p}(1 + n\hat{q})}}{2(n + Z_{1-\alpha/2}^2)}$$

$$P_U = \frac{(2n\hat{p} + Z_{1-\alpha/2}^2 + 1) + Z_{1-\alpha/2} \sqrt{Z_{1-\alpha/2}^2 + \left(2 - \frac{1}{n}\right) + 4\hat{p}(n\hat{q} - 1)}}{2(n + Z_{1-\alpha/2}^2)}$$

โดย $\hat{p} = \frac{y}{n}$

10. เกณฑ์ที่ใช้เปรียบเทียบการประมาณแบบช่วง

ทำการเปรียบเทียบค่าประมาณสัมประสิทธิ์ความเชื่อมั่นและค่าความกว้างเฉลี่ย โดยจะนำเฉพาะวิธีการประมาณที่ให้ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนดนี้ไปทำการเปรียบเทียบค่าความกว้างเฉลี่ยของค่าประมาณแบบช่วงที่ได้

ตารางที่ 2 ตัวประมาณแบบช่วงสำหรับการแจกแจงทวินาม

วิธี	ตัวประมาณ
วิธีปกติ	$P_L = \hat{p} - Z_{1-\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ $P_U = \hat{p} + Z_{1-\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$
วิธีโลจิส	$P_L = \frac{e^{\lambda_L}}{1 + e^{\lambda_L}}$ $P_U = \frac{e^{\lambda_U}}{1 + e^{\lambda_U}}$
วิธีสคอร์แบบปรับความต่อเนื่อง	$P_L = \frac{(2n\hat{p} + Z_{1-\alpha/2}^2 - 1) - Z_{1-\alpha/2} \sqrt{Z_{1-\alpha/2}^2 - \left(2 + \frac{1}{n}\right) + 4\hat{p}(1 + n\hat{q})}}{2(n + Z_{1-\alpha/2}^2)}$ $P_U = \frac{(2n\hat{p} + Z_{1-\alpha/2}^2 + 1) + Z_{1-\alpha/2} \sqrt{Z_{1-\alpha/2}^2 + \left(2 - \frac{1}{n}\right) + 4\hat{p}(n\hat{q} - 1)}}{2(n + Z_{1-\alpha/2}^2)}$

สำหรับวิธี โลจิส

$$\lambda_L = \log\left(\frac{y+0.5}{n-y+0.5}\right) - Z_{\alpha/2} \left[\frac{(n+1)(n+2)}{n(y+1)(n-y+1)} \right]^{1/2}$$

$$\lambda_U = \log\left(\frac{y+0.5}{n-y+0.5}\right) + Z_{\alpha/2} \left[\frac{(n+1)(n+2)}{n(y+1)(n-y+1)} \right]^{1/2}$$

อุปกรณ์และวิธีการ

อุปกรณ์

1. เครื่องไมโครคอมพิวเตอร์ของภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์
2. โปรแกรมภาษา Microsoft Visual Basic 6

วิธีการ

1. สร้างตัวแปรสุ่มที่มีการแจกแจงแบบสม่ำเสมอในช่วง $(0,1)$
2. สร้างตัวแปรสุ่มที่มีการแจกแจงทวินามที่มีพารามิเตอร์ n และ p โดยการสร้างตัวแปรสุ่มเบอร์นูลลีที่มีพารามิเตอร์ p จำนวน n ตัว ซึ่งเป็นอิสระกัน กำหนดให้เป็น $X_1, \dots, X_n$ ตัวแปรสุ่มเบอร์นูลลีแต่ละตัวมีค่าเป็น 1 หรือ 0 สร้างโดยใช้ค่าของตัวเลขสุ่มในข้อ 1. มาสร้างเงื่อนไขคือจะนับเป็นผลสำเร็จ (ให้ค่าเป็น 1) ถ้าตัวเลขสุ่มที่ถูกสุ่มมามีค่าน้อยกว่าหรือเท่ากับค่าพารามิเตอร์ p ที่กำหนด และตัวแปรสุ่มมีค่าเป็น 0 ถ้าตัวเลขสุ่มที่ถูกสุ่มมามีค่ามากกว่าค่าพารามิเตอร์ p ที่กำหนด เมื่อได้ตัวแปรสุ่มที่มีการแจกแจงเบอร์นูลลีโดยมีพารามิเตอร์ คือ p จำนวน n ตัว แล้วจะนำค่าของ $X_1, \dots, X_n$ มาบวกเข้าด้วยกัน ผลบวกสะสมที่ได้จะให้ป็นค่าของตัวแปรสุ่ม Y หรือ จำนวนครั้งของผลสำเร็จในการทดลอง n ครั้ง ดังนั้น ตัวแปรสุ่ม Y มีการแจกแจงทวินาม ที่มีพารามิเตอร์ n และ p ในแต่ละขนาดตัวอย่างที่กำหนด
3. เปรียบเทียบวิธีการประมาณค่าพารามิเตอร์แบบจุดและแบบช่วงโดยเกณฑ์ที่ใช้ในการตัดสินใจว่า วิธีประมาณใดเป็นวิธีที่มีความเหมาะสมมากที่สุด พิจารณาดังนี้
 - 3.1 วิธีการประมาณแบบจุด ทำการเปรียบเทียบค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean absolute error) ที่คำนวณได้ในแต่ละสถานการณ์การทดลอง

$$|e_i| = |p - \hat{p}|$$

โดยที่ e_i คือ ค่าคลาดเคลื่อนสัมบูรณ์

p คือ ค่าพารามิเตอร์ที่แท้จริง

$\hat{p}$ คือ ค่าพารามิเตอร์ที่ได้จากการประมาณ

และหาค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ย ซึ่งคำนวณจากการนำผลรวมค่าคลาดเคลื่อนสัมบูรณ์มาหารด้วยจำนวนรอบที่ทำซ้ำ (5,000) นำมาเปรียบเทียบกับ ถ้าวิธีการใดให้ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยต่ำกว่าแสดงว่าเป็นวิธีที่เหมาะสมกว่า

3.2 การประมาณค่าแบบช่วง จะทำการเปรียบเทียบค่าประมาณสัมประสิทธิ์ความเชื่อมั่นและค่าความกว้างเฉลี่ยของค่าประมาณแบบช่วง ที่คำนวณได้ในแต่ละสถานการณ์การทดลอง โดยมีรายละเอียดการเปรียบเทียบดังนี้

เกณฑ์ที่ใช้ในการเปรียบเทียบค่าประมาณสัมประสิทธิ์ความเชื่อมั่น เพื่อตรวจสอบว่าวิธีการประมาณใดให้ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนดนั้นจะอาศัยการทดสอบสมมติฐานโดยใช้ตัวสถิติทดสอบ Z มีรูปแบบดังนี้

(Ghosh, 1979)

H_0 : ค่าสัมประสิทธิ์ความเชื่อมั่นจะมีค่าไม่น้อยกว่าค่าที่กำหนด ($c \geq c_0$)

H_1 : ค่าสัมประสิทธิ์ความเชื่อมั่นจะมีค่าน้อยกว่าค่าที่กำหนด ($c < c_0$)

กำหนดระดับนัยสำคัญ $\alpha = 0.05$

$$\text{ตัวสถิติที่ใช้ทดสอบ } Z = \frac{\hat{c} - c_0}{\sqrt{c_0(1 - c_0)/n}}$$

บริเวณวิกฤติ คือ $Z < -Z_\alpha$

$$\frac{\hat{c} - c_0}{\sqrt{c_0(1-c_0)/n}} < -Z_{0.05}$$

จะปฏิเสธสมมติฐาน H_0 เมื่อ $\hat{c} < c_0 - 1.645\sqrt{\frac{c_0(1-c_0)}{n}}$

หรือ จะยอมรับสมมติฐาน H_0 เมื่อ $\hat{c} \geq c_0 - 1.645\sqrt{\frac{c_0(1-c_0)}{n}}$

นั่นคือจะยอมรับสมมติฐาน H_0 เมื่อ $\hat{c} \in \left(c_0 - 1.645\sqrt{\frac{c_0(1-c_0)}{n}}, 1 \right)$

ดังนั้นจะได้ว่า วิธีการประมาณใดให้ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่าเกณฑ์ที่กำหนด ถ้า $\hat{c}$ มีค่าอยู่ในช่วง $\left(c_0 - 1.645\sqrt{\frac{c_0(1-c_0)}{n}}, 1 \right)$

โดยที่ α_0 คือ ระดับนัยสำคัญหรือความน่าจะเป็นที่จะเกิดความผิดพลาดที่กำหนด

ในการทดสอบ สำหรับการวิจัยครั้งนี้กำหนดระดับนัยสำคัญของการทดสอบเท่ากับ 0.05

c คือ ค่าสัมประสิทธิ์ความเชื่อมั่น

$\hat{c}$ คือ ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการทดลอง

c_0 คือ ค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด (0.90, 0.95 และ 0.99)

n คือ จำนวนครั้งของการทดลอง

พิจารณา การตรวจสอบค่าระดับความเชื่อมั่นของช่วงความเชื่อมั่น 90%

$$H_0 : c \geq 0.90$$

$$H_1 : c < 0.90$$

กำหนดระดับนัยสำคัญ $\alpha = 0.05$

$$\text{ตัวสถิติที่ใช้ทดสอบ} \quad Z = \frac{\hat{c} - c_0}{\sqrt{c_0(1-c_0)/n}}$$

$$\begin{aligned} \text{บริเวณวิกฤติ คือ} \quad \hat{c} &< c_0 - 1.645 \sqrt{\frac{c_0(1-c_0)}{n}} \\ \hat{c} &< 0.90 - 1.645 \sqrt{\frac{0.90 \times 0.10}{5,000}} \end{aligned}$$

นั่นคือ จะปฏิเสธสมมติฐาน H_0 เมื่อ $\hat{c} < 0.8930$
 หรือ หรือจะยอมรับสมมติฐาน H_0 เมื่อ $\hat{c}$ อยู่ในช่วง $(0.8930, 1)$

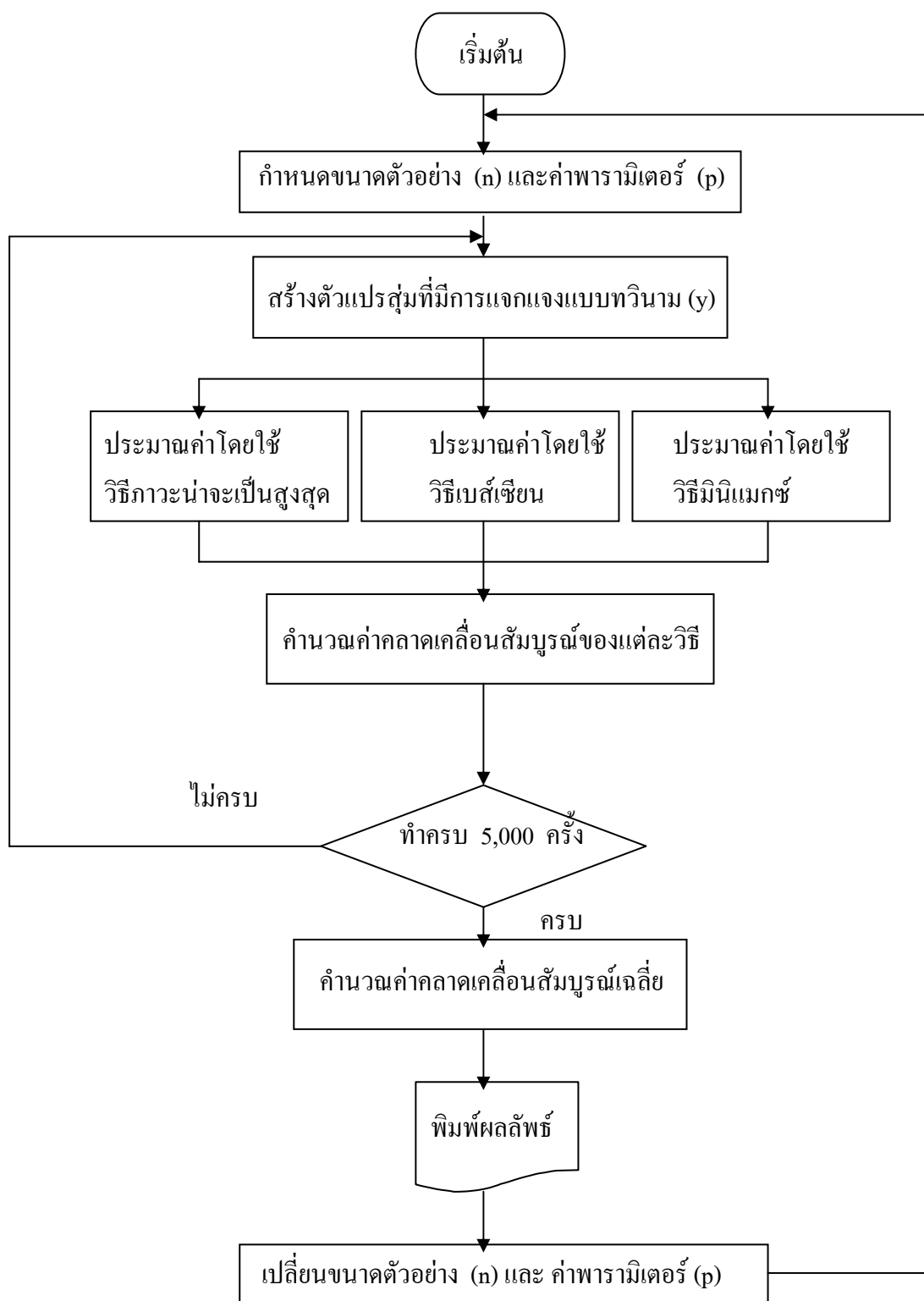
จะได้ว่า จากการทดสอบสมมติฐานที่ระดับนัยสำคัญ 0.05 ณ. ช่วงความเชื่อมั่น 90% วิธีใดมีค่าประมาณสัมประสิทธิ์ความเชื่อมั่นไม่น้อยกว่า 0.8930 แสดงว่าวิธีนั้นให้ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นไม่น้อยกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด และ ณ. ระดับความเชื่อมั่น 95% และ 99% มีวิธีการพิจารณาเช่นเดียวกันกับระดับความเชื่อมั่น 90%

นั่นคือ ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการทดลองจะต้องมีค่าไม่ต่ำกว่า 0.8930 , 0.9449 และ 0.9877 ที่ระดับความเชื่อมั่น 90%, 95% และ 99% ตามลำดับ จึงจะถือว่าวิธีการประมาณนั้นให้ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด และจะนำเฉพาะวิธีการประมาณที่ให้ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่าเกณฑ์ที่กำหนดนี้ไปทำการเปรียบเทียบค่าความกว้างเฉลี่ยของค่าประมาณแบบช่วงที่ได้ ซึ่งคำนวณจากการนำผลรวมของค่าความกว้างช่วงทั้งหมดที่ให้ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนดมาหารด้วยจำนวนซ้ำทั้งหมด

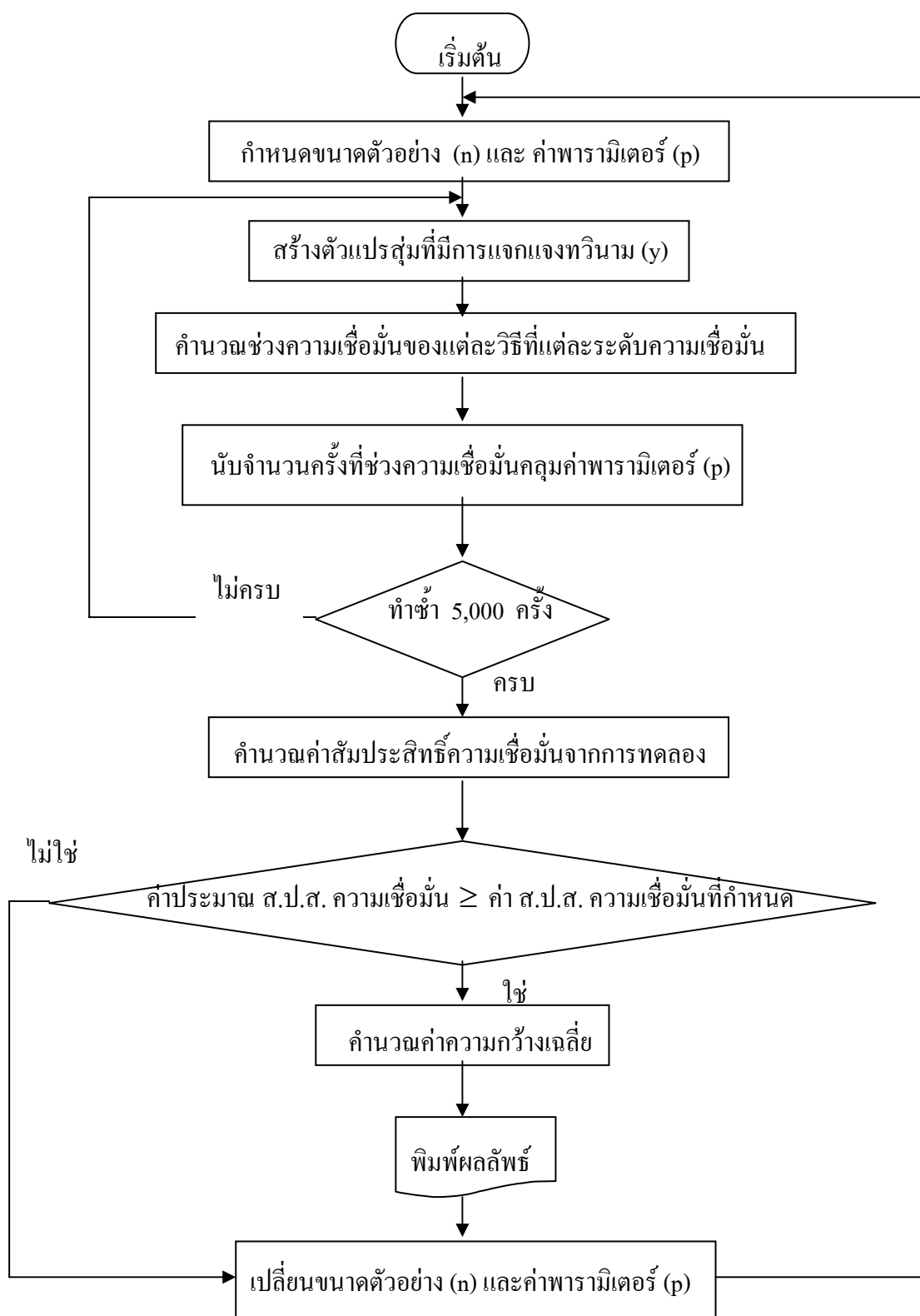
ดังนั้น วิธีการประมาณใดสามารถให้ค่าความกว้างเฉลี่ยของค่าประมาณแบบช่วงต่ำสุด จะถือว่าวิธีประมาณนั้นให้ช่วงประมาณที่เหมาะสมที่สุดสำหรับสถานการณ์นั้นๆ

4. สรุปผลและอภิปรายงานวิจัย

ขั้นตอนการวิจัยทั้งหมดสามารถสรุปเป็นแผนผัง ได้ ดังนี้



ภาพที่ 3 ขั้นตอนการประมาณค่าแบบจุด



ภาพที่ 4 ขั้นตอนการประมาณค่าแบบช่วง

สถานที่และระยะเวลาทำการวิจัย

สถานที่ทำการวิจัย ภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ วิทยาเขต
บางเขน มีระยะเวลาทำการวิจัยตั้งแต่ เดือน พฤศจิกายน พ.ศ. 2549 สิ้นสุด เดือน พฤศจิกายน พ.ศ.
2550

ผลและวิจารณ์

การนำเสนอผลการวิจัยในครั้งนี้ แบ่งการนำเสนอออกเป็น 3 ส่วน คือ

ส่วนที่ 1 เปรียบเทียบการประมาณค่าแบบจุด ซึ่งมี 3 วิธี คือ วิธีภาวนาจะเป็นสูงสุด วิธีเบส์เซียน และวิธีมินิแมกซ์ โดยวิธีที่ให้ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยต่ำสุด ในแต่ละค่า n และ p จะเป็นวิธีที่เหมาะสมที่สุด

ส่วนที่ 2 เปรียบเทียบค่าประมาณสัมประสิทธิ์ความเชื่อมั่น 3 ระดับ คือ 90%, 95% และ 99% โดยวิธีที่ให้ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด จะนำไปหาค่าความกว้างเฉลี่ยต่อไป

ส่วนที่ 3 เปรียบเทียบค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่น 3 ระดับ คือ 90%, 95% และ 99% โดยวิธีที่ให้ค่าความกว้างเฉลี่ยต่ำสุด จะเป็นวิธีที่เหมาะสมที่สุด

1. เปรียบเทียบการประมาณค่าแบบจุด

พิจารณาค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ย สรุปผลได้ดังตารางที่ 3

ตารางที่ 3 ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยของการประมาณทั้ง 3 วิธี

ขนาดตัวอย่าง ค่าพารามิเตอร์		วิธีการประมาณค่าแบบจุด		
(n)	(p)	วิธีภาวน่าจะเป็นสูงสุด	วิธีเบส์เซียน	วิธีมินิเมกซ์
5	0.05	0.0778	0.0752*	0.1392
	0.07	0.0979	0.0725*	0.1336
	0.09	0.1133	0.0768*	0.1268
	0.10	0.1193	0.0868*	0.1238
	0.30	0.1765	0.1181*	0.1183
	0.50	0.1880	0.0854*	0.1299
10	0.05	0.0600	0.0463*	0.1080
	0.07	0.0683	0.0629*	0.1031
	0.09	0.0723*	0.0727	0.0978
	0.10	0.0709*	0.0748	0.0955
	0.30	0.1158	0.0945*	0.0985
	0.50	0.1270	0.0794*	0.0965
30	0.05	0.0316	0.0293*	0.0695
	0.07	0.0344*	0.0356	0.0665
	0.09	0.0401	0.0395*	0.0645
	0.10	0.0399*	0.0413	0.0634
	0.30	0.0712	0.0639	0.0638*
	0.50	0.0784	0.0653	0.0651*
70	0.05	0.0198	0.0194*	0.0483
	0.07	0.0225*	0.0230	0.0465
	0.09	0.0258	0.0254*	0.0455
	0.10	0.0271*	0.0274	0.0451
	0.30	0.0501	0.0471	0.0470*
	0.50	0.0552	0.0508	0.0493*

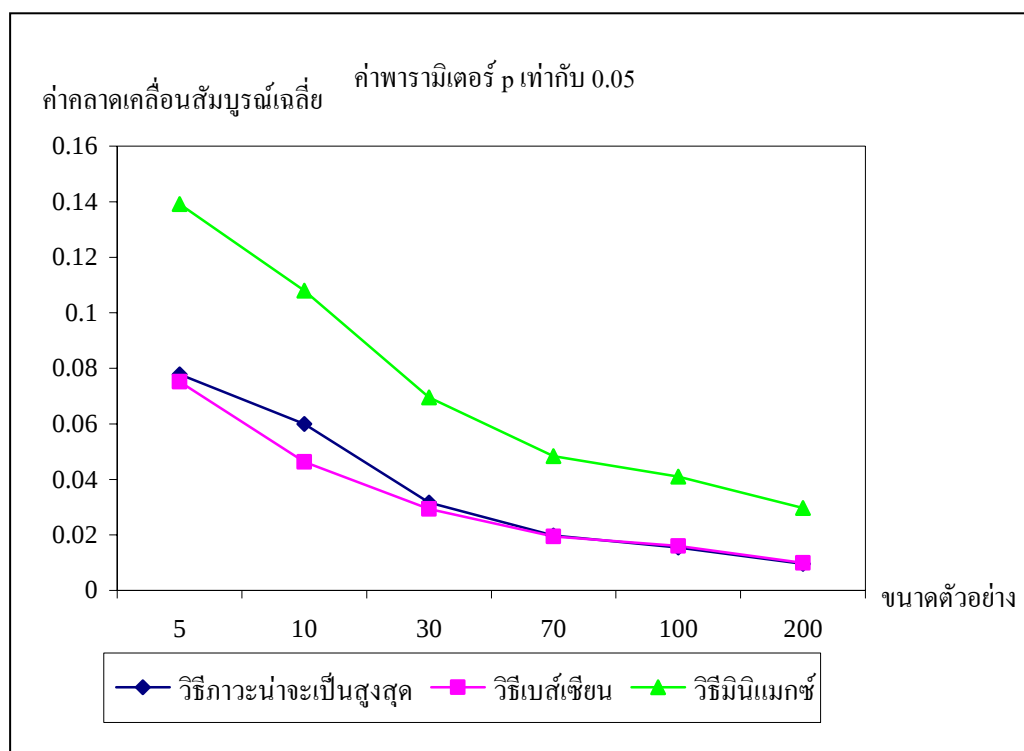
ตารางที่ 3 (ต่อ)

ขนาดตัวอย่าง (n)	ค่าพารามิเตอร์ (p)	วิธีการประมาณค่าแบบจุด		
		วิธีภาวะน่าจะเป็นสูงสุด	วิธีเบย์เซียน	วิธีมินิแมกซ์
100	0.05	0.0154*	0.0160	0.0410
	0.07	0.0183*	0.0187	0.0393
	0.09	0.0208*	0.0212	0.0387
	0.10	0.0220*	0.0223	0.0381
	0.30	0.0403	0.0392*	0.0394
	0.50	0.0452	0.0426	0.0411*
200	0.05	0.0095*	0.0099	0.0297
	0.07	0.0123*	0.0126	0.0287
	0.09	0.0142*	0.0145	0.0297
	0.10	0.0156*	0.0158	0.0279
	0.30	0.0329	0.0327*	0.0333
	0.50	0.0381	0.0369	0.0355*

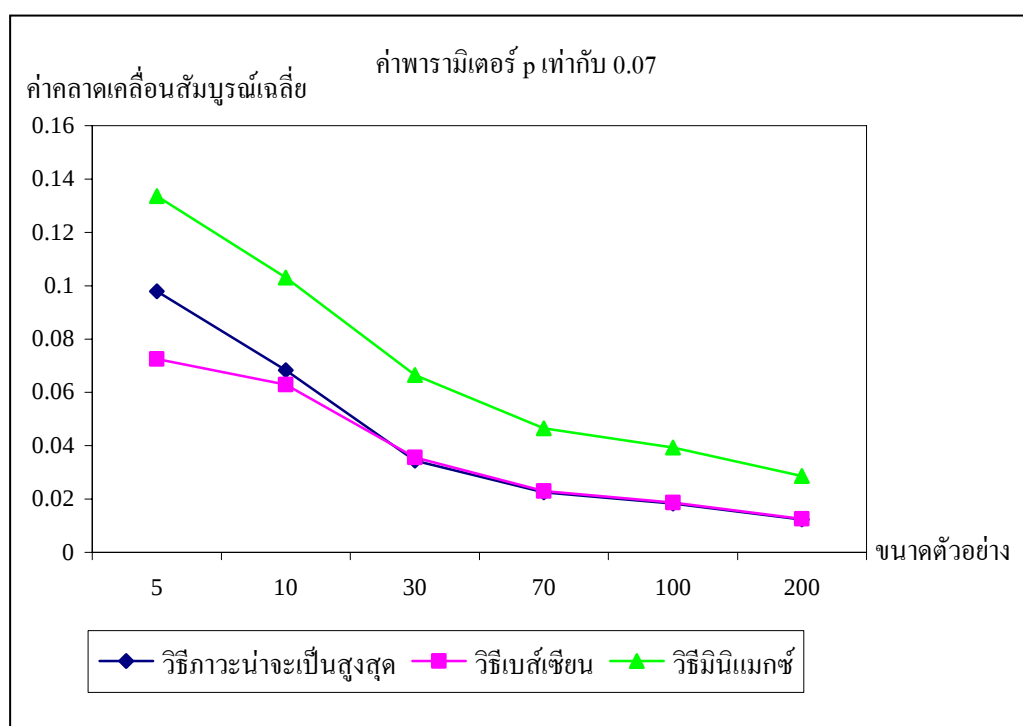
หมายเหตุ * กรณีที่วิธีการประมาณให้ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยต่ำสุด

พิจารณาปัจจัยที่มีผลต่อประสิทธิภาพในการประมาณแบบจุด

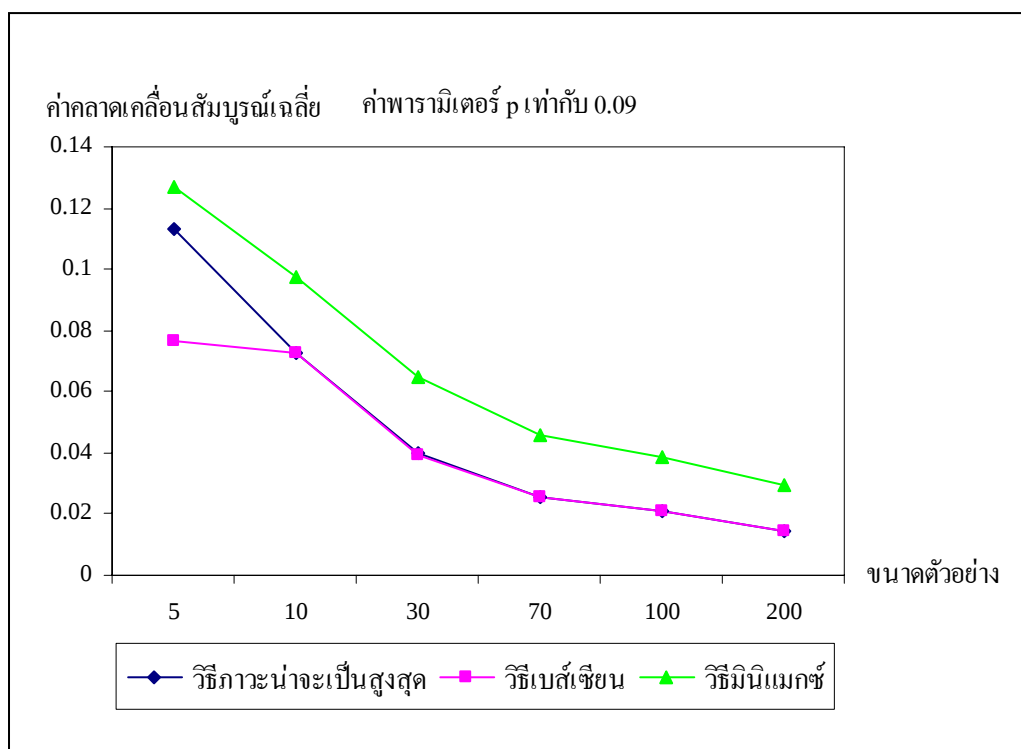
เมื่อตัวอย่างมีขนาดเพิ่มขึ้น พบว่า ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยของวิธีภาวะน่าจะเป็นสูงสุด วิธีเบย์เซียน และ วิธี มินิแมกซ์ มีค่าลดลง ดังภาพที่ 5-10



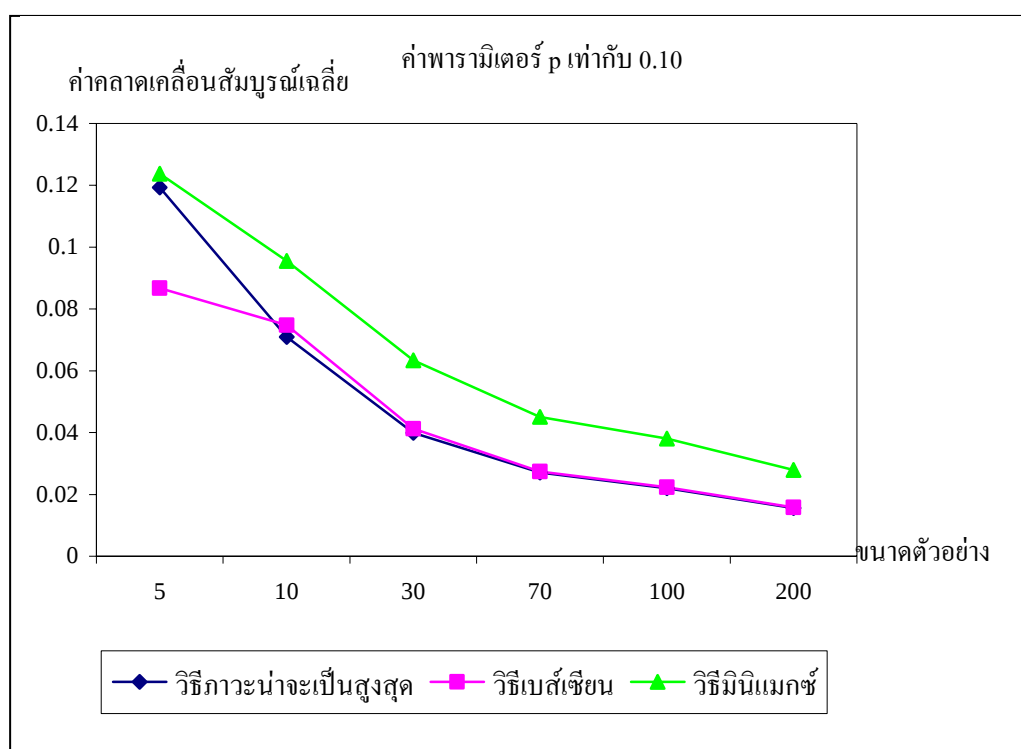
ภาพที่ 5 ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยสำหรับค่าพารามิเตอร์ p เท่ากับ 0.05



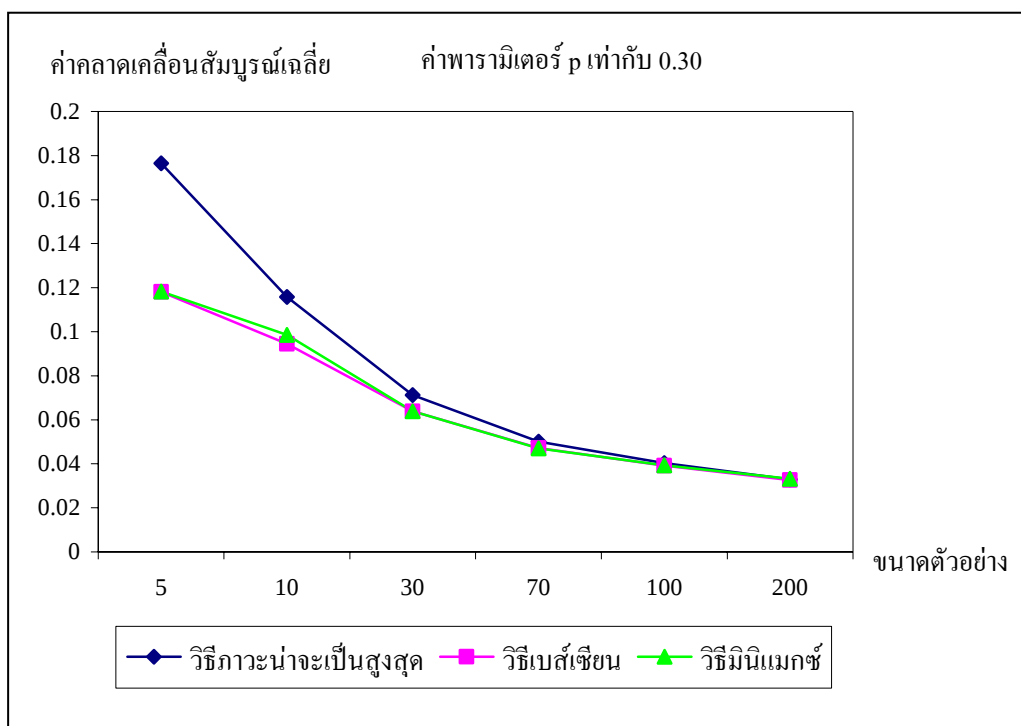
ภาพที่ 6 ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยสำหรับค่าพารามิเตอร์ p เท่ากับ 0.07



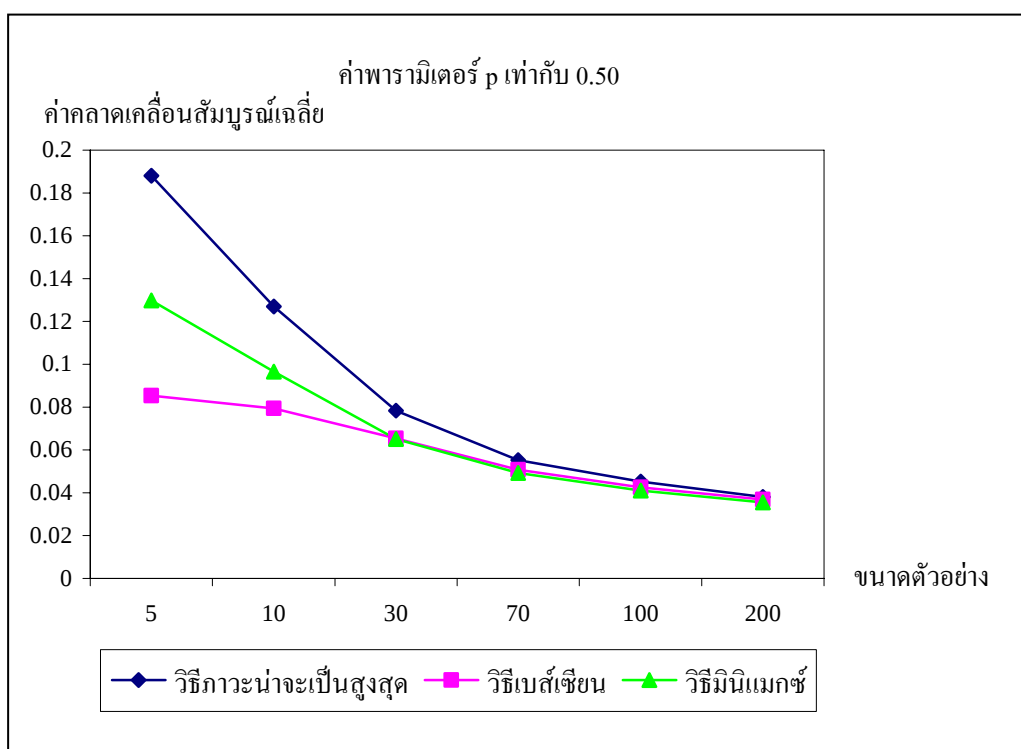
ภาพที่ 7 ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยสำหรับค่าพารามิเตอร์ p เท่ากับ 0.09



ภาพที่ 8 ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยสำหรับค่าพารามิเตอร์ p เท่ากับ 0.10



ภาพที่ 9 ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยสำหรับค่าพารามิเตอร์ p เท่ากับ 0.30



ภาพที่ 10 ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยสำหรับค่าพารามิเตอร์ p เท่ากับ 0.50

จากตารางที่ 3 และภาพ ที่ 5 ถึง 10 สามารถสรุปได้ว่า

สำหรับตัวอย่างขนาดเล็ก (n เท่ากับ 5 และ 10) วิธีเบสเซียน ให้ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยต่ำสุด ทุกค่าพารามิเตอร์ p ที่ทำการศึกษา ยกเว้นกรณี n เท่ากับ 5 และค่าพารามิเตอร์ p เท่ากับ 0.10 วิธีภาวน่าจะเป็นสูงสุดให้ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยต่ำกว่าแต่ก็ใกล้เคียงกับวิธีเบสเซียน

สำหรับตัวอย่างขนาดปานกลางและขนาดใหญ่ (n เท่ากับ 30, 70, 100 และ 200) กรณีค่าพารามิเตอร์ p เล็ก (p เท่ากับ 0.05, 0.07, 0.09 และ 0.10) วิธีเบสเซียนและวิธีภาวน่าจะเป็นสูงสุดให้ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยใกล้เคียงกัน ส่วนวิธีมินิแมกซ์ให้ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยสูงกว่า 2 วิธีแรก กรณีค่าพารามิเตอร์ p ใหญ่ขึ้น (p เท่ากับ 0.30 และ 0.50) การประมาณค่าทั้งสามวิธีให้ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยใกล้เคียงกัน

2. เปรียบเทียบค่าประมาณสัมประสิทธิ์ความเชื่อมั่น

สำหรับการเปรียบเทียบค่าประมาณสัมประสิทธิ์ความเชื่อมั่น ถ้าวิธีการประมาณใดให้ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นจากการทดลองไม่ต่ำกว่า 0.8930, 0.9449 และ 0.9877 ที่ระดับความเชื่อมั่นเท่ากับ 90%, 95% และ 99% ตามลำดับ จะถือว่าวิธีนั้นให้ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด

วิธีการประมาณทั้ง 3 วิธี ให้ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด ในแต่ละขนาดตัวอย่าง ค่าพารามิเตอร์ p และระดับความเชื่อมั่น ที่แตกต่างกันดังตารางต่อไปนี้

ตารางที่ 4 ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นของช่วงความเชื่อมั่นของพารามิเตอร์ p ที่
ระดับความเชื่อมั่น 90%

ขนาดตัวอย่าง ค่าพารามิเตอร์		วิธี		
(n)	(p)	วิธี ปกติ	วิธี โลจิส	วิธีสคอร์แบบปรับความต่อเนื่อง
5	0.05	0.2302	0.7694	0.9748*
	0.07	0.3002	0.9540*	0.9540*
	0.09	0.3628	0.9274*	0.9930*
	0.10	0.3950	0.9164*	0.9920*
	0.30	0.7918	0.9670*	0.9670*
	0.50	0.6142	0.9458*	0.9458*
10	0.05	0.3872	0.8900	0.9904*
	0.07	0.5074	0.9200*	0.9714*
	0.09	0.6028	0.9388*	0.9388*
	0.10	0.6308	0.9236*	0.9852*
	0.30	0.7932	0.9208*	0.9208*
	0.50	0.8882	0.8988*	0.9776*
30	0.05	0.7872	0.9498*	0.9498*
	0.07	0.8846	0.9530*	0.9530*
	0.09	0.7832	0.9172*	0.9586*
	0.10	0.8216	0.8930*	0.9798*
	0.30	0.8578	0.8920	0.9126*
	0.50	0.8722	0.8722	0.9560*
70	0.05	0.8696	0.9250*	0.9250*
	0.07	0.8746	0.9000*	0.9418*
	0.09	0.8940*	0.8930*	0.9388*
	0.10	0.9296*	0.8930*	0.9330*
	0.30	0.8754	0.8698	0.8964*
	0.50	0.8186	0.8186	0.8946*

ตารางที่ 4 (ต่อ)

ขนาดตัวอย่าง ค่าพารามิเตอร์		วิธี		
(n)	(p)	วิธี ปกติ	วิธี โลจิส	วิธีสคอร์แบบปรับความต่อเนื่อง
100	0.05	0.9034*	0.9424*	0.9878*
	0.07	0.9326*	0.9062*	0.9518*
	0.09	0.8946*	0.9322*	0.9628*
	0.10	0.8936*	0.9144*	0.9450*
	0.30	0.7806	0.8150	0.8978*
	0.50	0.8134	0.8134	0.8930*
200	0.05	0.9170*	0.9676*	0.9808*
	0.07	0.9158*	0.9500*	0.9684*
	0.09	0.9156*	0.9580*	0.9690*
	0.10	0.9030*	0.9610*	0.9664*
	0.30	0.8622	0.8412	0.9002*
	0.50	0.8292	0.8292	0.9194*

หมายเหตุ * กรณีที่วิธีการประมาณให้ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่าค่า
สัมประสิทธิ์ความเชื่อมั่นที่กำหนด

ตารางที่ 5 ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นของช่วงความเชื่อมั่นของพารามิเตอร์ p ที่ระดับความเชื่อมั่น 95%

ขนาดตัวอย่าง ค่าพารามิเตอร์		วิธี		
(n)	(p)	วิธี ปกติ	วิธี โลจิส	วิธีสคอร์แบบปรับความต่อเนื่อง
5	0.05	0.2264	0.7722	0.9754*
	0.07	0.3046	0.9550*	0.9550*
	0.09	0.3650	0.9447*	0.9932*
	0.10	0.3978	0.9600*	0.9924*
	0.30	0.7986	0.9656*	0.9988*
	0.50	0.9384	0.9456*	1.0000*
10	0.05	0.3986	0.9124	0.9896*
	0.07	0.5076	0.9674*	0.9674*
	0.09	0.6002	0.9460*	0.9924*
	0.10	0.6458	0.9500*	0.9868*
	0.30	0.8268	0.9449*	0.9890*
	0.50	0.8878	0.9858*	0.9858*
30	0.05	0.7978	0.9494*	0.9908*
	0.07	0.8946	0.9562*	0.9898*
	0.09	0.9448	0.9550*	0.9886*
	0.10	0.8334	0.9530*	0.9802*
	0.30	0.9300	0.9282	0.9610*
	0.50	0.8864	0.9419	0.9512*
70	0.05	0.8802	0.9500*	0.9868*
	0.07	0.8934	0.9449*	0.9748*
	0.09	0.9532*	0.9450*	0.9716*
	0.10	0.9449*	0.9460*	0.9682*
	0.30	0.9136	0.9152	0.9486*
	0.50	0.8788	0.9280	0.9808*

ตารางที่ 5 (ต่อ)

ขนาดตัวอย่าง ค่าพารามิเตอร์		วิธี		
(n)	(p)	วิธี ปกติ	วิธี โลจิส	วิธีสคอร์แบบปรับความต่อเนื่อง
100	0.05	0.9500*	0.9592*	0.9862*
	0.07	0.9470*	0.9638*	0.9840*
	0.09	0.9504*	0.9580*	0.9858*
	0.10	0.9449*	0.9546*	0.9728*
	0.30	0.8918	0.8868	0.9449*
	0.50	0.8716	0.8716	0.9526*
200	0.05	0.9666*	0.9890*	0.9948*
	0.07	0.9592*	0.9848*	0.9902*
	0.09	0.9558*	0.9722*	0.9818*
	0.10	0.9478*	0.9854*	0.9884*
	0.30	0.9200	0.9362	0.9498*
	0.50	0.9094	0.9094	0.9450*

หมายเหตุ * กรณีที่วิธีการประมาณให้ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่าค่า
สัมประสิทธิ์ความเชื่อมั่นที่กำหนด

ตารางที่ 6 แสดงค่าประมาณสัมประสิทธิ์ความเชื่อมั่นของช่วงความเชื่อมั่นของพารามิเตอร์ p ที่ระดับความเชื่อมั่น 99%

ขนาดตัวอย่าง ค่าพารามิเตอร์		วิธี		
(n)	(p)	วิธี ปกติ	วิธี โลจิส	วิธีสคอร์แบบปรับความต่อเนื่อง
5	0.05	0.2322	0.9784	0.9986*
	0.07	0.2980	0.9900*	0.9978*
	0.09	0.3658	0.9942*	0.9942*
	0.10	0.4108	0.9922*	0.9922*
	0.30	0.7890	0.9978*	0.9978*
	0.50	0.9378	1.0000*	1.0000*
10	0.05	0.3984	0.9864	0.9888*
	0.07	0.5104	0.9900*	0.9974*
	0.09	0.6094	0.9902*	0.9902*
	0.10	0.6464	0.9884*	0.9884*
	0.30	0.9626	0.9982*	0.9982*
	0.50	0.9816	1.0000*	1.0000*
30	0.05	0.7966	0.9902*	0.9994*
	0.07	0.8996	0.9880*	0.9986*
	0.09	0.9456	0.9900*	0.9990*
	0.10	0.9608	0.9964*	0.9964*
	0.30	0.9586	0.9800	0.9877*
	0.50	0.9846	0.9850	0.9985*
70	0.05	0.9795	0.9877*	0.9990*
	0.07	0.9632	0.9966*	0.9966*
	0.09	0.9900*	0.9912*	0.9954*
	0.10	0.9900*	0.9922*	0.9922*
	0.30	0.9700	0.9770	0.9955*
	0.50	0.9800	0.9826	0.9936*

ตารางที่ 6 (ต่อ)

ขนาดตัวอย่าง ค่าพารามิเตอร์		วิธี		
(n)	(p)	วิธี ปกติ	วิธี โลจิส	วิธีสคอร์แบบปรับความต่อเนื่อง
100	0.05	0.9877*	0.9990*	1.0000*
	0.07	0.9877*	0.9964*	0.9998*
	0.09	0.9877*	0.9962*	0.9962*
	0.10	0.9878*	0.9948*	0.9980*
	0.30	0.9470	0.9680	0.9932*
	0.50	0.9856	0.9872	0.9930*
200	0.05	0.9914*	0.9994*	1.0000*
	0.07	0.9877*	0.9990*	0.9996*
	0.09	0.9877*	0.9988*	0.9988*
	0.10	0.9906*	0.9990*	0.9994*
	0.30	0.9848	0.9764	0.9930*
	0.50	0.9654	0.9825	0.9970*

หมายเหตุ * กรณีที่วิธีการประมาณให้ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่าค่า
สัมประสิทธิ์ความเชื่อมั่นที่กำหนด

จากตารางที่ 4, 5 และ 6 เมื่อพิจารณากรณีที่ให้ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นต่ำกว่า
ค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด ได้แก่

วิธีปกติ กรณีตัวอย่างขนาดเล็กและปานกลางทุกค่าพารามิเตอร์ p ที่ทำการศึกษา ยกเว้น n
เท่ากับ 70 ค่าพารามิเตอร์ p เท่ากับ 0.09 ถึง 0.10 และตัวอย่างขนาดใหญ่ ค่าพารามิเตอร์ p เท่ากับ
0.30 ถึง 0.50

วิธีโลจิส กรณีตัวอย่างขนาดเล็ก ค่าพารามิเตอร์ p เท่ากับ 0.05 ตัวอย่างขนาดปานกลางและ
ใหญ่ ค่าพารามิเตอร์ p เท่ากับ 0.30 ถึง 0.50

พิจารณากรณีที่ให้ค่าประมาณสัมประสิทธิ์ความเชื่อมั่น ไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด ได้แก่

วิธีปกติ กรณีตัวอย่างขนาดปานกลาง (เฉพาะ n เท่ากับ 70) ค่าพารามิเตอร์ p เท่ากับ 0.09 ถึง 0.10 และ ตัวอย่างขนาดใหญ่ ค่าพารามิเตอร์ p เท่ากับ 0.05 ถึง 0.10

วิธีโลจิส กรณีตัวอย่างขนาดเล็ก ค่าพารามิเตอร์ p เท่ากับ 0.07 ถึง 0.09 ตัวอย่างขนาดปานกลางและใหญ่ ค่าพารามิเตอร์ p เท่ากับ 0.05 ถึง 0.10

วิธีสคอร์แบบปรับความต่อเนื่อง ทุกขนาดตัวอย่างและทุกค่าพารามิเตอร์ p ที่ทำการศึกษา

3. การเปรียบเทียบค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่น

ในกรณีที่วิธีการประมาณใดให้ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นจากการทดลองไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนดแล้ว จะนำวิธีการเหล่านั้นมาเปรียบเทียบหาวิธีที่เหมาะสมที่สุด โดยศึกษาค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นที่ได้จากการทดลองนั้นว่าวิธีการใดให้ช่วงความกว้างเฉลี่ยต่ำสุดจะเป็นวิธีที่เหมาะสมที่สุด

การประมาณทั้ง 3 วิธี ให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นต่ำสุด ในแต่ละขนาดตัวอย่าง ค่าพารามิเตอร์ p และระดับความเชื่อมั่นที่แตกต่างกัน ดังตารางต่อไปนี้

ตารางที่ 7 ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นจากวิธีการประมาณทั้ง 3 วิธี ที่ระดับความเชื่อมั่น 90%

ขนาดตัวอย่าง ค่าพารามิเตอร์		วิธี		
(n)	(p)	วิธี ปกติ	วิธี โลจิส	วิธีสคอร์แบบปรับความต่อเนื่อง
5	0.05	-	-	0.4634*
	0.07	-	0.3990*	0.4650
	0.09	-	0.3952*	0.5080
	0.10	-	0.3945*	0.5149
	0.30	-	0.5110*	0.6107
	0.50	-	0.5364*	0.6451
10	0.05	-	-	0.3212*
	0.07	-	0.2514*	0.3316
	0.09	-	0.2786*	0.3345
	0.10	-	0.2788*	0.3656
	0.30	-	0.3850*	0.4519
	0.50	-	0.4045*	0.5100
30	0.05	-	0.1313*	0.1561
	0.07	-	0.1485*	0.1738
	0.09	-	0.1487*	0.1907
	0.10	-	0.1626*	0.2027
	0.30	-	-	0.2656*
	0.50	-	-	0.3002*
70	0.05	-	0.0824*	0.0935
	0.07	-	0.0914*	0.1073
	0.09	0.1017	0.1014*	0.1181
	0.10	0.1097	0.1052*	0.1231
	0.30	-	-	0.1589*
	0.50	-	-	0.1839*

ตารางที่ 7 (ต่อ)

ขนาดตัวอย่าง ค่าพารามิเตอร์		วิธี		
(n)	(p)	วิธี ปกติ	วิธี โลจิส	วิธีสคอร์แบบปรับความต่อเนื่อง
100	0.05	0.0657*	0.0700	0.0818
	0.07	0.0768*	0.0786	0.0900
	0.09	0.0850*	0.0882	0.1004
	0.10	0.0893*	0.0905	0.1026
	0.30	-	-	0.1424*
	0.50	-	-	0.1511*
200	0.05	0.0471*	0.0501	0.0553
	0.07	0.0550*	0.0570	0.0628
	0.09	0.0614*	0.0642	0.0695
	0.10	0.0637*	0.0675	0.0725
	0.30	-	-	0.1011*
	0.50	-	-	0.1106*

หมายเหตุ - กรณีที่ไม่ได้หาค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่น เนื่องจากวิธีการประมาณนั้น
ให้ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด

* กรณีที่วิธีการประมาณให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นต่ำสุด

ตารางที่ 8 ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นจากวิธีการประมาณทั้ง 3 วิธี ที่ระดับความเชื่อมั่น 95%

ขนาดตัวอย่าง ค่าพารามิเตอร์		วิธี		
(n)	(p)	วิธี ปกติ	วิธี โลจิส	วิธีสคอร์แบบปรับความต่อเนื่อง
5	0.05	-	-	0.5406*
	0.07	-	0.4881*	0.5404
	0.09	-	0.4903*	0.5819
	0.10	-	0.4972*	0.5885
	0.30	-	0.5926*	0.6913
	0.50	-	0.6126*	0.7226
10	0.05	-	-	0.3831*
	0.07	-	0.3402*	0.3890
	0.09	-	0.3473*	0.4209
	0.10	-	0.3515*	0.4255
	0.30	-	0.4593*	0.5369
	0.50	-	0.5086*	0.5701
30	0.05	-	0.1645*	0.1983
	0.07	-	0.1849*	0.2169
	0.09	-	0.2007*	0.2331
	0.10	-	0.2053*	0.2378
	0.30	-	-	0.3218*
	0.50	-	-	0.3439*
70	0.05	-	0.1031*	0.1184
	0.07	-	0.1163*	0.1322
	0.09	0.1293	0.1274*	0.1448
	0.10	0.1347	0.1321*	0.1496
	0.30	-	-	0.2103*
	0.50	-	-	0.1176*

ตารางที่ 8 (ต่อ)

ขนาดตัวอย่าง ค่าพารามิเตอร์		วิธี		
(n)	(p)	วิธี ปกติ	วิธี โลจิส	วิธีสคอร์แบบปรับความต่อเนื่อง
100	0.05	0.0822*	0.0862	0.0969
	0.07	0.0952*	0.0993	0.1097
	0.09	0.1070*	0.1098	0.1209
	0.10	0.1115*	0.1137	0.1243
	0.30	-	-	0.1749*
	0.50	-	-	0.1915*
200	0.05	0.0585*	0.0617	0.0663
	0.07	0.0681*	0.0711	0.0758
	0.09	0.0760*	0.0779	0.0831
	0.10	0.0792*	0.0827	0.0874
	0.30	-	-	0.1241*
	0.50	-	-	0.1340*

หมายเหตุ - กรณีที่ไม่ได้หาค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่น เนื่องจากวิธีการประมาณนั้น
ให้ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด

* กรณีที่วิธีการประมาณให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นต่ำสุด

ตารางที่ 9 ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นจากวิธีการประมาณทั้ง 3 วิธี ที่ระดับความเชื่อมั่น 99%

ขนาดตัวอย่าง ค่าพารามิเตอร์		วิธี		
(n)	(p)	วิธี ปกติ	วิธี โลจิส	วิธีสคอร์แบบปรับความต่อเนื่อง
5	0.05	-	-	0.6713*
	0.07	-	0.6664*	0.6805
	0.09	-	0.6860*	0.6879
	0.10	-	0.6893*	0.6931
	0.30	-	0.7456*	0.7714
	0.50	-	0.7646*	0.7983
10	0.05	-	-	0.4899*
	0.07	-	0.4959*	0.5118
	0.09	-	0.5107*	0.5232
	0.10	-	0.5154*	0.5284
	0.30	-	0.6091*	0.6362
	0.50	-	0.6363*	0.6694
30	0.05	-	0.2551*	0.2677
	0.07	-	0.2742*	0.2875
	0.09	-	0.2921*	0.3057
	0.10	-	0.3018*	0.3123
	0.30	-	-	0.4115*
	0.50	-	-	0.4454*
70	0.05	-	0.1525*	0.1603
	0.07	-	0.1709*	0.1769
	0.09	0.1849	0.1744*	0.1920
	0.10	0.1912	0.1829*	0.1982
	0.30	-	-	0.3036*
	0.50	-	-	0.3028*

ตารางที่ 9 (ต่อ)

ขนาดตัวอย่าง ค่าพารามิเตอร์		วิธี		
(n)	(p)	วิธี ปกติ	วิธี โลจิส	วิธีสคอร์แบบปรับความต่อเนื่อง
100	0.05	0.1098*	0.1256	0.1301
	0.07	0.1281*	0.1403	0.1458
	0.09	0.1439*	0.1535	0.1590
	0.10	0.1115*	0.1588	0.1651
	0.30	-	-	0.2361*
	0.50	-	-	0.2556*
200	0.05	0.0784*	0.0847	0.0877
	0.07	0.0911*	0.0964	0.0997
	0.09	0.1023*	0.1067	0.1101
	0.10	0.1079*	0.1113	0.1148
	0.30	-	-	0.1679*
	0.50	-	-	0.1829*

หมายเหตุ - กรณีที่ไม่ได้หาค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่น เนื่องจากวิธีการประมาณนั้น
ให้ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด
* กรณีที่วิธีการประมาณให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นต่ำสุด

จากตารางที่ 7,8 และ 9 เมื่อพิจารณาค่าประมาณสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่าค่า
สัมประสิทธิ์ความเชื่อมั่นที่กำหนด พบว่ากรณีที่แต่ละวิธีไม่สามารถหาค่าความกว้างเฉลี่ยได้ คือ

วิธีปกติ กรณีตัวอย่างขนาดเล็กและปานกลางทุกค่าพารามิเตอร์ p ที่ทำการศึกษา
ยกเว้นขนาดตัวอย่าง n เท่ากับ 70 ค่าพารามิเตอร์ p เท่ากับ 0.09 ถึง 0.10 และตัวอย่างขนาดใหญ่
ค่าพารามิเตอร์ p เท่ากับ 0.30 ถึง 0.50

วิธีโลจิส กรณีตัวอย่างขนาดเล็ก ค่าพารามิเตอร์ p เท่ากับ 0.05 และตัวอย่างขนาดปานกลาง
และใหญ่ ค่าพารามิเตอร์ p เท่ากับ 0.30 ถึง 0.50

กรณีแต่ละวิธีให้ค่าความกว้างเฉลี่ยต่ำสุด คือ

วิธีปกติ กรณีตัวอย่างขนาดใหญ่ ค่าพารามิเตอร์ p เท่ากับ 0.05 ถึง 0.10

วิธีโลจิส กรณีตัวอย่างขนาดเล็ก ค่าพารามิเตอร์ p เท่ากับ 0.07 ถึง 0.50 และตัวอย่างขนาดปานกลาง ค่าพารามิเตอร์ p เท่ากับ 0.05 ถึง 0.10

วิธีสคอร์แบบปรับความต่อเนื่อง กรณีตัวอย่างขนาดเล็ก ค่าพารามิเตอร์ p เท่ากับ 0.05 และตัวอย่างขนาดปานกลางและใหญ่ ค่าพารามิเตอร์ p เท่ากับ 0.30 ถึง 0.50

สรุปและข้อเสนอแนะ

สรุป

การวิจัยนี้เป็นการวิจัยเชิงทดลอง เพื่อศึกษาและเปรียบเทียบวิธีการประมาณค่าแบบจุดและการประมาณค่าแบบช่วง ผลการวิจัยสรุปได้ดังนี้

1. เปรียบเทียบการประมาณแบบจุด

จากการพิจารณาการประมาณค่าแบบจุด 3 วิธี คือ วิธีภาวน่าจะเป็นสูงสุด วิธีเบส์เซียน และ วิธีมินิแมกซ์ โดยใช้ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยต่ำสุดเป็นเกณฑ์ในการเลือกวิธีประมาณ สามารถสรุปได้ว่า แต่ละกรณีที่ทำการศึกษา ควรเลือกใช้วิธีการประมาณค่าที่เหมาะสม ดังตารางที่ 10

ตารางที่ 10 วิธีที่เหมาะสมในการประมาณค่าแบบจุด จำแนกตามขนาดตัวอย่าง (n) และค่าพารามิเตอร์ (p)

ขนาดตัวอย่าง (n)	พารามิเตอร์ (p)					
	0.05	0.07	0.09	0.10	0.30	0.50
5	B	B	B	B	B	B
10	B	B	B	B	B	B
30	M	M	M	M	M	M
70	M	M	M	M	M	M
100	M	M	M	M	M	M
200	M	M	M	M	M	M

หมายเหตุ M แทน วิธีภาวน่าจะเป็นสูงสุด, B แทน วิธีเบส์เซียน

ข้อสังเกตในการประมาณค่าแบบจุด

ขนาดตัวอย่าง

เมื่อพิจารณาขนาดตัวอย่าง พบว่า การประมาณค่าของทั้งสามวิธีให้ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้น

ค่าพารามิเตอร์

เมื่อค่าพารามิเตอร์ p มีค่าน้อย วิธีเบส์เซียนและวิธีภาวน่าจะเป็นสูงสุดให้ผลที่ไม่แตกต่างกัน ส่วนวิธีมินิเมกซ์ให้ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยสูงกว่าวิธีอื่นอย่างชัดเจน แต่เมื่อค่าพารามิเตอร์ p เพิ่มขึ้น วิธีการทั้งสามวิธีให้ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยที่ไม่แตกต่างกัน

2. การประมาณค่าแบบช่วง

จากการพิจารณาการประมาณค่าแบบช่วง 3 วิธี คือ วิธีปกติ วิธีโลจิส และ วิธีสคอร์แบบปรับความต่อเนื่อง พบว่า สำหรับทุกระดับความเชื่อมั่น 90%, 95% และ 99% กรณีที่แต่ละวิธีให้ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด แสดงได้ดังตารางที่ 11

ตารางที่ 11 กรณีที่แต่ละวิธีให้ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนดจำแนกตามขนาดตัวอย่าง (n) และ ค่าพารามิเตอร์ (p)

ขนาดตัวอย่าง (n)	พารามิเตอร์ (p)					
	0.05	0.07	0.09	0.10	0.30	0.50
5	S	L,S	L,S	L,S	L,S	L,S
10	S	L,S	L,S	L,S	L,S	L,S
30	L,S	L,S	L,S	L,S	S	S
70	L,S	L,S	N,L,S	N,L,S	S	S
100	N,L,S	N,L,S	N,L,S	N,L,S	S	S
200	N,L,S	N,L,S	N,L,S	N,L,S	S	S

หมายเหตุ N แทน วิธีปกติ, L แทน วิธีโลจิส, S แทน วิธีสคอร์แบบปรับความต่อเนื่อง

เมื่อนำทุกกรณีที่แต่ละวิธีให้ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนดมาทำการเปรียบเทียบค่าความกว้างเฉลี่ยต่ำสุด สรุปได้ว่าแต่ละกรณีที่ทำการศึกษา ควรเลือกใช้วิธีการประมาณค่าแบบช่วง ดังตารางที่ 12

ตารางที่ 12 วิธีที่เหมาะสมในการประมาณค่าแบบช่วง จำแนกตามขนาดตัวอย่าง (n) และ ค่าพารามิเตอร์ (p)

ขนาดตัวอย่าง (n)	พารามิเตอร์ (p)					
	0.05	0.07	0.09	0.10	0.30	0.50
5	S	L	L	L	L	L
10	S	L	L	L	L	L
30	L	L	L	L	S	S
70	L	L	L	L	S	S
100	N	N	N	N	S	S
200	N	N	N	N	S	S

หมายเหตุ N แทน วิธีปกติ, L แทน วิธีโลจิส, S แทน วิธีสคอร์แบบปรับความต่อเนื่อง

ข้อเสนอแนะ

จากการวิจัยเพื่อศึกษาวิธีการประมาณค่าพารามิเตอร์ สำหรับการแจกแจงทวินาม มีข้อเสนอแนะจากการวิจัยและข้อเสนอแนะสำหรับการวิจัยครั้งต่อไป ดังนี้

ข้อเสนอแนะจากการวิจัย

จากการวิจัยพบว่า

1. การประมาณค่าแบบจุด

วิธีเบส์เซียน เหมาะสำหรับตัวอย่างขนาดเล็ก เนื่องจากให้ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยต่ำสุดทุกค่าพารามิเตอร์ p ที่ทำการศึกษา วิธีภาวน่าจะเป็นสูงสุด เหมาะสำหรับตัวอย่างขนาดปานกลางและขนาดใหญ่ เนื่องจากเมื่อค่าพารามิเตอร์ p มีค่าน้อย วิธีเบส์เซียนและวิธีภาวน่าจะเป็นสูงสุดให้ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยไม่แตกต่างกัน และเมื่อค่าพารามิเตอร์ p เพิ่มขึ้น ทั้งสามวิธีให้ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยที่ไม่แตกต่างกัน และวิธีมินิแมกซ์เป็นวิธีที่ไม่ควรเลือกใช้ เพราะให้ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยสูงกว่า 2 วิธีแรกและมีการคำนวณที่ซับซ้อนกว่า

2. การประมาณค่าแบบช่วง

วิธีปกติ เหมาะสำหรับตัวอย่างมีขนาดใหญ่ เมื่อค่าพารามิเตอร์ p เท่ากับ 0.05 ถึง 0.10 วิธีโลจิส เหมาะสำหรับตัวอย่างมีขนาดเล็ก เมื่อค่าพารามิเตอร์ p เท่ากับ 0.07 ถึง 0.50 และเหมาะสำหรับตัวอย่างมีขนาดปานกลาง เมื่อค่าพารามิเตอร์ p เท่ากับ 0.05 ถึง 0.10 วิธีสคอร์แบบปรับความต่อเนื่อง เหมาะสำหรับตัวอย่างมีขนาดเล็ก เมื่อค่าพารามิเตอร์ p เท่ากับ 0.05 และเหมาะสำหรับตัวอย่างมีขนาดปานกลางและใหญ่เมื่อค่าพารามิเตอร์ p เท่ากับ 0.30 ถึง 0.50

ข้อเสนอแนะสำหรับการวิจัยครั้งต่อไป

1. ศึกษาวิธีอื่น นอกเหนือจาก 3 วิธีดังกล่าวในการประมาณค่าแบบจุด อาทิเช่น วิธีไคกำลังสองต่ำสุด (Minimum chisquares method) วิธีการถ่วงน้ำหนัก (Weighted method) และวิธีเบย์เซียน (Bayesian method) โดยกำหนดความน่าจะเป็นก่อน (Prior probability) เป็นการแจกแจงแบบอื่นๆ
2. ศึกษาวิธีอื่น นอกเหนือจาก 3 วิธีดังกล่าวในการประมาณค่าแบบช่วง อาทิเช่น วิธีอัตราส่วนภาวน่าจะเป็น (Likelihood ratio interval) วิธีหาค่าต่ำสุด (Minimization) ของความกว้างของช่วงความเชื่อมั่น และวิธี Jeffreys interval
3. ศึกษาการประมาณค่าพารามิเตอร์สำหรับกรณีที่มีการแจกแจงแบบอื่นๆ เช่น การแจกแจงเรขาคณิต การแจกแจงทวินามนิเสธ การแจกแจงแบบปัวส์ซอง เป็นต้น

เอกสารและสิ่งอ้างอิง

- จอมขวัญ พวงดอก. 2548. การเปรียบเทียบวิธีการประมาณช่วงความเชื่อมั่นของสัดส่วน
ประชากร. วิทยานิพนธ์ปริญญาโท, สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ.
- พลาวัช ศรีสมบูรณ์. 2534. การประมาณค่าแบบช่วงสำหรับพารามิเตอร์ของการแจกแจงทวินาม
ภายใต้เงื่อนไขขนาดเล็ก. วิทยานิพนธ์ปริญญาโท, จุฬาลงกรณ์มหาวิทยาลัย.
- สงขลา สำเภาเงิน. 2542. การเปรียบเทียบวิธีการประมาณค่าแบบช่วงสำหรับค่าสัดส่วนของการ
แจกแจงแบบทวินาม. วิทยานิพนธ์ปริญญาโท, จุฬาลงกรณ์มหาวิทยาลัย.
- สายชล สันสมบูรณ์. 2544. สถิติคณิตศาสตร์ 1. สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหาร
ลาดกระบัง, กรุงเทพฯ.
- สารินิย์ คงกัน. 2546. การเปรียบเทียบวิธีประมาณค่าแบบช่วงสำหรับพารามิเตอร์ในการแจกแจง
ทวินาม. วิทยานิพนธ์ปริญญาโท, มหาวิทยาลัยเกษตรศาสตร์.
- อโนทัย ตรีวานิช. 2539. ทฤษฎีการอนุมานเชิงสถิติ. โรงพิมพ์คลังนานาวิทยา, ขอนแก่น.
- Aczel, A.D. 1993. **Complete Business Statistics**. Homewood, Irwin.
- Agresti, A. and B. Caffo. 2000. Simple and effective confidence intervals for proportion and
differences of proportion result from adding two successes and two failures. **The
American Statistician**. 54: 280-288.
- _____. and B.A. Coull. 1998. Approximate is better than “exact for interval estimation of
binomial proportions”. **The American Statistician**. 52: 19-126.

Anderson, A.D., D.J. Sweeney and T.A. William. 1994. **Introduction to Statistics: Concepts and Applications** . St. Paul MN., West Publishing.

Brown, L.D., T.Tony Cai and A. Dasgupta . 2001. Interval estimation for a binomial proportion. **Statistical Science**. 16: 101-133.

_____. 2002. Confidence interval for a binomial proportion and asymptotic expansion. **The Annals of Statistics**. 30: 160-201.

_____. 2003. Interval estimation in exponential families. **Statistica Sinica** .13: 19-49.

Casella,G. and Berger,R.L. 2002. **Statistical Inference**. John Wiley and Sons,Inc., New York.

Chen,H. 1990. Accuracy of approximate interval for binomial parameter. **Journal of the American Statistical Association**. 85: 514-518.

Cox,D.R. and E.J. Snell. 1989. **Analysis of Binary Data**. Chapman and Hall, London.

Creighton, J.H.C. 1994. **A First Course in Probability Models and Statistical Inference**. Springer-Verlag., New York

Fleiss, J.L. 2003. **Statistical Inference for a Singer Proportion**. John Wiley and Sons,Inc., New York.

Freund, J.E. 1992. **Mathematical Statistics**. Prentice-Hall, London.

Gart, J.J. and J.R. Zweifel 1967. O n the bias of various estimators of the logit and its variance with application to quantal bioassay. **Biometrika**. 54: 181-185.

Ghosh,B.K. 1979. Comparison of some approximate confidence intervals for binomial parameter. **Journal of the American Statistical Association.** 74:894-900.

Larson, H.J. 1995. **Introduction to Probability.** Reading, MA., Addison Wesley.

Mood, A.M., F.A. Graybill and D.C. Boes. 1974. **Introduction to the Theory of Statistics.** McGrawHill, Tokyo.

Rahatgi,V.K. and A.K. Saleh and M.D.Ehsanes. 2001. **An Introduction to Probability Theory and Mathematical Statistics.** John Wiley and Sons,Inc., New York.

Roger, V.H. and E.A. Tanis. 2006. **Probability and Statistical Inference.** Pearson Education,Inc., New Jersey.

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

William M.B. 2004. **Introduction to Bayesian Statistics.** John Wiley and Sons,Inc., New York.

ภาคผนวก

ภาคผนวก ก

โปรแกรมที่ใช้จำลองชุดข้อมูลในการวิจัย

โปรแกรมที่ใช้จำลองชุดข้อมูลในการวิจัย

Option Explicit

Dim Error1, Error2, Error3 As Double ' เก็บค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยในการ
ประมาณค่าแบบจุด

Dim Sum1, Sum2, Sum3 As Double ' เก็บค่าประมาณสัมประสิทธิ์ความเชื่อมั่นจาก
การทดลอง

Dim SumT1, SumT2, SumT3 As Double ' เก็บค่าความกว้างเฉลี่ยในการประมาณค่า
แบบช่วง

Dim SumL, SumU As Double ' เก็บค่าผลรวมของขอบเขตล่างและขอบเขตบน

Dim Z(3) As Double

Dim T(3) As Double

ค่าคงที่ในการประมาณค่าแบบช่วง (สำหรับทุกวิธีและทุกระดับความเชื่อมั่น)

$Z(0)=1.645$

$Z(1)=1.96$

$Z(2)=2.576$

ค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด (สำหรับทุกวิธีและทุกระดับความเชื่อมั่น)

$T(0)=0.8930$

$T(1)=0.9449$

$T(2)=0.9877$

การคำนวณวิธีประมาณค่าแบบจุด

Private Sub Command1_Click()

Dim N, L, I, J, SumY As Integer

Dim Error1, Error2, Error3, vRnd, P As Double

```
Dim str As String
```

```
If varN.Text <> "" Then
```

```
    P = Val(Combo1.List(Combo1.ListIndex))
```

```
    N = Int(varN.Text)
```

```
    L = Int(varL.Text)
```

```
    Pro.Max = L
```

```
    Pro.Value = 0
```

```
    Error1=0
```

```
    Error2=0
```

```
    Error3=0
```

ขั้นที่ 1 สร้างตัวแปรสุ่มที่มีการแจกแจงทวินาม

```
For I = 1 To L
```

```
    SumY = 0
```

```
    For J = 1 To N
```

```
        Randomize Timer
```

```
        vRnd = Rnd
```

```
        If vRnd <= P Then
```

```
            SumX = SumX + 1
```

```
        End If
```

```
    Next
```

ขั้นที่ 2 คำนวณผลรวมของค่าความคลาดเคลื่อนสัมบูรณ์

```
Error1 = Error1 + Cal_MLE (SumX, N, P) ' คำนวณค่าจากฟังก์ชัน Cal_MLE
```

```
Error2 = Error2 + Cal_Bayes (SumX, N, P) ' คำนวณค่าจากฟังก์ชัน Cal_Bayes
```

```
Error3 = Error3 + Cal_Minimax (SumX, N, P) ' คำนวณค่าจากฟังก์ชัน Cal_Minimax
```

```
Next
```

ขั้นที่ 3 คำนวณค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ย

$Error1 = Error1 / L$ ' นำผลรวมค่าคลาดเคลื่อนสัมบูรณ์มาหารด้วยจำนวนรอบ สำหรับวิธี
ภาวะน่าจะเป็นสูงสุด

reCal(0).Text = Format(Error1, "0.0000")

$Error2 = Error2 / L$ ' นำผลรวมค่าคลาดเคลื่อนสัมบูรณ์มาหารด้วยจำนวนรอบ สำหรับวิธี
เบย์เซียน

reCal(1).Text = Format(Error2, "0.0000")

$Error3 = Error3 / L$ ' นำผลรวมค่าคลาดเคลื่อนสัมบูรณ์มาหารด้วยจำนวนรอบ สำหรับวิธี
มินิแมกซ์

reCal(3).Text = Format(Error3, "0.0000")

End If

End Sub

ฟังก์ชันการคำนวณวิธีภาวะน่าจะเป็นสูงสุด

Private Function Cal_MLE(ByVal Y As Integer, ByVal N As Integer, ByVal P As Double) As
Double

Dim P_hat, PL, PU, Error_MLE As Double

If Y >= 0 Then

P_hat = Y / N

Error_MLE = Abs(P - P_hat) ' คำนวณค่าคลาดเคลื่อนสัมบูรณ์

Cal_MLE = Error_MLE

Else

Cal_MLE = 0

End If

End Function

ฟังก์ชันการคำนวณวิธีเบส์เขียน

Private Function Cal_bayes(ByVal Y As Integer, ByVal N As Integer, ByVal P As Double) As

Double

Dim P_hat, PL, PU, Error_Bayes As Double

If Y >= 0 Then

Select Case P ' เลือกค่า P จากหน้าจอ

Case Is = 0.05

P_hat = (Y + 0.5)/(N+1)

Case Is = 0.07

P_hat = (Y + 0.5)/(N+1)

Case Is = 0.09

P_hat = (Y + 0.5)/(N+1)

Case Is = 0.1

P_hat = (Y + 0.5)/(N+1)

Case Is = 0.3

P_hat = (Y + 3) / (N + 6)

Case Is = 0.5

P_hat = (Y + 3) / (N + 6)

End Select

Error_Bayes = Abs(P - P_hat) ' คำนวณค่าคลาดเคลื่อนสัมบูรณ์

Cal_Bayes = Error_Bayes

Else

Cal_Bayes = 0

End If

End Function

ฟังก์ชันการคำนวณวิธีมินิแมกซ์

Private Function Cal_minimax(ByVal Y As Integer, ByVal N As Integer, ByVal P As Double)

As Double

Dim P_hat, PL, PU, Error_Minimax As Double

If Y >= 0 Then

$P_hat = (Y + (Sqr(N) / 2)) / (N + Sqr(N))$

Error_Minimax = Abs(P - P_hat) ' คำนวณค่าคลาดเคลื่อนสัมบูรณ์

Cal_minimax = Error_Minimax

Else

Cal_minimax = 0

End If

End Function

การคำนวณวิธีการประมาณค่าแบบช่วง

Private Sub Command1_Click()

Dim N, L, I, J As Integer

Dim SumX, Sum1, Sum2, Sum3, vRnd, SumT1, SumT2, SumT3, P As Double

Dim str As String

If varN.Text <> "" Then

P = Val(varP.Text)

N = Int(varN.Text)

L = Int(varL.Text)

Pro.Max = L

Pro.Value = 0

Sum1 = 0

Sum2 = 0

Sum3 = 0

SumU = 0

SumL = 0

SumT1 = 0

SumT2 = 0

SumT3 = 0

ขั้นที่ 1 สร้างตัวแปรสุ่มที่มีการแจกแจงทวินาม

For I = 1 To L

SumX = 0

For J = 1 To N

Randomize Timer

vRnd = Rnd

If vRnd <= P Then

SumX = SumX + 1

End If

Next

ขั้นที่ 2 นับจำนวนครั้งที่ช่วงความเชื่อมั่นคลุมค่าพารามิเตอร์ p

If Cal_Normal(SumX, N, P) = True Then ' ถ้าค่าจากฟังก์ชัน Cal_Normal เป็นจริง
จะนับจำนวนครั้งที่ช่วงความเชื่อมั่นคลุม
ค่าพารามิเตอร์ p

Sum1 = Sum1 + 1

End If

SumT1 = SumT1 + (SumU - SumL) ' คำนวณผลรวมของค่าความกว้างในการประมาณแบบ
ช่วงสำหรับวิธีปกติ

If Cal_logit(SumX, N, P) = True Then ' ถ้าค่าจากฟังก์ชัน Cal_logits เป็นจริงจะนับจำนวน
ครั้งที่ช่วงความเชื่อมั่นคลุมค่าพารามิเตอร์

Sum2 = Sum2 + 1

End If

SumT2 = SumT2 + (SumU - SumL) ' คำนวณผลรวมของค่าความกว้างในการประมาณแบบ
ช่วงสำหรับวิธีโลจิส

If Cal_score(SumX, N, P) = True Then ' ถ้าค่าจากฟังก์ชัน Cal_scores เป็นจริงจะนับ
จำนวนครั้งที่ช่วงความเชื่อมั่นคลุมค่าพารามิเตอร์ p

Sum3 = Sum3 + 1

End If

SumT3 = SumT3 + (SumU - SumL) ' คำนวณผลรวมของค่าความกว้างในการประมาณแบบ
ช่วงสำหรับวิธีสคอร์แบบปรับความต่อเนื่อง

Pro.Value = Pro.Value + 1

Next

ขั้นที่ 3 คำนวณค่าประมาณสัมประสิทธิ์ความเชื่อมั่นจากการทดลอง

Sum1 = Sum1 / L ' นำผลรวมของจำนวนที่ช่วงความเชื่อมั่นคลุมค่า p หารด้วยจำนวนรอบ
สำหรับวิธีปกติ

Sum2 = Sum2 / L ' นำผลรวมของจำนวนที่ช่วงความเชื่อมั่นคลุมค่า p หารด้วยจำนวนรอบ
สำหรับวิธีโลจิส

Sum3 = Sum3 / L ' นำผลรวมของจำนวนที่ช่วงความเชื่อมั่นคลุมค่า p หารด้วยจำนวนรอบ
สำหรับวิธีสคอร์แบบปรับความต่อเนื่อง

reCal(0).Text = Format(Sum1, "0.0000")

reCal(1).Text = Format(Sum2, "0.0000")

reCal(3).Text = Format(Sum3, "0.0000")

ขั้นที่ 4 คำนวณค่าความกว้างเฉลี่ย

If Sum1 >= T(Combo1.ListIndex) Then ' หากค่าประมาณสัมประสิทธิ์ความเชื่อมั่นจากการทดลองในวิธีปกติมีค่ามากกว่าหรือเท่ากับค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนดแล้วทำการหาค่าความกว้างเฉลี่ยต่อไป

SumT1 = SumT1 / L ' คำนวณค่าความกว้างเฉลี่ย สำหรับวิธีปกติ

reCalT(0).Text = Format(SumT1, "0.0000")

Else

reCalT(0).Text = "ไม่ผ่าน"

End If

If Sum2 >= T(Combo1.ListIndex) Then

SumT2 = SumT2 / L ' คำนวณค่าความกว้างเฉลี่ย สำหรับวิธีโลจิส

reCalT(1).Text = Format(SumT2, "0.0000")

Else

reCalT(1).Text = "ไม่ผ่าน"

End If

If Sum3 >= T(Combo1.ListIndex) Then

SumT3 = SumT3 / L ' คำนวณค่าความกว้างเฉลี่ย สำหรับวิธีสคอ์แบบปรับความต่อเนื่อง

reCalT(2).Text = Format(SumT3, "0.0000")

Else

```

reCalT(2).Text = “ไม่ผ่าน”
End If
End If
End Sub

```

ฟังก์ชันการคำนวณวิธีปกติ

```

Private Function Cal_Normal(ByVal Y As Integer, ByVal N As Integer, ByVal P As Double) As
Boolean
Dim P_hat, PL, PU, N1 As Double
SumU = 0
SumL = 0

If Y >= 0 Then
P_hat = Y / N
N1 = Z(Combo1.ListIndex) * Sqr(P_hat * (1 - P_hat) / N)
PL = P_hat - N1 ' คำนวณค่าขอบเขตล่าง
PU = P_hat + N1 ' คำนวณค่าขอบเขตบน

If PL <= P And PU >= P Then
SumL = PL
SumU = PU
Cal_Normal = True
Else
Cal_Normal = False
End If
Else
Cal_Normal = False

```

End If

End Function

ฟังก์ชันการคำนวณวิธีโลจิส

Private Function Cal_logit(ByVal Y As Integer, ByVal N As Integer, ByVal P As Double) As

Boolean

Dim P_hat, Q_hat, PL, PU, L1, LL, UU As Double

SumL = 0

SumU = 0

If Y >= 0 Then

L1 = N * (Y + 1)

L1 = L1 * (N - Y + 1)

L1 = (N + 1) / L1

L1 = (N + 2) * L1

L1 = Sqr L1

L1 = Z(Combo1.ListIndex) * L1

LL = Log((Y + 0.5) / (N - Y + 0.5)) - L1

UU = Log((Y + 0.5) / (N - Y + 0.5)) + L1

PL = (Exp(LL)) / (1 + Exp(LL)) ' คำนวณค่าขอบเขตล่าง

PU = (Exp(UU)) / (1 + Exp(UU)) ' คำนวณค่าขอบเขตบน

If PL <= P And PU >= P Then

SumL = PL

SumU = PU

```

    Cal_logit = True
Else
    Cal_logit = False
End If
Else
    Cal_logit = False
End If
End Function

```

ฟังก์ชันการคำนวณวิธีสคอร์แบบปรับความต่อเนื่อง

```

Private Function Cal_score(ByVal Y As Integer, ByVal N As Integer, ByVal P As Double) As
Boolean

```

```

Dim P_hat, Q_hat, PL, PU, S1 As Double

```

```

    SumL = 0

```

```

    SumU = 0

```

```

If Y >= 0 Then

```

```

    P_hat = Y / N

```

```

    Q_hat = 1 - P_hat

```

```

    S1 = (2 * (N + (Z(Combo1.ListIndex) * Z(Combo1.ListIndex))))

```

```

    PL = (((2 * N * P_hat) + (Z(Combo1.ListIndex) * Z(Combo1.ListIndex)) - 1) -
Z(Combo1.ListIndex) * Sqr((Z(Combo1.ListIndex) * Z(Combo1..ListIndex)) - (2 + (1 / N)) +
(4 * P_hat * ((N * Q_hat) + 1)))) / S1      ' คำนวณค่าขอบเขตล่าง

```

```

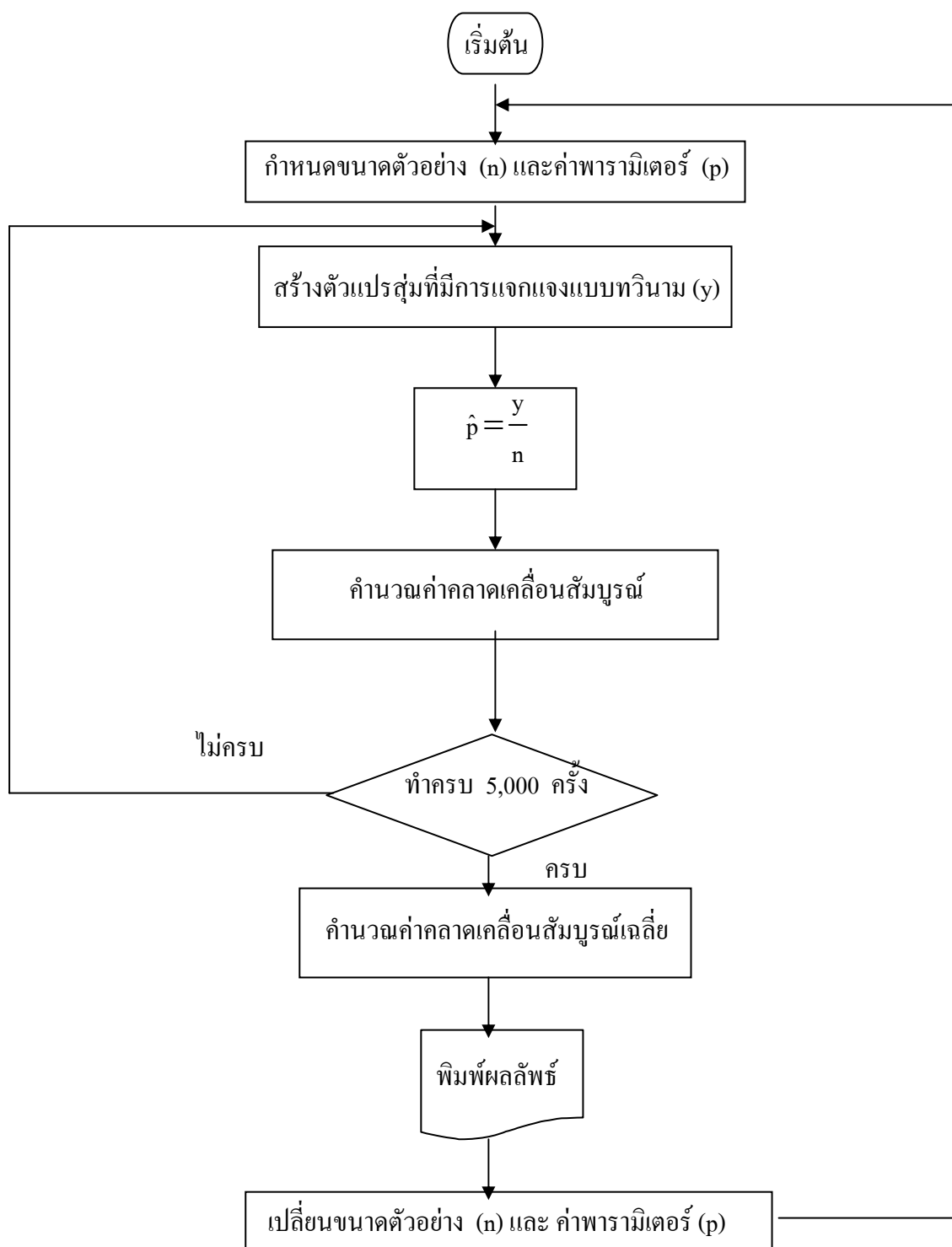
    PU = (((2 * N * P_hat) + (Z(Combo1.ListIndex) * Z(Combo1..ListIndex)) + 1)
++Z(Combo1.ListIndex) * Sqr((Z(Combo1.ListIndex) * Z(Combo1.ListIndex)) + (2 - (1 / N)) +
(4 * P_hat * ((N * Q_hat) - 1)))) / S1      ' คำนวณค่าขอบเขตบน

```

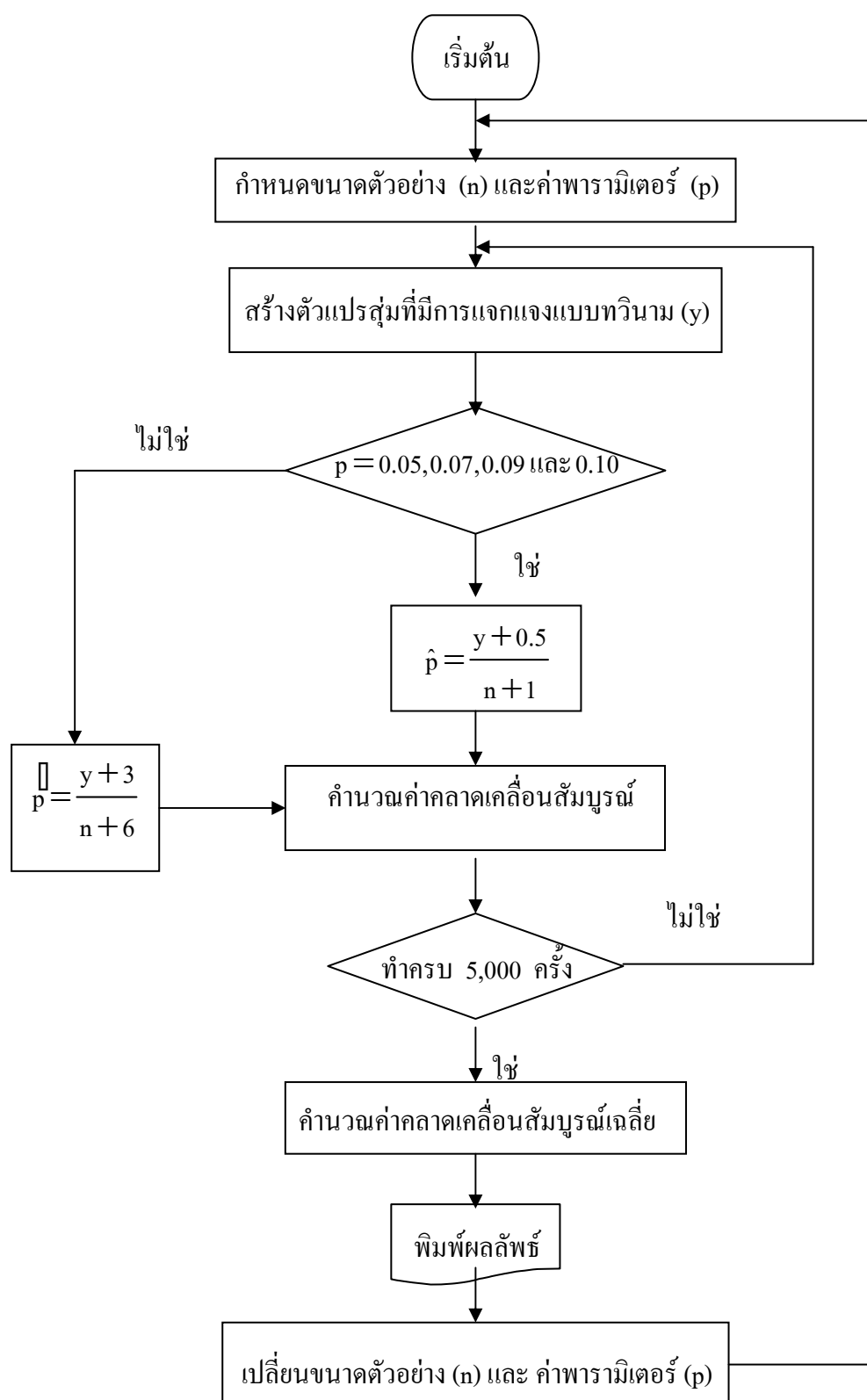
```
If PL <= P And PU >= P Then
    SumL = PL
    SumU = PU
    Cal_score = True
Else
    Cal_score = False
End If
Else
    Cal_score = False
End If
End Function
```

ภาคผนวก ข

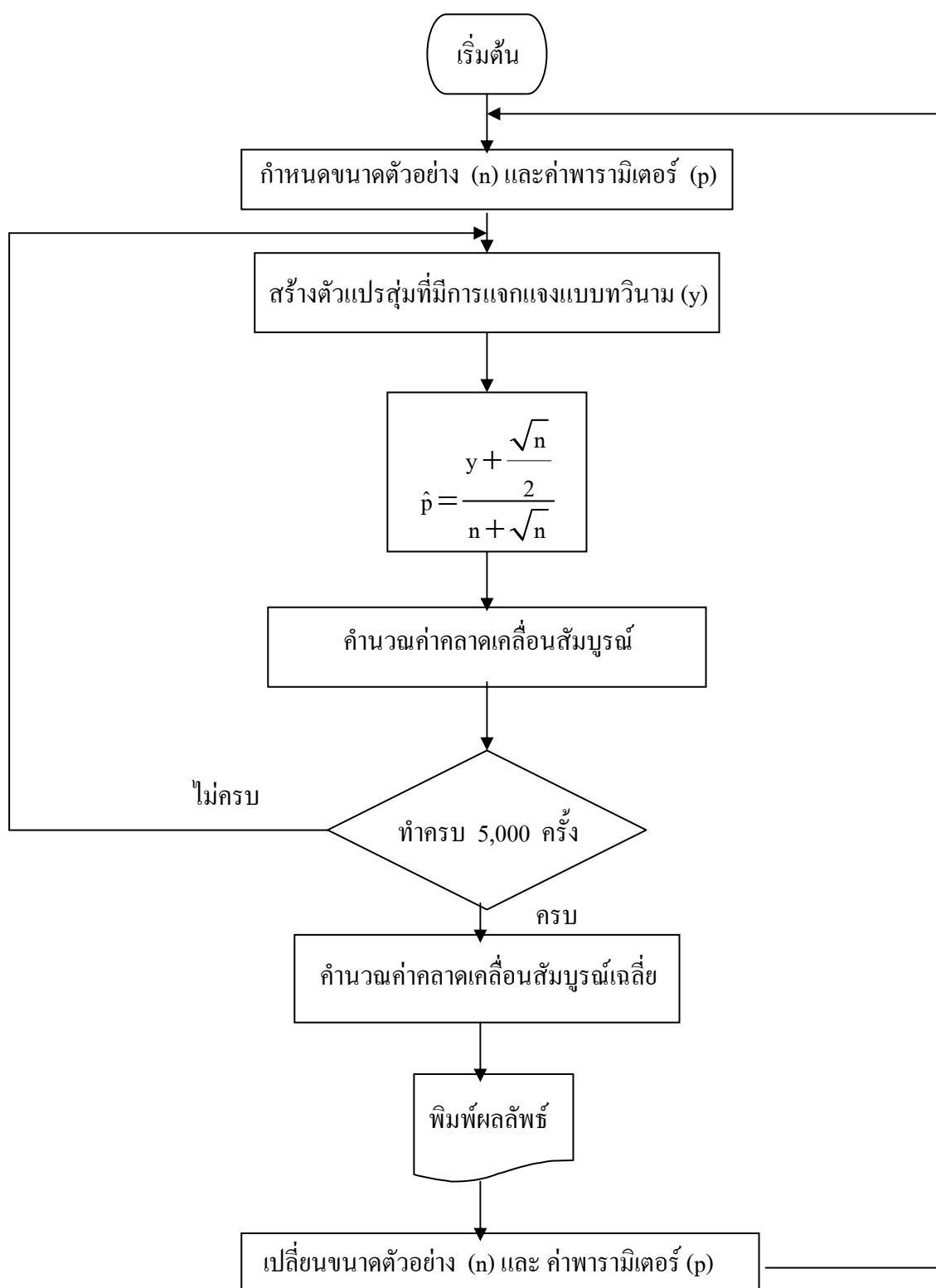
แผนผังการคำนวณแต่ละวิธีสำหรับการประมาณค่าแบบจุดและแบบช่วงที่ใช้ในงานวิจัย



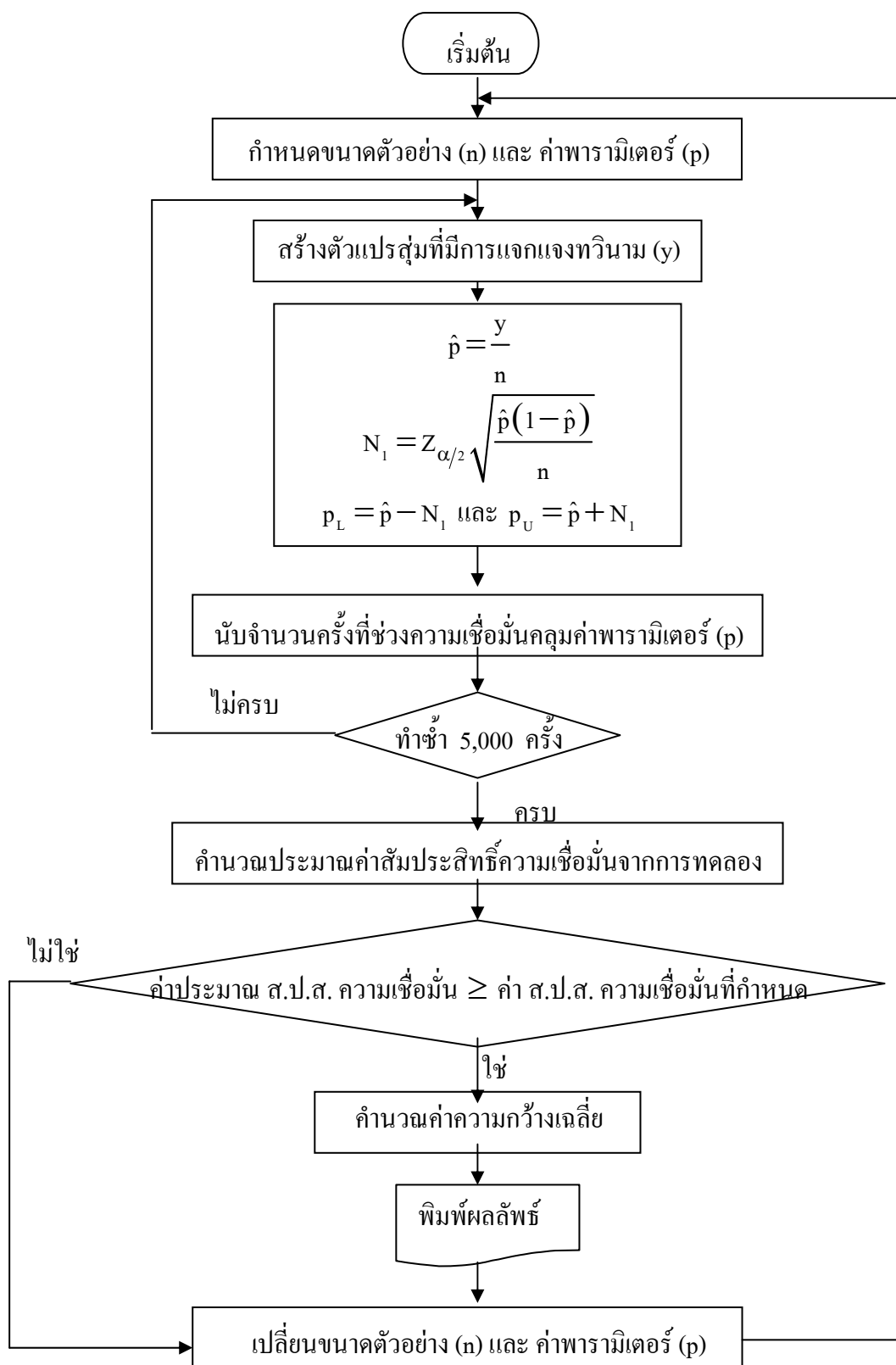
ภาพผนวกที่ ข1 ขั้นตอนการประมาณค่าแบบจุดโดยวิธีภาวะน่าจะเป็นสูงสุด



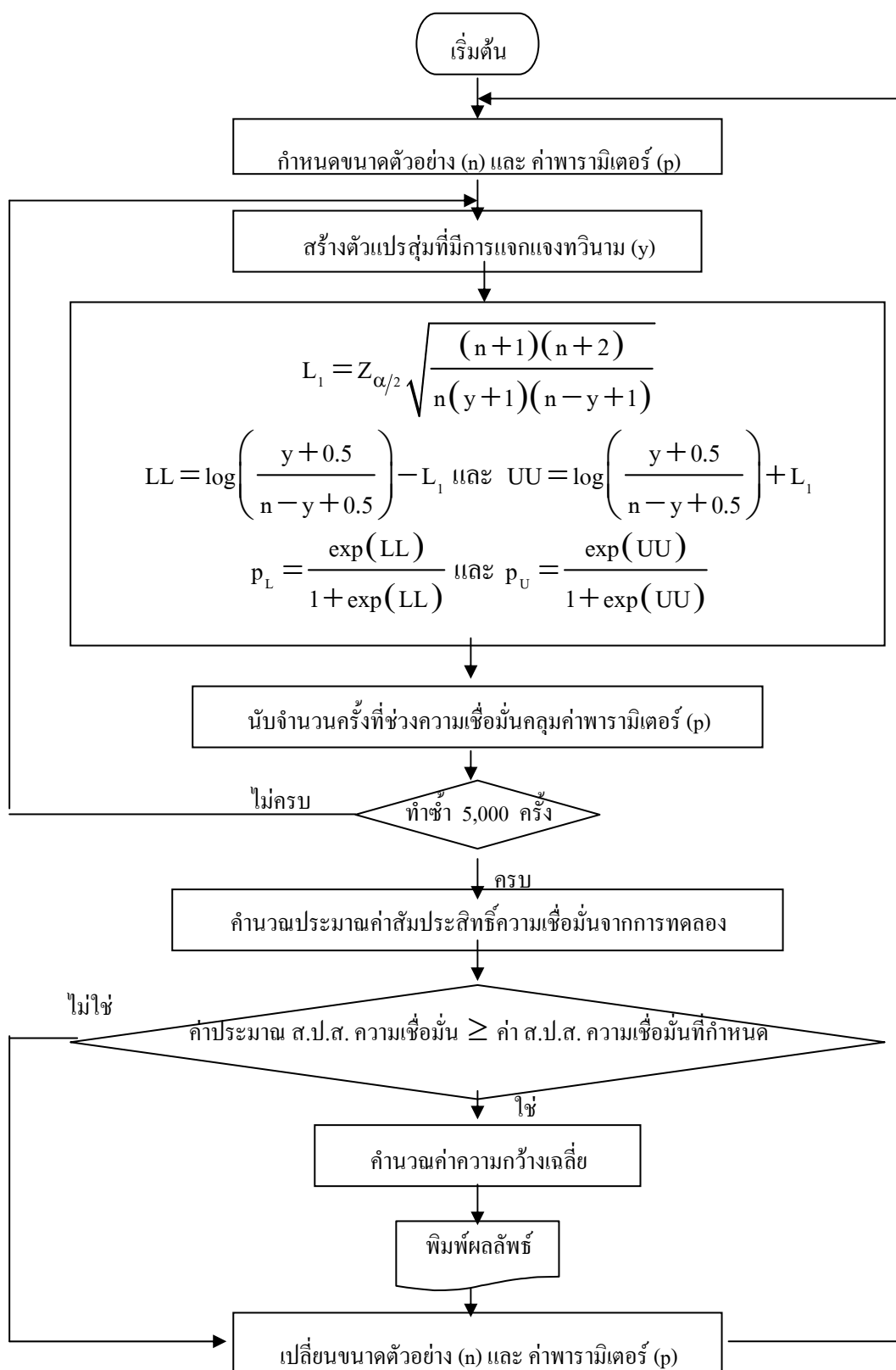
ภาพผนวกที่ ข2 ขั้นตอนการประมาณค่าแบบจุดโดยวิธีเบสส์เชียน



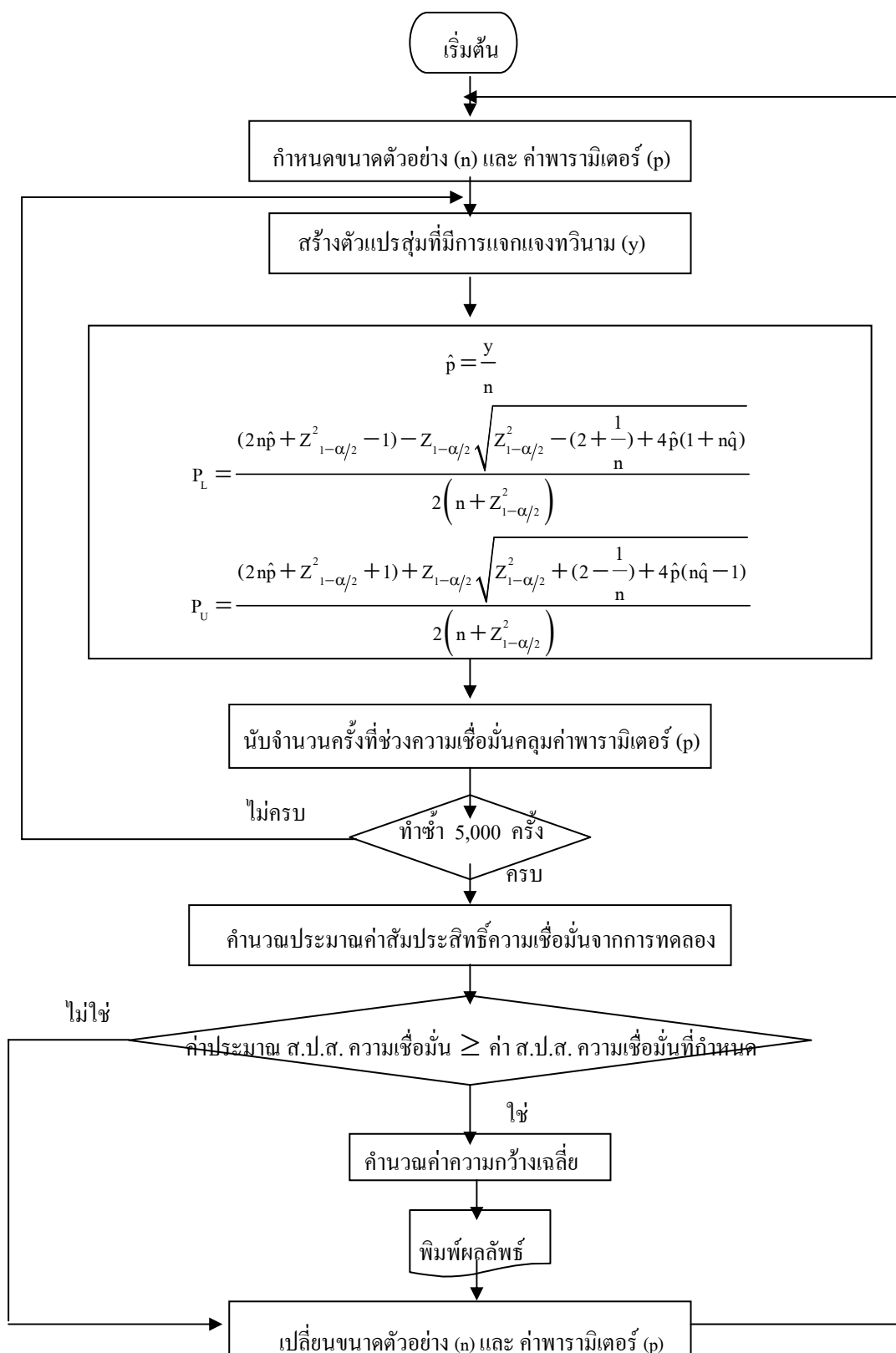
ภาพผนวกที่ ข3 ขั้นตอนการประมาณค่าแบบจุดโดยวิธีมินิแมกซ์



ภาพผนวกที่ ข4 ขั้นตอนการประมาณค่าแบบช่วงโดยวิธีปกติ



ภาพผนวกที่ ข5 ขั้นตอนการประมาณค่าแบบช่วงโดยวิธีโลจิส



ภาพผนวกที่ ข6 ขั้นตอนการประมาณค่าแบบช่วงโดยวิธีสก็อร์แบบปรับความต่อเนื่อง

ประวัติการศึกษา และการทำงาน

ชื่อ –นามสกุล	นางสาวมัลลิกา ทานุสิทธิ์
วัน เดือน ปี ที่เกิด	วันที่ 21 ธันวาคม 2525
สถานที่เกิด	อำเภอเมือง จังหวัดลำปาง
ประวัติการศึกษา	วท.บ.(คณิตศาสตร์) มหาวิทยาลัยเชียงใหม่
ตำแหน่งหน้าที่การงานปัจจุบัน	
สถานที่ทำงานปัจจุบัน	
ผลงานดีเด่นและรางวัลทางวิชาการ	
ทุนการศึกษาที่ได้รับ	