บทที่ 2
วรรณกรรมและงานวิจัยที่เกี่ยวข้อง

การทบทวนวรรณกรรมและงานวิจัยที่เกี่ยวข้องกับการศึกษาขนาดตัวอย่างที่เหมาะสมในการประมาณค่าขีดจำกัดบนช่วงเชื่อถือเบย์เซียน เมื่อผลลัพธ์ในตัวอย่างเป็นศูนย์ สำหรับการรับรองคุณภาพโรงพยาบาล ผู้ศึกษาได้ทบทวนจากเอกสาร ทฤษฎีต่างๆ และงานวิจัยที่เกี่ยวข้องเพื่อเป็นแนวทางในการศึกษา โดยแบ่งเนื้อหาสาระออกเป็นประเด็นหลักดังนี้

1. แนวคิดการพัฒนาและรับรองคุณภาพโรงพยาบาล

1.1 ความหมายของการรับรองคุณภาพโรงพยาบาล

1.2 คำจำกัดความของคุณภาพในงานบริการทางการแพทย์

1.3 การพัฒนาคุณภาพ (Quality Improvement)

1.4 การประเมินคุณภาพ (Quality Assessment)

1.5 การรับรองโรงพยาบาล

1.6 การกำหนดเกณฑ์ชี้วัดคุณภาพโรงพยาบาล

2. ความรู้พื้นฐานที่เกี่ยวข้องกับงานวิจัย

2.1 ทฤษฎีการแจกแจงของประชากร

2.2 ความน่าจะเป็นแบบมีเงื่อนไข (conditional probability)

2.3 ทฤษฎีบทของเบย์ (Bayes' theorem)

2.4 วิธีการและการอ้างอิงทางสถิติแบบเบย์เซียน

2.5 การกำหนดขนาดตัวอย่างแบบเบย์เซียน

2.6 การกำหนดขนาดตัวอย่างของการสุ่มตัวอย่างเพื่อการยอมรับ

3. วิธีการประมาณค่าเมื่อไม่พบผลลัพธ์ในตัวอย่าง

3.1 การประมาณค่าด้วย Exact binomial

3.2 การประมาณค่าแบบเบย์เซียน (Bayesian estimate)

3.2.1 กรณีทราบความน่าจะเป็นเบื้องต้น (Informative prior)

3.2.2 กรณีไม่ทราบความน่าจะเป็นเบื้องต้น (Non-informative prior)

1. แนวคิดการพัฒนาและรับรองคุณภาพโรงพยาบาล (สถาบันพัฒนาและรับรองคุณภาพโรงพยาบาล, 2550)

1.1 ความหมายของการรับรองคุณภาพโรงพยาบาล

การรับรองคุณภาพโรงพยาบาล คือ กลไกกระตุ้นให้เกิดการพัฒนาคุณภาพของโรงพยาบาล ควบคู่ไปกับการเรียนรู้แลกเปลี่ยน และการรับรองจากองค์กรภายนอก การรับรองเป็นเพียงส่วนเดียวและส่วนสุดท้ายของกระบวนการ แต่จุดสำคัญคือการกำหนดมาตรฐาน ตรวจสอบ และพัฒนาตนเองอย่างต่อเนื่องของโรงพยาบาล

การรับรองคุณภาพจะกระทำโดยองค์กรภายนอกที่เป็นกลาง เพื่อเป็นหลักประกันว่าผลการรับรองนั้นจะเป็นที่น่าเชื่อถือ สถาบันพัฒนาและรับรองคุณภาพโรงพยาบาลเป็นหน่วยงานที่จัดตั้งขึ้นโดยข้อบังคับตามพระราชบัญญัติสถาบันวิจัยระบบสาธารณสุข ทำหน้าที่ส่งเสริมการพัฒนาคุณภาพของโรงพยาบาล โดยอาศัยการประเมินตนเองร่วมกับการประเมินจากภายนอกเป็นกลไกกระตุ้นที่สำคัญ

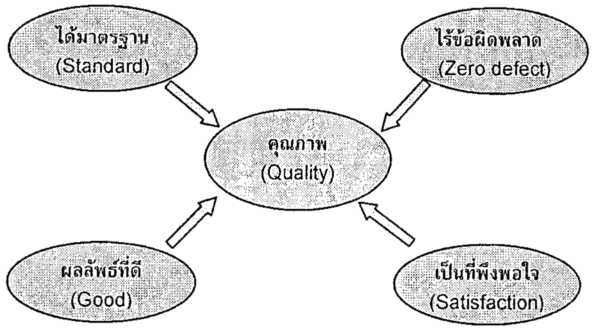
ก่อนที่จะขอรับรองคุณภาพจากสถาบันพัฒนาและรับรองคุณภาพโรงพยาบาลเป็นหน้าที่ของโรงพยาบาลที่จะต้องประเมินและพัฒนาคุณภาพตามกรอบมาตรฐานโรงพยาบาลมาก่อน

1.2 คำจำกัดความของคุณภาพในงานบริการทางการแพทย์

Institute of Medicine กล่าวไว้ว่า คุณภาพควรประกอบไปด้วยความพึงพอใจด้านสุขภาพของประชาชนหรือแต่ละบุคคลต่อระดับการให้บริการสุขภาพที่จัดให้สอดคล้องกับความรู้ที่มีอยู่ในปัจจุบัน (Lohr, 1990 อ้างใน อนุวัฒน์ ศุภชุติกุล และจิรุตม์ ศรีรัตนบัลล์, 2543)

อนุวัฒน์ ศุภชุติกุล และคณะ (2542) คุณภาพบริการ หมายถึง การตอบสนองความต้องการที่จำเป็นของลูกค้าโดยอยู่บนพื้นฐานของมาตรฐานวิชาชีพ ซึ่งประกอบด้วยความถูกต้องตามมาตรฐาน และความถูกต้องตรงตามความต้องการและความคาดหวังของผู้รับบริการ

อย่างไรก็ตาม คุณภาพสำหรับบริการทางการแพทย์ อาจหมายถึง ลักษณะของบริการที่เป็นไปตามมาตรฐานที่เหมาะสม ปราศจากข้อผิดพลาด ส่งผลให้เกิดผลลัพธ์ดีต่อคุณภาพชีวิต สามารถตอบสนองความต้องการการดูแล และเป็นที่พึงพอใจของผู้ใช้บริการ ดังภาพที่ 2



ภาพที่ 2 กรอบคำจำกัดความของคุณภาพงานบริการทางการแพทย์

1.3 การพัฒนาคุณภาพ (Quality Improvement)

การพัฒนาคุณภาพ คือ การจัดระบบบริหารและระบบการทำงานในโรงพยาบาลตามแนวทางที่กำหนดไว้ในมาตรฐานโรงพยาบาล ซึ่งมุ่งเน้นการทำงานด้วยใจที่มุ่งมั่นต่อคุณภาพของเจ้าหน้าที่ การทำงานเป็นทีม การตอบสนองความต้องการของผู้ป่วยและผู้รับผลงาน มีระบบตรวจสอบเพื่อแก้ไขปรับปรุงด้วยการประสานกิจกรรมการบริหารความเสี่ยง (risk management-RM) การประกันคุณภาพ (quality assurance-QA) และการพัฒนาคุณภาพอย่างต่อเนื่อง (continual quality improvement-CQI) เข้าด้วยกัน

การบริหารความเสี่ยง คือ การค้นหาโอกาสที่จะเกิดความเสี่ยง การประเมินความเสี่ยง การวางมาตรการป้องกันความเสี่ยง และการดำเนินการเมื่อเกิดความสูญเสียขึ้น

การประกันคุณภาพ คือ การวางระบบเพื่อเป็นหลักประกันว่าจะมีการปฏิบัติงานตามมาตรฐาน หรือแนวทางปฏิบัติที่กำหนดและมีผลลัพธ์ตามที่คาดไว้ ซึ่งประกอบด้วยการกำหนดมาตรฐาน การวัดผลการปฏิบัติ และการปรับปรุงแก้ไขเมื่อไม่มีการปฏิบัติตามแนวทางที่กำหนดหรือผลลัพธ์ไม่เป็นไปตามที่คาด

การพัฒนาคุณภาพอย่างต่อเนื่อง คือ การใช้กระบวนการทางวิทยาศาสตร์และความคิดสร้างสรรค์ในการปรับปรุงระบบงานเพื่อตอบสนองความต้องการของผู้รับผลงานอย่างไม่หยุดยั้งโดยมุ่งไปสู่ความเป็นเลิศ

1.4 การประเมินคุณภาพ (Quality Assessment)

การประเมินคุณภาพ คือ การตรวจสอบระบบงานและสิ่งที่ปฏิบัติ กับข้อกำหนดในมาตรฐานโรงพยาบาล ซึ่งจะทำโดยโรงพยาบาลและโดยผู้ประเมินภายนอก

1.4.1 การประเมินคุณภาพโดยโรงพยาบาล (Self Assessment)

เป็นการประเมินตนเองเพื่อตรวจสอบความก้าวหน้าในการพัฒนาคุณภาพ และตรวจสอบความพร้อมที่จะได้รับการประเมินและรับรองจากภายนอก การประเมินตนเองของโรงพยาบาลควรประเมินโดยทีมที่เกี่ยวข้อง โดยควรครอบคลุมการประเมินเพื่อค้นหาโอกาสพัฒนา การตรวจเยี่ยมเพื่อสังเกตการปฏิบัติงานจริง การทบทวนแนวคิด แนวทางปฏิบัติ การปฏิบัติงานจริง และผลลัพธ์ที่เกิดขึ้น เพื่อนำไปสู่การแก้ปัญหาและพัฒนาวิธีการทำงานอย่างต่อเนื่อง การใช้แบบประเมินตนเองเพื่อบันทึกและวิเคราะห์ระบบงานตามข้อกำหนดในมาตรฐานโรงพยาบาล

1.4.2 การประเมินโดยผู้ประเมินภายนอก (External Survey)

การประเมินโดยผู้ประเมินภายนอกมี 3 ลักษณะ คือ การประเมินความพร้อมของโรงพยาบาล การประเมินเพื่อพิจารณารับรอง และการประเมินหลังการรับรอง

(1) การประเมินความพร้อมของโรงพยาบาล (Preparation Survey) เป็นการประเมินเพื่อดูว่าโรงพยาบาลได้ดำเนินการพัฒนาคุณภาพตามข้อกำหนดในกรอบมาตรฐานโรงพยาบาลได้ครบถ้วนแล้วหรือไม่ มีประเด็นความเสี่ยงที่ชัดเจนหลงเหลืออยู่หรือไม่

โรงพยาบาลถ้วนแล้วหรือไม่ มีประเด็นความเสี่ยงที่ชัดเจนหลงเหลืออยู่หรือไม่ โรงพยาบาลจะขอให้มีการทำ presurvey ต่อเมื่อผลการประเมินตนเองอยู่ในระดับที่มั่นใจว่าได้มีการพัฒนาตามมาตรฐานโรงพยาบาลในประเด็นสำคัญ ๆ ครบถ้วน ผลประเมินในขั้นตอนนี้คือการให้คำแนะนำเพื่อให้โรงพยาบาลนำไปปรับปรุง การประเมินความพร้อมอาจจะทำเป็นระยะ ๆ หลายครั้ง จนกว่าจะมั่นใจว่าโรงพยาบาลมีความพร้อมเต็มที่สำหรับการประเมินเพื่อรับรอง

(2) การประเมินเพื่อพิจารณารับรอง (Accreditation Survey) คือ การไปรับทราบหลักฐานและความจริงว่าโรงพยาบาลได้มีการปฏิบัติตามข้อกำหนดในมาตรฐานโรงพยาบาล สิ่งที่ระบุไว้ในนโยบาย/คู่มือการปฏิบัติงานของโรงพยาบาล คำแนะนำของผู้เชี่ยวชาญเฉพาะทาง และข้อเสนอเพื่อการปรับปรุงจากการประเมินความพร้อม เครื่องมือสำคัญที่ผู้ประเมินภายนอกจะใช้คือข้อมูลที่โรงพยาบาลประเมินตนเองตามแบบฟอร์มที่กำหนดไว้ซึ่งผู้ประเมินภายนอกจะต้องศึกษาล่วงหน้าก่อนที่จะไปประเมินในพื้นที่

(3) การประเมินหลังการรับรอง มี 3 ลักษณะได้แก่

1) การประเมินเฝ้าระวัง (Surveillance Survey) เป็นการประเมินตามกำหนดเวลาทุก 6-12 เดือน โดยเน้นประเด็นสำคัญหรือประเด็นที่มีแนวโน้มจะมีปัญหาในภาพรวม

2) การประเมินเมื่อมีปัญหา (Unscheduled Survey) เป็นการประเมินเมื่อได้รับทราบว่าอาจจะมีปัญหารุนแรงเกี่ยวกับการดูแลผู้ป่วยหรือความปลอดภัย

3) การประเมินเมื่อมีการปรับเปลี่ยน (Verification Survey) ได้แก่ การเปิดบริการ การขยายบริการ การเปลี่ยนเจ้าของหรือผู้บริหารระดับสูง เมื่อมีการเปลี่ยนแปลงเกิดขึ้น โรงพยาบาลจะต้องแจ้งให้ทราบภายใน 30 วัน และจะมีการประเมินซ้ำเพื่อยืนยันการรับรองหากอายุการรับรองยังมีเหลือมากกว่า 9 เดือน โดยอายุการรับรองจะไม่ขยายมากกว่าเดิม

1.5 การรับรองโรงพยาบาล

การรับรองโรงพยาบาล คือ การรับรองว่าโรงพยาบาลมีการจัดระบบงานที่ดี เอื้อต่อการให้บริการได้อย่างมีคุณภาพ และปลอดภัย มีความมุ่งมั่นที่จะทำงานให้มีคุณภาพและพัฒนาคุณภาพอย่างต่อเนื่อง มีการตรวจสอบตนเองอย่างสม่ำเสมอ

ผู้ตัดสินให้การรับรองโรงพยาบาล คือ คณะกรรมการผู้ทรงคุณวุฒิซึ่งสถาบันพัฒนาและรับรองคุณภาพแต่งตั้งขึ้น โดยคณะกรรมการจะตัดสินใจจากการพิจารณาข้อมูลที่ได้รับจากผู้ประเมินภายนอก ลักษณะของการรับรองมีดังนี้

1.5.1 รับรอง 2 ปี สำหรับโรงพยาบาลที่สามารถปฏิบัติตามข้อกำหนดในมาตรฐานได้ครบถ้วน บริการส่วนใหญ่อยู่ในระดับดี ไม่มีความเสี่ยงที่ชัดเจน มีหลักฐานของความพยายามในการปรับปรุงอย่างต่อเนื่อง

1.5.2 ไม่รับรอง สำหรับโรงพยาบาลที่ไม่สามารถปฏิบัติตามมาตรฐานโรงพยาบาลได้ครบถ้วนยังมีความเสี่ยงปรากฏอย่างชัดเจน

1.6 การกำหนดเกณฑ์ชี้วัดคุณภาพโรงพยาบาล

ก่อนที่จะมีการกำหนดเกณฑ์ชี้วัดคุณภาพจะต้องมีการกำหนดตัวชี้วัดคุณภาพโรงพยาบาล ซึ่งตัวชี้วัดคุณภาพเป็นข้อความที่บ่งชี้ถึงผลที่คาดว่าจะเกิดขึ้นจากการปฏิบัติงาน โดยใช้เป็นตัววัดและประเมินผลการปฏิบัติงานว่าบรรลุถึงคุณภาพที่กำหนดไว้หรือไม่ การกำหนดตัวชี้วัดสามารถกำหนดได้จากมาตรฐานการดูแลหรือมาตรฐานการปฏิบัติงาน การศึกษา วิจัย หรือข้อเสนอแนะขององค์กรที่ทำหน้าที่ในการควบคุมคุณภาพ โดยสมาชิกในทีมพัฒนาคุณภาพหรือคณะทำงานประกันคุณภาพร่วมกันวิเคราะห์และกำหนดตัวชี้วัดคุณภาพโรงพยาบาล

ตารางที่ 2 ตัวอย่างการกำหนดเกณฑ์ชี้วัดคุณภาพการพยาบาลที่มีโอกาสเกิดน้อย

ตัวชี้วัด	เกณฑ์	แหล่งข้อมูล/วิธีตรวจสอบ
1. อุบัติการณ์การให้การรักษาพยาบาลผิดคน เนื่องจากการระบุตัวผู้ป่วยผิดคน (Mistaken patient identification)	0	- สุ่มตรวจสอบกระบวนการระบุตัวผู้ป่วยขณะให้การพยาบาล - ตรวจสอบเอกสารมาตรฐานวิธีปฏิบัติเกี่ยวกับการระบุตัวบุคคลผู้ป่วยไม่รู้สึกตัวหรือผู้ป่วยไม่ทราบชื่อ - ตรวจสอบจากรายงานอุบัติการณ์
2. อุบัติการณ์ความผิดพลาดในการให้ยา เลือด หรือส่วนประกอบของเลือด	0	- สุ่มตรวจสอบวิธีปฏิบัติในการให้ยา เลือด หรือส่วนประกอบของเลือด และตรวจสอบจากรายงานอุบัติการณ์
3. จำนวนอุบัติการณ์การบาดเจ็บจากการผูกมัด (Restraint) หรือการใช้อุปกรณ์/เครื่องมือทางการแพทย์ หรือแผลไหม้จากความร้อนของเครื่องมือ	0	- สุ่มสังเกตวิธีการผูกมัดผู้ป่วย สุ่มตรวจสภาพผิวหนังผู้ป่วย และตรวจสอบรายงานอุบัติการณ์
4. อุบัติการณ์การส่งผู้ป่วยผ่าตัดผิดคน	0	- ตรวจสอบจากเอกสารวิธีปฏิบัติการตรวจสอบผู้ป่วยก่อนทำการผ่าตัด และสุ่มสอบถามพยาบาลห้องผ่าตัด และวิสัญญีพยาบาลเกี่ยวกับวิธีปฏิบัติในการตรวจสอบผู้ป่วย ร่วมกับตรวจสอบรายงานอุบัติการณ์
5. จำนวนอุบัติการณ์การมีสิ่งของหรืออุปกรณ์ตกค้างในร่างกายผู้ป่วยหลังผ่าตัด	0	- ตรวจสอบจากเอกสารวิธีปฏิบัติในการตรวจนับอุปกรณ์ (Counting procedure) และสุ่มตรวจสอบจากบันทึกการตรวจนับ

หลังจากที่มีการกำหนดตัวชี้วัดคุณภาพโรงพยาบาลแล้วจะต้องมีการกำหนดเกณฑ์ชี้วัดคุณภาพ เพื่อการวัดและประเมินคุณภาพการปฏิบัติตามมาตรฐานที่กำหนดขึ้น ซึ่งเกณฑ์ชี้วัดคุณภาพจะเป็นตัวที่ใช้เปรียบเทียบกับผลการปฏิบัติว่าได้ผลลัพธ์ตามที่คาดหวังไว้หรือไม่ การกำหนดเกณฑ์ชี้วัดคุณภาพจะขึ้นอยู่กับแต่ละหน่วยงาน ซึ่งโรงพยาบาลแต่ละแห่งจะต้องมีการศึกษาข้อมูลทางด้านสถิติ รายงานอุบัติการณ์ต่าง ๆ เพื่อนำมากำหนดเกณฑ์ชี้วัดคุณภาพ สำหรับการศึกษาครั้งนี้ขอยกตัวอย่างการกำหนดเกณฑ์ชี้วัดคุณภาพการพยาบาลที่สามารถใช้เป็นเกณฑ์กลางในการเปรียบเทียบระดับคุณภาพการพยาบาลของแต่ละหน่วยงานของโรงพยาบาล (กฤษดา แสวงดี และคณะ, 2542) ดังตารางที่ 2

จากตัวอย่างการกำหนดเกณฑ์ชี้วัดคุณภาพการพยาบาลที่มีโอกาสเกิดน้อย พบว่าเกณฑ์คุณภาพที่ยอมรับมีการกำหนดเกณฑ์ไว้ในระดับสูง คือการไม่พบผลลัพธ์ในตัวชี้วัด ทั้งนี้การกำหนดเกณฑ์คุณภาพขึ้นอยู่กับความสำคัญของงานบริการ และผลลัพธ์ที่อาจเป็นอันตรายต่อผู้ป่วย แม้ว่าจะมีการกำหนดเกณฑ์คุณภาพไว้ที่ 0 แต่จากรายงานผลคุณภาพจะเห็นได้ว่าอุบัติการณ์ของการเกิดผลลัพธ์มีอุบัติการณ์เกิดน้อยมาก เช่น ความผิดพลาดในการให้ยาร้อยละ 0.01 การบาดเจ็บจากการผูกมัด การใช้อุปกรณ์หรือเครื่องมือทางการแพทย์หรือแผลไหม้จากความร้อนของเครื่องมือร้อยละ 5.10 ความผิดพลาดในการประเมินอาการเพื่อจำแนกและส่งผู้ป่วยไปยังห้องตรวจโรคร้อยละ 0.005 (โรงพยาบาลบุรีรัมย์, 2547) การให้เลือดผิดหมู่ในผู้ป่วยที่ได้รับระงับความรู้สึก 0.18 ต่อ 10,000 ราย (สมรัตน์ จารุลักษณานันท์ และคณะ, 2005) เป็นต้น เพื่อให้การประเมินสามารถที่จะคำนวณโอกาสของการเกิดผลลัพธ์ได้ แต่ละหน่วยงานจึงควรกำหนดเกณฑ์ชี้วัดคุณภาพที่สามารถยอมรับได้ ถ้าผลลัพธ์ที่เกิดขึ้นก่อให้เกิดผลเสียต่อผู้ป่วยอย่างรุนแรง และไม่ควรเกิดขึ้น แต่ในความเป็นจริงอาจมีการเกิดขึ้น ดังนั้นหน่วยงานอาจจะกำหนดเกณฑ์ไว้ในระดับสูง เช่น การลืมอุปกรณ์การแพทย์ไว้ในตัวผู้ป่วยที่ได้รับการผ่าตัดช่องท้องไม่ควรเกิดขึ้นมากกว่า 1 ต่อ 100,000 รายที่ได้รับการผ่าตัด หรือกำหนดให้มาตรฐานสูงมากขึ้นอาจจะกำหนดเกณฑ์เท่ากับ 1 ต่อ 1,000,000 ราย เป็นต้น ซึ่งการกำหนดเกณฑ์คุณภาพที่ชัดเจน แทนการกำหนดเกณฑ์คุณภาพเท่ากับ 0 เป็นการสร้างระดับเกณฑ์คุณภาพที่สามารถประเมินได้ และทราบโอกาสของการเกิดผลลัพธ์ที่ชัดเจนกว่า ซึ่งการกำหนดเกณฑ์ชี้วัดคุณภาพอาจมีการเปลี่ยนแปลงได้ และการประเมินผลคุณภาพงานบริการควรกำหนดเกณฑ์คุณภาพในระดับที่สูงกว่าสภาพในปัจจุบัน เพื่อเป็นจุดในการเปรียบเทียบผลการประเมิน

สำหรับการศึกษาในครั้งนี้ผู้วิจัยได้คำนึงถึงการนำไปประยุกต์ใช้ได้กับสถานการณ์จริง จึงได้จำลองสถานการณ์ให้มีลักษณะเช่นเดียวกันกับการกำหนดเกณฑ์ชี้วัดคุณภาพจริง เพื่อให้สามารถนำผลการศึกษาไปใช้ได้กับทุกหน่วยงาน ผู้วิจัยจึงจำลองเกณฑ์ชี้วัดคุณภาพให้มีความครอบคลุมโดยกำหนดเป็นค่าสัดส่วนตั้งแต่ 0.00001 ถึง 0.01

2. ความรู้พื้นฐานที่เกี่ยวข้องกับงานวิจัย

2.1 ทฤษฎีการแจกแจงของประชากร (ประสิทธิ์ พยัคฆพงษ์, 2545; Freund, 1992)

การวิจัยครั้งนี้ได้ทบทวนวรรณกรรมเกี่ยวกับทฤษฎีการแจกแจงที่เกี่ยวข้องกับการประมาณค่าแบบเบย์เซียน ดังนี้

2.1.1 การแจกแจงทวินาม (Binomial Distribution)

การแจกแจงทวินาม เกิดจากการทดลองแบบเบอร์นูลี (Bernoulli) n ครั้ง สรุปการทดลองแบบทวินามได้ดังนี้

(1) การสังเกต มีซ้ำกัน n ครั้ง

(2) ในแต่ละครั้งของการ สังเกต แบ่งออกเป็น 2 ลักษณะ คือ เกิดเหตุการณ์ที่สนใจ (x=1) และ ไม่เกิดเหตุการณ์ที่สนใจ (x=0)

(3) ความน่าจะเป็นของการเกิดเหตุการณ์ที่สนใจเท่ากับ p มีค่าเท่ากันในแต่ละครั้งของการสังเกต

(4) การสังเกต n ครั้งเป็นอิสระต่อกัน

ถ้าให้ X แทนจำนวนครั้งของการเกิดเหตุการณ์ที่สนใจ n ครั้ง ดังนั้น x=0, 1, 2,..., n ส่วนจำนวนเหตุการณ์ที่ไม่ประสบผลสำเร็จ คือ n-x และเนื่องจากจำนวนวิธีที่จะเกิดเหตุการณ์ที่สนใจครั้งที่ x จากการเกิดเหตุการณ์ที่สนใจ x ครั้งในการสังเกต n ครั้ง คือ

$$\binom{n}{x} = \frac{n!}{x!(n-x)!} \quad (1)$$

เนื่องจากการสังเกตในแต่ละครั้งเป็นอิสระจากกัน และความน่าจะเป็นที่แต่ละครั้งของการสังเกตการเกิดเหตุการณ์ที่สนใจ คือ p และไม่เกิดเหตุการณ์ที่สนใจเท่ากับ 1-p ` ดังนั้นความน่าจะเป็นที่จะเกิดเหตุการณ์ที่สนใจ x ครั้ง คือ $\binom{n}{x}p^x(1-p)^{n-x}$

ถ้ากำหนดให้ X เป็นตัวแปรสุ่มที่มีการแจกแจงแบบทวินามฟังก์ชันความน่าจะเป็น (probability function) ของ X คือ

$$f(x) = \binom{n}{x}p^x(1-p)^{n-x} \qquad \text{เมื่อ } x=0,\ 1,\ 2,\ \ldots,\ n \quad (2)$$

ถ้า X มีการแจกแจงแบบทวินาม ค่าเฉลี่ยของความแปรปรวนของ X คือ

$$E(X) = np \qquad \text{และ} \qquad Var(X) = np(1-p) \quad (3)$$

2.1.2 การแจกแจงแบบบีต้า (Beta distribution)

ถ้ากำหนดให้ X เป็นตัวแปรสุ่มมีการแจกแจงบีต้าที่มีพารามิเตอร์ a (scale parameter) และพารามิเตอร์ b (shape parameter) จะได้ฟังก์ชันความน่าจะเป็นของ X คือ

$$f(x) = \frac{\Gamma(a+b)}{\Gamma(a).\Gamma(b)}x^{a-1}(1-x)^{b-1} \qquad \text{เมื่อ } 0 < x < 1 \quad (4)$$

จะได้ว่า

$$\int_0^1 x^{a-1}(1-x)^{b-1}\,dx = \frac{\Gamma(a).\Gamma(b)}{\Gamma(a+b)}$$

ซึ่งอินทิกรัลนี้เรียกว่า บีต้าฟังก์ชัน (beta function) และแทนด้วย $B(a,b)$

โดยที่

$$B(a,b) = \frac{(a-1)!(b-1)!}{(a+b-2)!}$$ เมื่อ a, b เป็นจำนวนเต็ม

ดังนั้น

$$f(x) = \frac{x^{a-1}(1-x)^{b-1}}{B(a,b)} \quad \text{เมื่อ } 0 < x < 1 \qquad (5)$$

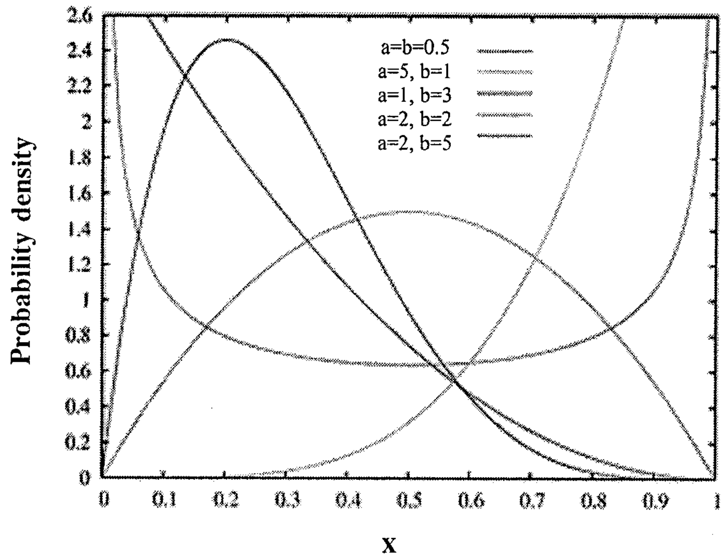
ค่าเฉลี่ยและความแปรปรวนของการแจกแจงบีต้า คือ

$$E(X) = \frac{a}{a+b} \qquad (6)$$

$$Var(X) = \frac{ab}{(a+b)^2(a+b+1)} \qquad (7)$$

ค่าฐานนิยมของการแจกแจงบีต้า (Xycoon, 2006)

ค่าฐานนิยม	ค่าพารามิเตอร์ของบีต้า
$M_0 = \dfrac{a-1}{a+b-2}$	a>1, b>1
M_0 = 0 และ 1	a<1, b<1
M_0 = 0	a=1, b>1
M_0 = 1	a$\geq$1, b<1
	a>1, b=1



ภาพที่ 3 ลักษณะการแจกแจงแบบบีต้า (Wikipedia, 2007)

2.2 ความน่าจะเป็นแบบมีเงื่อนไข (รัชกมล กบิลจิตต์, 2545)

ความน่าจะเป็นของเหตุการณ์ A เมื่อทราบว่าเหตุการณ์ B ได้เกิดขึ้นแล้ว เรียกว่า ความน่าจะเป็นแบบมีเงื่อนไข โดยแทนความน่าจะเป็นของเหตุการณ์ A ภายใต้เงื่อนไขของการเกิดเหตุการณ์ B ด้วย P(A|B) (conditional probability of A given B) และนิยาม P(A|B) ดังนี้

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \quad ; \quad P(B) \neq 0$$

ทำนองเดียวกัน หาก P(B|A) เป็นความน่าจะเป็นแบบมีเงื่อนไขของเหตุการณ์ B เมื่อกำหนดว่าเหตุการณ์ A ได้เกิดขึ้นแล้วจะได้

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} \quad ; \quad P(A) \neq 0$$

หากพิจารณาเหตุการณ์ $A \cap B$ และ B จะเห็นได้ว่า $A \cap B \subseteq B$ และ $P(A \cap B) \leq P(B)$ ดังนั้นผลที่ได้ตามมาคือ

1) $0 \leq P(A \mid B) \leq 1$

2) P(S|B) = 1, P(ϕ|B) = 0 และ P(B|B) = 1

3) ถ้า A_1, A_2, ... เป็นเหตุการณ์ที่ไม่มีสมาชิกร่วมกัน (disjoint) และ B เป็นสับเซตของ Union ของ A_1, A_2, ... แล้ว $P(A_1 \cup A_2 \cup \ldots \cup A_k \mid B) = P(A_1 \mid B) + P(A_2 \mid B) + \ldots$

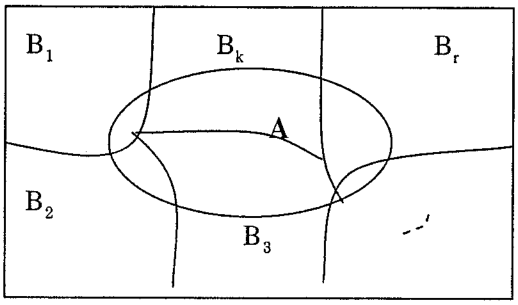
2.3 ทฤษฎีบทของเบย์ (Bayes' Theorem)

ทฤษฎีบทของเบย์เป็นทฤษฎีสำหรับหาค่าความน่าจะเป็นแบบมีเงื่อนไขของเหตุการณ์ที่สนใจจากเหตุการณ์ที่เป็นเงื่อนไข

เมื่อกำหนดให้ $B_1, B_2, ..., B_k$ เป็นเหตุการณ์ที่ไม่เกิดขึ้นร่วมกัน (Mutually Exclusive event) และครอบคลุมทั่ว Sample space (Exhaustive) เมื่อกำหนดให้ A เป็นเหตุการณ์ใด ๆ ใน Sample space (S)

$B_i \cap B_j = \phi$ เมื่อ $i \neq j$

$B_1 \cup B_2 \cup ... \cup B_k =$ Sample Space (S)



จะได้ความน่าจะเป็นแบบมีเงื่อนไขของทฤษฎีบทของเบย์ ดังนี้ (Freund, 1992: p.70-72)

$$P(B_r \mid A) = \frac{P(B_r) \cdot P(A \mid B_r)}{\sum_{i=1}^{k} P(B_i) \cdot P(A \mid B_i)} \tag{8}$$

ซึ่งทฤษฎีความน่าจะเป็นแบบมีเงื่อนไขหรือทฤษฎีบทของเบย์สามารถนำไปประยุกต์ใช้ในการหาการแจกแจงของตัวแปรได้

ถ้าให้ $B_1, B_2, ..., B_n$ เป็นเหตุการณ์ที่ถูกสังเกต n เหตุการณ์ โดยที่เหตุการณ์นี้ไม่เกิดร่วมกัน ซึ่งมีการแจกแจงความน่าจะเป็นคือ P(B|A) ที่ขึ้นอยู่กับค่าพารามิเตอร์ A และมีค่าความน่าจะเป็นคือ P(A) แล้วจะได้ความน่าจะเป็นแบบมีเงื่อนไขของ A เมื่อ P(A) > 0 คือ

$$P(B_i \mid A) = \frac{P(A \mid B_i) P(B_i)}{P(A)} \tag{9}$$

เมื่อ

$$P(A) = \sum_{i=1}^{k} P(B_i) \cdot P(A \mid B_i)$$

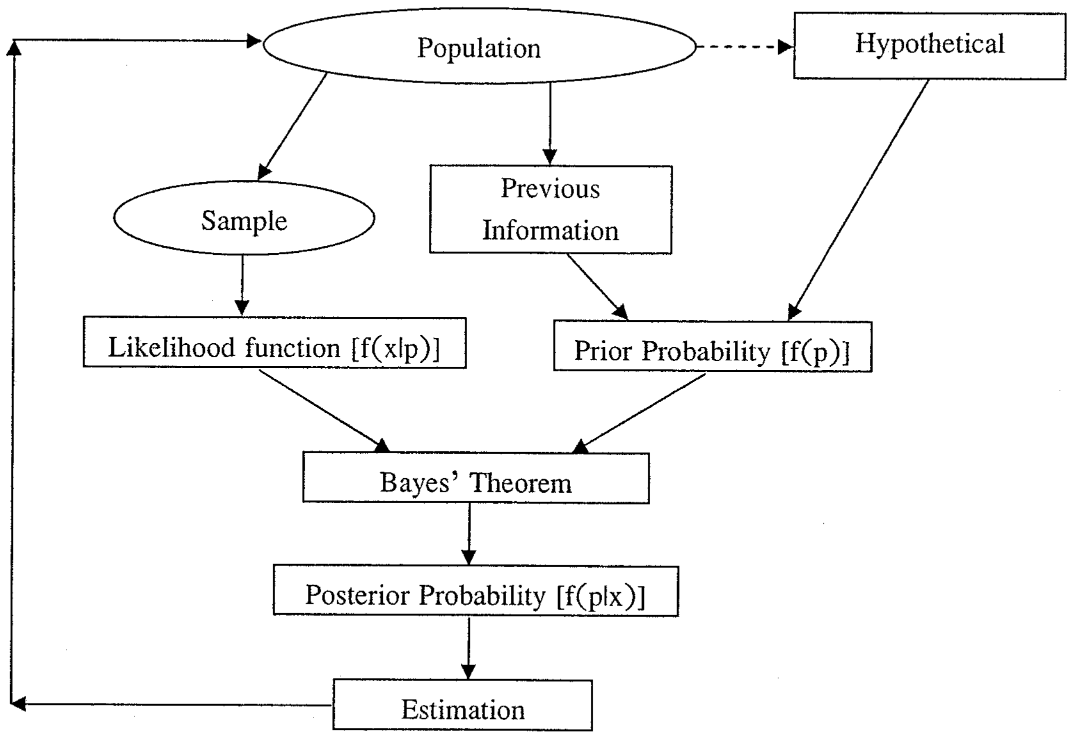
จากสมการ (9) $P(B_i|A)$ คือ การแจกแจงภายหลัง (Posterior Distribution) ของเหตุการณ์ B_i และ $P(B_i)$ คือ การแจกแจงเบื้องต้น (Prior Distribution) ของเหตุการณ์ B_i ก่อนที่ข้อมูลจะถูกสังเกตหรือทดลอง $P(A|B_i)$ คือ likelihood function

กล่าวโดยสรุป แนวความคิดการประมาณค่าแบบเบย์เซียนนี้จะเป็นการอาศัยข้อมูลที่ได้จากการสังเกตมาเปลี่ยนแปลงความเชื่อดั้งเดิมเกี่ยวกับค่าที่แตกต่างกันของ p โดยการเปลี่ยนฟังก์ชันความน่าจะเป็นเบื้องต้นของ p ไปสู่ฟังก์ชันความน่าจะเป็นภายหลังของ p

$$\text{Posterior Probability} = \frac{\text{Prior Probability x Likelihood}}{\text{Marginal Distribution of the data}}$$

2.4 วิธีการและการอ้างอิงทางสถิติแบบเบย์เซียน (Bayesian process and inference)

ในกระบวนการอ้างอิงทางสถิติแบบเบย์เซียนจะถือว่าค่าพารามิเตอร์ของประชากรนั้นเป็นตัวแปร (variable) ซึ่งสถิติแบบดั้งเดิมถือว่าค่าพารามิเตอร์เป็นค่าคงที่ (constant) ดังนั้นในสถิติแบบเบย์เซียนค่าพารามิเตอร์จะมีการแจกแจง (Distribution) และข้อมูลเบื้องต้นที่มีอยู่ หรือความเชื่อของบุคคลที่มีเกี่ยวกับการแจกแจงของพารามิเตอร์ดังกล่าวเรียกว่า การแจกแจงเบื้องต้น (Prior Distribution) เมื่อทำการศึกษาจะทำการสุ่มตัวอย่างมาเพื่อศึกษาคุณลักษณะของประชากรที่ต้องการ อาศัยทฤษฎีเบย์รวมข้อมูลเบื้องต้นและข้อมูลที่ได้จากตัวอย่างเข้าด้วยกันจะได้การแจกแจงที่เรียกว่า การแจกแจงภายหลัง (Posterior Distribution) แล้วทำการประมาณ (estimation) ค่าพารามิเตอร์จากการแจกแจงดังกล่าว ดังภาพที่ 4



ภาพที่ 4 การอ้างอิงทางสถิติแบบเบย์เซียน

2.5 การกำหนดขนาดตัวอย่างแบบเบย์เซียน (Bayesian sample size determination) (Pham-Gia และ Turkkan, 1992; Joseph et al., 1995a)

การกำหนดขนาดตัวอย่างแบบเบย์เซียนสามารถแบ่งออกเป็น 2 แนวคิด (Adcock, 1997) ได้แก่ การคำนวณขนาดตัวอย่างแบบเบย์เซียนที่คำนึงถึงฟังก์ชันอรรถประโยชน์ โดยคำนึงถึงต้นทุนหรือค่าใช้จ่ายในการสุ่มตัวอย่าง ซึ่งในงานวิจัยที่ต้องใช้ขนาดตัวอย่างในการทดสอบสมมติฐานจะไม่สามารถกำหนดขนาดตัวอย่างด้วยวิธีนี้ได้ เนื่องจากวิธีที่คำนึงถึงฟังก์ชันอรรถประโยชน์จะไม่สามารถกำหนดระดับนัยสำคัญ α (fixed significance level α) สำหรับรายละเอียดผู้สนใจสามารถศึกษาได้ใน Lindley (1997) แนวคิดที่สองคือการคำนวณขนาด

ตัวอย่างแบบเบย์เซียนที่คำนึงถึงการอนุมานเกี่ยวกับค่าพารามิเตอร์ p ซึ่งในที่นี้จะนำเสนอวิธีการกำหนดขนาดตัวอย่างแบบเบย์เซียน 5 วิธี ดังนี้

2.5.1 วิธีครอบคลุมโดยเฉลี่ย (Average Coverage Criterion; ACC)

เมื่อกำหนดค่าความกว้างของความคลาดเคลื่อน (Length; L) ช่วงความน่าจะเป็นภายหลังสูงสุด (highest posterior density: HPD) การหาขนาดตัวอย่างแบบเบย์เซียนที่น้อยที่สุดที่จะทำให้ค่าความน่าจะเป็นครอบคลุมโดยเฉลี่ยอย่างน้อย $1-\alpha$ คือ

$$\int_{x}\left\{\int_{a(x,n)}^{a(x,n)+L} f(p \mid x,n)dp\right\}f(x)dx \geq 1-\alpha \tag{10}$$

เมื่อ f(x) คือ preposterior marginal distribution of x โดยที่

$$f(x) = \int_{\Theta} f(x \mid p)f(p)dp \tag{11}$$

และ $f(p \mid x,n)$ คือ การแจกแจงภายหลังของ p เมื่อกำหนด x ซึ่งมีค่า

$$f(p \mid x,n) = \frac{f(x \mid p)f(p)}{\int_{\Theta} f(x \mid p)f(p)dp} \tag{12}$$

$f(x \mid p)$ คือ ฟังก์ชันภาวะน่าจะเป็น (likelihood) และ a(x,n) คือ ลิมิตขอบล่างของ HPD ที่มีความกว้างขนาด L สำหรับ posterior density $f(p \mid x,n)$

เมื่อผลลัพธ์ (x) มีลักษณะการแจกแจงแบบทวินามที่มีพารามิเตอร์ n และ p แล้ว การหาขนาดตัวอย่างด้วยวิธีครอบคลุมโดยเฉลี่ย ขนาดตัวอย่างที่น้อยที่สุดหาได้จาก

ให้ a(x,n) คือ ลิมิตขอบล่างของ HPD เมื่อกำหนดขนาดตัวอย่าง (n) และค่าสังเกตของจำนวนครั้งของความสำเร็จ (x)

L คือ ความกว้างที่กำหนดให้ล่วงหน้า

กำหนด f(p) คือ ฟังก์ชันความหนาแน่นของการแจกแจงเบื้องต้นของ p โดยที่มีการแจกแจงเบื้องต้นแบบบีต้าที่มีพารามิเตอร์ B(a, b) โดยที่

$$f(p) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}p^{a-1}(1-p)^{b-1}, \qquad 0<p<1 \tag{13}$$

ดังนั้น การแจกแจงภายหลังของ P เมื่อกำหนด x คือ

$$f(p \mid x,n,a,b) = \frac{\Gamma(n+a+b)}{\Gamma(x+a)\Gamma(n-x+b)}p^{x+a-1}(1-p)^{n-x+b-1} \tag{14}$$

และ a(x,n) คือ preposterior marginal distribution of x โดยที่

$$a(x,n) = \binom{n}{x}B(x+a,n-x+b)/B(a,b) \tag{15}$$

ดังนั้น การหาขนาดตัวอย่างที่มีค่าน้อยที่สุดด้วยวิธีความครอบคลุมโดยเฉลี่ยหาได้จาก a(x,n) ที่สอดคล้องกับสมการดังนี้

$$\sum_{x=0}^{n}\left(\binom{n}{x}/B(a,b)\right)\int_{a(x,n)}^{a(x,n)+L} p^{(x+a-1)}(1-p)^{n-x+b-1)}dp \geq 1-\alpha \tag{16}$$

2.5.2 วิธีความกว้างโดยเฉลี่ย (Average Length Criterion; ALC)

เป็นการกำหนดค่าความน่าจะเป็นที่ครอบคลุม $1-\alpha$ ไว้แน่นอน โดยที่แต่ละค่าที่เป็นไปได้ของ x จะต้องการความกว้าง $L'(x,n)$ ที่ครอบคลุมความน่าจะเป็นตามที่ต้องการ

$$\int_{a(x,n)}^{a(x,n)+L'(x,n)} f(p \mid x,n)dp = 1-\alpha \qquad (17)$$

โดยที่ความกว้างโดยเฉลี่ยที่น้อยกว่าหรือเท่ากับ L ต้องการตัวอย่างที่น้อยที่สุด

$$\sum_{x=0}^{n} L'(x,n)a(x,n) \le L \qquad (18)$$

เมื่อ a(x,n) คือ preposterior marginal distribution of x เมื่อกำหนดความกว้าง $L'(x,n)$ จะได้ว่า

$$\int_{a(x,n)}^{a(x,n)+L'(x,n)} f(p \mid x,n,a,b)d\theta = 1-\alpha \qquad (19)$$

เมื่อ $f(p \mid x,n,a,b)$ คือ การแจกแจงภายหลังของ p ดังสมการ (14)

a(x,n) และ $a(x+n)+L'(x,n)$ คือ ลิมิตล่างและบนของช่วง HPD

2.5.3 วิธีผลลัพธ์แย่ที่สุด (Worst Outcome Criterion; WOC)

เป็นการหาขนาดตัวอย่างที่ทำให้ได้ค่าครอบคลุมไม่น้อยกว่า $1-\alpha$ และมีค่าความกว้างไม่มากกว่า L ที่กำหนดไว้ ดังนั้น

$$\inf_{x\in X}\left\{\int_{a(x,n)}^{a(x,n)+L} f(p \mid x,n,a,b)dp\right\} \ge 1-\alpha \qquad (20)$$

เมื่อ $f(p \mid x,n,a,b)$ คือ การแจกแจงภายหลังของ p ดังสมการ (14)

สำหรับค่าใด ๆ ของ n, a, b, L ค่าของอินทิกรัล คือ

$$\int_{a(x,n)}^{a(x,n)+L} f(p \mid x,n,a,b)dp$$

จะมีค่าน้อยที่สุด สำหรับ $x \in (0,1,2,\ldots,n)$ โดยที่ $x^* = x^*(n,c,d)$ (21)

$$= \begin{cases} \dfrac{n+a+b+1}{2}-a \quad \text{หรือ} \quad \dfrac{n+a+b-1}{2}-a & \text{ถ้า } n+a+b \text{ เป็นเลขคี่ และ } n \ge |b-a|, \\ \dfrac{n+a+b+1}{2}-a & \text{ถ้า } n+a+b \text{ เป็นเลขคู่ และ } n \ge |b-a|, \\ n & \text{ถ้า } 0 \le n \le |b-a| \end{cases}$$

สมมติให้การเลือก x* ถูกต้องจากสมการ (20) ขนาดตัวอย่างอย่างน้อยที่คำนวณด้วยวิธีผลลัพธ์แย่ที่สุดได้จาก

$$\int_{a(x^*,n)}^{a(x^*,n)+L} f(p \mid x^*,n,c,d)dp \ge 1-\alpha \qquad (22)$$

2.5.4 วิธีความแปรปรวนภายหลังสูงสุด (Maximum posterior variance criterion)

ในกรณีเมื่อความแปรปรวนภายหลังมีค่าลดต่ำลงเสมอ หรือเมื่อ n>b-a สามารถกำหนดค่าความกว้างแล้วหาค่าความแปรปรวนภายหลัง ณ ความกว้างที่กำหนดได้จากความสัมพันธ์ระหว่างความกว้างและความแปรปรวนภายหลัง เมื่อความแปรปรวนภายหลังมีค่าสูงสุด $(\max_{0\le x\le n}\lfloor \mathrm{var}_{post}(n)\rfloor)$ ดังนี้

$$L = 2Z_{\alpha/2}\sqrt{W} \qquad \text{หรือ} \qquad W = \sigma^2_{post} = (L/[2Z_{\alpha/2}])^2$$

จึงหาค่าขนาดตัวอย่างได้จาก

$$n_{pham\,max} = [4W]^{-1} - (a+b+1)\,; \quad n > b-a \tag{23}$$

2.5.5 วิธีความแปรปรวนภายหลังโดยเฉลี่ย (Average posterior variance criterion)

เป็นการหาขนาดตัวอย่างจากความแปรปรวนเบื้องต้นและความแปรปรวนภายหลัง เมื่อกำหนดค่าความกว้างและระดับนัยสำคัญ สามารถหาขนาดตัวอย่างได้จาก

$$n_{Pham\,aver} = (a+b)([\sigma^2_{prior}/\sigma^2_{post}]-1) \tag{24}$$

เมื่อ $U = \sigma^2_{prior} = ab/[(a+b)^2(a+b+1)] \quad W = \sigma^2_{post} = (L/[2Z_{\alpha/2}])^2$

ในการกำหนดขนาดตัวอย่างแบบเบย์เซียน จีรภา สิมะจารึก (2545) ได้สรุปว่าปัจจัยที่มีอิทธิพลต่อการกำหนดขนาดตัวอย่างจะประกอบไปด้วยค่าพารามิเตอร์ (a, b) ของการแจกแจงความน่าจะเป็นเบื้องต้น ความกว้างของความคลาดเคลื่อน และระดับนัยสำคัญ โดยค่าพารามิเตอร์ของการแจกแจงความน่าจะเป็นเบื้องต้นเป็นปัจจัยที่มีอิทธิพลมากที่สุดในการกำหนดขนาดตัวอย่างแบบเบย์เซียน จากการทบทวนวรรณกรรมที่เกี่ยวข้อง ดังตารางที่ 3 พบว่าการกำหนดขนาดตัวอย่างกรณีที่ผลลัพธ์มีโอกาสเกิดน้อยด้วยวิธีความกว้างโดยเฉลี่ยจะให้ขนาดตัวอย่างที่เล็กสุด (Joseph et al., 1995a; Dendukuri et al., 2004) ซึ่งการศึกษาส่วนใหญ่จะกำหนดความน่าจะเป็นเบื้องต้นแบบยูนิฟอร์มจึงทำให้ตัวอย่างที่ได้มีขนาดเล็ก แต่จากการศึกษาของ จีรภา สิมะจารึก (2545) พบว่าวิธีความแปรปรวนภายหลังโดยเฉลี่ยจะให้ขนาดตัวอย่างที่เล็กกว่าวิธีความกว้างโดยเฉลี่ยเมื่อผลลัพธ์มีโอกาสเกิดน้อย อย่างไรก็ตามการกำหนดขนาดตัวอย่างด้วยวิธีความกว้างโดยเฉลี่ยที่ใช้ความน่าจะเป็นเบื้องต้นแบบยูนิฟอร์ม และวิธีความกว้างโดยเฉลี่ยที่อาศัยแนวคิดการแจกแจงปกติ เป็นวิธีการกำหนดขนาดตัวอย่างที่ไม่เหมาะสมสำหรับกรณีที่ผลลัพธ์ในตัวอย่างเป็นศูนย์

ดังนั้น ในงานวิจัยนี้จึงไม่สามารถที่จะคำนวณขนาดตัวอย่างจากวิธีที่นำเสนอได้ แต่จะทำการสร้างขนาดตัวอย่างแล้วทำการประมาณค่าขีดจำกัดบนช่วงเชื่อถือเบย์เซียน จนกว่าค่าขีดจำกัดบนช่วงเชื่อถือเบย์เซียนจะมีค่าน้อยกว่าหรือเท่ากับเกณฑ์คุณภาพที่กำหนด โดยจะถือว่าขนาดตัวอย่าง ณ จุดนั้นเป็นขนาดตัวอย่างที่เหมาะสม

ตารางที่ 3 การทบทวนวรรณกรรมที่เกี่ยวข้องกับการกำหนดขนาดตัวอย่างด้วยวิธีเบย์เซียน (ต่อ)

ชื่อผู้แต่ง	ปี	การศึกษา	ผลการศึกษา	สรุปข้อคิดเห็นที่ได้
Lindley, D.V.	1997	- ศึกษาเกี่ยวกับการกำหนดขนาดตัวอย่างแบบเบย์เซียนที่คำนึงถึงฟังก์ชันอรรถประโยชน์ (utility function) ซึ่งเรียกว่า maximization of expected utility (MEU) โดยทำการเปรียบเทียบกับการกำหนดขนาดตัวอย่างด้วยวิธี ALC ข้อแตกต่างระหว่าง MEU และ ALC 1. MEU คำนึงถึงต้นทุน/ค่าใช้จ่ายในการสุ่มตัวอย่าง 2. ALC ไม่มีความชัดเจนของการได้มาซึ่งค่าพารามิเตอร์ว่าขึ้นอยู่กับช่วงเชื่อมั่น (interval) หรือไม่ ในขณะที่ MEU มีความชัดเจนว่าขึ้นอยู่กับอรรถประโยชน์	- พบว่าการกำหนดขนาดตัวอย่างด้วย MEU และ ALC ให้ขนาดตัวอย่างที่ใกล้เคียงกัน โดย MEU>ALC เพียงเล็กน้อย โดยการกำหนดต้นทุนที่ต่ำจะทำให้ขนาดตัวอย่างเพิ่มขึ้น ในทางตรงกันข้ามหากต้นทุนสูงจะทำให้ขนาดตัวอย่างเล็กลง	- การกำหนดขนาดตัวอย่างด้วย MEU เป็นทางเลือกอีกทางในสำหรับงานวิจัยที่ต้องคำนึงถึงต้นทุนในงานวิจัย ซึ่งจะเห็นได้ว่าให้ขนาดตัวอย่างที่ไม่แตกต่างจากการกำหนดด้วยวิธี ALC แต่ในทางปฏิบัติแล้วบางครั้งไม่สามารถที่จะคำนวณขนาดตัวอย่างด้วย MEU ได้ ทั้งนี้เนื่องจาก MEU จะไม่สามารถกำหนดระดับนัยสำคัญ α ดังนั้นในกรณีที่กำหนดขนาดตัวอย่างเพื่อการตัดสินใจที่จะต้องคำนึงถึงการยอมรับ หรือปฏิเสธสมมติฐานจะไม่สามารถใช้การกำหนดขนาดตัวอย่างด้วย MEU ได้ (Pham-Gia, T., 1997)

ตารางที่ 3 การทบทวนวรรณกรรมที่เกี่ยวข้องกับการกำหนดขนาดตัวอย่างด้วยวิธีเบย์เซียน (ต่อ)

ชื่อผู้แต่ง	ปี	การศึกษา	ผลการศึกษา	สรุปข้อคิดเห็นที่ได้
Rahme, E., Joseph L., and Gyorkos, T.W.	2000	- ศึกษาการกำหนดขนาดตัวอย่างแบบเบย์เซียนสำหรับการประมาณค่าพารามิเตอร์ของทวินามใน misclassify data โดยการกำหนดขนาดตัวอย่างด้วยวิธี ACC สำหรับการประมาณค่าความไว (sensitivity; S) และความจำเพาะ (specificity; C)	- เมื่อกำหนดให้ S และ C เท่ากับ 0.9 ขนาดตัวอย่างเท่ากับ 1473 จะทำให้มีความครอบคลุมโดยเฉลี่ย 0.95 เมื่อความกว้างเท่ากับ 0.4 เมื่อกำหนดให้ขนาดตัวอย่างคงที่ และเปลี่ยนค่า S และ C จะพบว่าเมื่อ S และ C มีค่าลดลงจะส่งผลให้ความครอบคลุมโดยเฉลี่ยมีค่าลดลงเช่นเดียวกัน เช่น S~B (0.85, 0.95) C~B(0.85, 0.95) จะได้ coverage=0.589 ซึ่งจากการศึกษาพบว่าการกำหนดความจำเพาะจะมีผลกระทบต่อความครอบคลุมมากกว่าการกำหนดค่าความไว	- เมื่อใช้ขนาดตัวอย่างเท่าเดิมที่ทำให้ความครอบคลุมโดยเฉลี่ยตามที่ต้องการ หากมีการเปลี่ยนแปลงค่าความไวและความจำเพาะพบว่าความครอบคลุมโดยเฉลี่ยเปลี่ยนแปลงไปเช่นเดียวกัน ซึ่งการประมาณค่าความชุกจึงขึ้นอยู่กับการกำหนดค่าความไวและความจำเพาะ และการเพิ่มขนาดตัวอย่างให้ใหญ่ขึ้นไม่ได้หมายความว่าจะเป็นการเพิ่มความถูกต้องในการประมาณค่าพารามิเตอร์เสมอไป - วิธีการที่ศึกษาจะกำหนดช่วง HPD คำนวณได้จาก $\hat{\theta} \pm d$ ซึ่งถ้าหากนำมาใช้ในกรณีผลลัพธ์มีโอกาสเกิดน้อยจะเป็นการกำหนดช่วงความน่าจะเป็นภายหลังที่ไม่เหมาะสม

21

ตารางที่ 3 การทบทวนวรรณกรรมที่เกี่ยวข้องกับการกำหนดขนาดตัวอย่างด้วยวิธีเบย์เซียน (ต่อ)

ชื่อผู้แต่ง	ปี	การศึกษา	ผลการศึกษา	สรุปข้อคิดเห็นที่ได้
Dendukuri, N., Rahme, E., Bélisle, Patrick., and Joseph, L.	2004	- ศึกษาการกำหนดขนาดตัวอย่างแบบเบย์เซียนสำหรับการศึกษาความชุก (Prevalence) และการวินิจฉัยโรค (Diagnostic Test) เมื่อไม่มี gold standard test โดยได้ทำการศึกษาเปรียบเทียบการกำหนดขนาดตัวอย่าง 3 วิธี ได้แก่ ACC, ALC และ MWOC (modified WOC)	- พบว่าขนาดตัวอย่างสำหรับการประมาณค่าความชุกแบบ single test เมื่อกำหนด l=0.2 $1-\alpha=0.95$ ตัวอย่างจะมีขนาดใหญ่ (finite sample size) ทั้ง 3 วิธี - ขนาดตัวอย่างสำหรับค่าความไว (sensitivity;S) และความจำเพาะ (specificity;C) โดยการกำหนด gold standard (S=C=1) และ non-gold standard S~B(155.63,66.15) C~B(906.98,3.63) พบว่า ขนาดตัวอย่างที่คำนวณจากการกำหนด gold standard ทั้ง S และ C จะได้ขนาดตัวอย่างเท่ากัน ซึ่งการกำหนด gold standard ด้วย Uniform prior จะให้ขนาดตัวอย่างเล็กกว่า non-gold standard ทั้ง 3 วิธี และเมื่อเปรียบเทียบทั้ง 3 วิธี พบว่า ALC จะให้ขนาดตัวอย่างน้อยที่สุด รองลงมาคือ ACC และ MWOC ตามลำดับ	- ถึงแม้ว่าการกำหนดขนาดตัวอย่าง โดยการใช้ Uniform prior เป็น gold standard จะให้ขนาดตัวอย่างที่เล็กกว่า non-gold standard แต่ในทางปฏิบัติการเกิดความชุก หรือการวินิจฉัยโรคใหม่ๆ อาจไม่สามารถที่จะกำหนดให้เป็นค่าเดียวกันได้ ทั้งนี้ขึ้นอยู่กับนักวิจัยที่จะมีข้อมูลสำหรับการกำหนดความน่าจะเป็นเบื้องต้นที่แตกต่างออกไป ซึ่งอาจจะเลือกให้เหมาะสมสำหรับแต่ละ case - เมื่อกำหนดค่าพารามิเตอร์ของบีต้าที่แตกต่างกันมากจะพบว่าการกำหนดขนาดตัวอย่างด้วยวิธี ALC จะให้ขนาดตัวอย่างน้อยที่สุด เช่นเดียวกับการศึกษาของ Joseph et al. (1995a)

ตารางที่ 3 การทบทวนวรรณกรรมที่เกี่ยวข้องกับการกำหนดขนาดตัวอย่างด้วยวิธีเบย์เซียน (ต่อ)

ชื่อผู้แต่ง	ปี	การศึกษา	ผลการศึกษา	สรุปข้อคิดเห็นที่ได้
จีรภา สิมะจารึก	2545	- ศึกษาการกำหนดขนาดตัวอย่างสำหรับการประมาณค่าสัดส่วนตามแนวคิดฟรีเควนทิสและเบย์เซียน 5 วิธี คือ Average coverage criterion (ACC), Average length criterion (ALC), Worst outcome criterion (WOC), Maximum posterior variance criteria ($Pham_{max}$) และ Average posterior variance criteria ($Pham_{aver}$)	- คุณสมบัติของการกำหนดขนาดตัวอย่างแต่ละวิธี พบว่า 1. ขนาดตัวอย่างที่ได้จากทุกวิธีเมื่อกำหนด B(a,b) จะให้ขนาดตัวอย่างเท่ากับ B(b,a) 2. ขนาดตัวอย่างจะแปรผันตามระดับความเชื่อมั่นในทิศทางเดียวกัน 3. ขนาดตัวอย่างจะแปรผกผันกับความกว้างของความคลาดเคลื่อน 4. ขนาดตัวอย่างที่ได้จากวิธี ALC จะแปรผันตามค่าพารามิเตอร์ของการแจกแจงเบื้องต้นในทิศทางเดียวกัน เมื่อ a หรือ b เพิ่มขึ้นจะให้ขนาดตัวอย่างมากขึ้น ส่วนวิธีอื่น ๆที่เหลือ คือ ฟรีเควนทิส, ACC, WOC, $Pham_{max}$, $Pham_{aver}$ ขนาดตัวอย่างจะแปรผกผันกับค่าพารามิเตอร์ของการแจกแจงเบื้องต้น โดยเมื่อ a หรือ b เพิ่มขึ้น ขนาดตัวอย่างจะลดลง	- ปัจจัยที่มีอิทธิพลต่อการกำหนดขนาดตัวอย่าง ประกอบด้วย ค่าพารามิเตอร์ (a,b) ของการแจกแจงเบื้องต้น ความกว้างของความคลาดเคลื่อนและระดับนัยสำคัญ โดยปัจจัยที่มีอิทธิพลมากที่สุดคือ ค่าพารามิเตอร์ของการแจกแจงเบื้องต้น - ผลการศึกษาเกี่ยวกับการกำหนดขนาดตัวอย่างด้วยวิธี ALC จะให้ขนาดตัวอย่างเพิ่มมากขึ้น เมื่อค่าพารามิเตอร์ของการแจกแจงเบื้องต้น a หรือ b มีค่าเพิ่มขึ้น ซึ่งขัดแย้งกับการศึกษาของ Joseph et al. (2005a และ 2005b) ที่กล่าวไว้ว่าวิธี ALC จะให้ขนาดตัวอย่างที่เล็กลงเมื่อ a และ b มีค่าแตกต่างกันมาก

23

ตารางที่ 3 การทบทวนวรรณกรรมที่เกี่ยวข้องกับการกำหนดขนาดตัวอย่างด้วยวิธีเบย์เซียน

ชื่อผู้แต่ง	ปี	การศึกษา	ผลการศึกษา	สรุปข้อคิดเห็นที่ได้
Joseph, L., Wolfson, D.B., and du Berger, R.	1995a	- ศึกษาเกี่ยวกับการกำหนดขนาดตัวอย่างของค่าสัดส่วนทวินามด้วยวิธีเบย์เซียน โดยอาศัยแนวคิด Highest posterior density (HPD) interval ซึ่งได้ทำการศึกษา 3 วิธี ได้แก่ Average coverage criterion (ACC), Average length criterion (ALC) และ Worst outcome criterion (WOC) - เปรียบเทียบวิธีการกำหนดขนาดตัวอย่าง 3 วิธีกับวิธีของ Adcock (1992) n_{CONS} ซึ่งอาศัยแนวคิด tolerance region และ Pham-Gia &Turkkan (1992) n_{VAR} ซึ่งอาศัยแนวคิด average posterior variance criterion	- ในกรณีที่ผลลัพธ์เกิดน้อย เมื่อกำหนด a=1 b=200 ที่ระดับความเชื่อมั่น 90% l=0.01 เมื่อ p และ SD. เท่ากับ 0.005 พบว่า n_{ACC}=164 n_{ALC}=96 และ n_{WOC}=26,854 ถ้าคำนวณขนาดตัวอย่างโดยอาศัยแนวคิด n_{VAR} จะพบว่าขนาดตัวอย่างเท่ากับ 332 - เมื่อกำหนดความน่าจะเป็นเบื้องต้นให้มีการแจกแจงแบบยูนิฟอร์ม [Uniform prior: B(1,1)] พบว่าเมื่อ $1-\alpha \geq 0.90$ จะพบว่า $n_{ALC} \leq n_{VAR} \leq n_{ACC} \leq n_{WOC} \approx n_{CONS}$	- ในกรณีที่ผลลัพธ์เกิดน้อย การคำนวณขนาดตัวอย่างโดยอาศัยแนวคิด average posterior variance criterion ไม่เหมาะสม เนื่องจากการกำหนดค่า l (ความกว้างของความคลาดเคลื่อน) มีแนวคิดมาจากการแจกแจงปกติ โดยที่ $l = 2Z_{\alpha/2}\sqrt{W}$ เมื่อ W คือความแปรปรวนภายหลัง ซึ่งในผลลัพธ์ที่เกิดน้อยจะมีลักษณะการแจกแจงเบ้ขวาจึงไม่เหมาะสมหากคำนวณขนาดตัวอย่างโดยอาศัยแนวคิดของการแจกแจงปกติ - การคำนวณขนาดตัวอย่างที่กำหนดให้ l>p ย่อมส่งผลให้ n มีขนาดเล็ก ซึ่งในการประมาณค่า ต้องการค่า l ที่มีค่าเล็ก เพื่อเพิ่มความแม่นยำในการประมาณค่า ดังนั้นการกำหนดให้ l>p จึงไม่เหมาะสม - ในกรณีที่ a และ b มีความแตกต่างกันมาก การกำหนดขนาดตัวอย่างด้วยวิธี ALC จะให้ขนาดตัวอย่างน้อยที่สุด

ตารางที่ 3 การทบทวนวรรณกรรมที่เกี่ยวข้องกับการกำหนดขนาดตัวอย่างด้วยวิธีเบย์เซียน (ต่อ)

ชื่อผู้แต่ง	ปี	การศึกษา	ผลการศึกษา	สรุปข้อคิดเห็นที่ได้
			- ผลการเปรียบเทียบขนาดตัวอย่าง 1. การกำหนดขนาดตัวอย่างด้วยวิธีเบย์เซียนให้ขนาดตัวอย่างน้อยกว่าฟรีเควนทิสในทุกสถานการณ์ 2. วิธี $Pham_{aver}$ ให้ขนาดตัวอย่างน้อยที่สุดเมื่อกำหนดให้ B(0.5,0.5), B(0.5,1), B(0.5,2), B(0.5,3), B(1,0.5), B(1,3), B(1,5), B(1,10), B(1,50), B(2,0.5), B(3,0.5), B(3,1), B(5,1), B(10,1), B(10,50), (50,10) และ B(50,10) ทุกระดับความกว้างของความคลาดเคลื่อน 3. วิธี ALC ให้ขนาดตัวอย่างน้อยสุด เมื่อกำหนด B(1,1), B(1,2), B(2,1), B(2,2), B(2,3), B(3,2), B(3,3), B(5,10) และ B(10,5) ทุกระดับความกว้างของความคลาดเคลื่อน 4. วิธี $Pham_{max}$ จะให้ขนาดตัวอย่างใกล้เคียงกับวิธี WOC ทุกระดับความกว้างและลักษณะการแจกแจงเบื้องต้น	ทั้งนี้เนื่องมาจากการศึกษานี้ได้มีการกำหนด HPD ให้มีค่าเท่ากับ $U \pm Z_{\alpha/2}\sqrt{W}$ โดยที่ U คือ ค่าเฉลี่ยภายหลัง W คือ ความแปรปรวนภายหลัง ซึ่งอาศัยแนวคิดของการแจกแจงปกติ ในขณะที่การศึกษาของ Joseph คำนวณด้วยวิธี Exact จึงให้ผลการศึกษาไม่สอดคล้องกัน

2.6 การกำหนดขนาดตัวอย่างของการสุ่มตัวอย่างเพื่อการยอมรับ (Lwanga และ Lemeshow, 1991)

การสุ่มตัวอย่างเพื่อการยอมรับ (Lot acceptance sampling) เป็นสถิติอนุมานแบบทดสอบสมมติฐานทางเดียว (one-sided test) เพื่อตัดสินใจยอมรับ หรือปฏิเสธผลลัพธ์ของตัวอย่างที่นำมาตรวจสอบว่าได้คุณภาพตามมาตรฐานที่กำหนดไว้หรือไม่ ซึ่งก่อนการทดสอบสมมติฐานผู้วิจัยจะต้องมีการขนาดตัวอย่างที่จะใช้ในการสุ่มเพื่อตรวจสอบ โดยสามารถกำหนดขนาดตัวอย่างของการสุ่มตัวอย่างเพื่อการยอมรับได้ดังนี้

$$\sum_{x=0}^{x^*} {}^{n}C_x p^x (1-p)^{n-x} < 100(1-\alpha)\% \tag{24}$$

เมื่อ x คือ จำนวนของการเกิดผลลัพธ์ที่สนใจ

n คือ จำนวนตัวอย่างที่ใช้ในการสังเกต

p คือ ค่าสัดส่วนของการเกิดผลลัพธ์ที่สนใจ

$100(1-\alpha)\%$ คือ ระดับความเชื่อมั่น

${}^{n}C_x$ คือ $\dfrac{n!}{x!(n-x)!}$

3. วิธีการประมาณค่าเมื่อไม่พบผลลัพธ์ในตัวอย่าง

สำหรับการประมาณค่าที่ใช้ทั่วไปในปัจจุบันเป็นวิธีการประมาณค่าโดยใช้แนวคิดของฟรีเควนทิส (Frequentist) ซึ่งอาศัยหลักการของ Maximum likelihood estimate ในการประมาณค่าสัดส่วน แต่ในกรณีที่ไม่พบผลลัพธ์ในตัวอย่าง การประมาณค่าด้วยวิธีนี้จะไม่สามารถประมาณค่าผลลัพธ์ได้ (Hanley และ Lippman-Hand, 1983; Jovanovic และ Zalenski, 1997) ดังนั้นวิธีการประมาณค่าเมื่อไม่พบผลลัพธ์ในตัวอย่างที่ถูกเสนอแนะมี 2 วิธี ดังนี้

3.1 การประมาณค่าด้วย Exact binomial

Tobi และคณะ (2005) ได้ทำการศึกษาแบบจำลองเหตุการณ์ เพื่อเปรียบเทียบวิธีการประมาณค่าเมื่อผลลัพธ์มีโอกาสเกิดน้อย พบว่าการประมาณค่าด้วย Exact method สามารถประมาณค่าได้ดีกว่าวิธีการอื่น เมื่อ $p \le 0.01$ แต่การแจกแจงแบบทวินามจะมีลักษณะคล้ายการแจกแจงแบบปกติในตัวอย่างที่มีขนาดใหญ่ จึงใช้การประมาณค่าด้วยการแจกแจงปกติ (normal approximation) หรือเรียกว่า Wald method เมื่อ np และ $n(1-p) \ge 5$ แต่ในกรณีที่ผลลัพธ์มีโอกาสเกิดน้อย หรือ np และ $n(1-p) < 5$ หากประมาณค่าด้วย Wald method จะให้ช่วงเชื่อมั่นที่ไม่เหมาะสม (Ghosh, 1979; Blyth, 1983 อ้างใน Agresti และ Coull, 1998) ดังนั้น ในกรณีที่ตัวอย่างมีขนาดเล็ก และความน่าจะเป็นของการเกิดผลลัพธ์มีน้อย จึงต้องใช้วิธีการประมาณค่าแบบ exact binomial (Vollset, 1993; Jovanovic และ Zalenski, 1997; Bailey, 1997)

ถ้ากำหนดให้ X แทนจำนวนครั้งของการเกิดเหตุการณ์ที่สนใจใน n ตัวอย่าง (X=0, 1, 2, ..., n) และความน่าจะเป็นที่แต่ละครั้งของการสังเกตเกิดเหตุการณ์ที่สนใจเป็น p และไม่เกิดเหตุการณ์ที่สนใจเป็น 1-p ดังนั้นจะได้ (Fliess, 2003: p. 17-49)

$$P(X = x) = b(x, n, p) := \binom{n}{x} p^x (1-p)^{n-x} \qquad (25)$$

ถ้า 0<X<n สามารถคำนวณ exact binomial $100(1-\alpha)\%$ confidence interval ของ p ได้เป็น p_l (lower bound) และ p_u (upper bound) โดยที่

ขีดจำกัดล่างช่วงเชื่อมั่น

$$\sum_{x=np}^{n} \frac{n!}{x!(n-x)!} p_l^x (1-p_l)^{n-x} = \alpha/2 \qquad (26)$$

ขีดจำกัดบนช่วงเชื่อมั่น

$$\sum_{x=0}^{np} \frac{n!}{x!(n-x)!} p_u^x (1-p_u)^{n-x} = \alpha/2 \qquad (27)$$

Hanley และ Lippman-Hand (1983) ได้เสนอแนะว่าในกรณีที่ x=0 การประมาณค่าด้วย exact binomial จะเป็นการประมาณค่าแบบทางเดียว (one-sided) โดยที่ค่าขีดจำกัดบนช่วงเชื่อมั่นเท่ากับ

$$(1-p_u)^n = \alpha$$

ซึ่ง Hanley และ Lippman-Hand เรียกวิธีการประมาณค่านี้ว่า "Rule of Three"

เมื่อกำหนดระดับความเชื่อมั่นที่ 95% หรือ $\alpha = 0.05$ จะได้ว่า

$$(1-p_u)^n = 0.05$$

$$n \ln(1-p_u) = \ln(0.05)$$

เมื่อ p มีค่าเข้าใกล้ 0 $[\ln(1-p) \approx -p]$ และ $\ln(0.05) \approx -2.996$

$$n(-p_u) \approx -2.996$$

$$p_u \approx \frac{3}{n}$$

เช่นเดียวกันกับกรณีที่กำหนดระดับความเชื่อมั่นที่ 99% จะได้ "rule of 4.6" (ln 0.01 =-4.6051...) นั่นคือ p = 4.6/n

ดังนั้น เมื่อ x=0 จะคำนวณช่วงเชื่อมั่นได้เท่ากับ $0 < \hat{p} < \frac{3}{n}$

จากการศึกษาของ Hanley และ Lippman-Hand (1983) พบว่าการประมาณค่าด้วย Rule of Three จะให้ค่าประมาณที่ใกล้เคียงกับการประมาณค่าด้วย Exact binomial เมื่อขนาดตัวอย่างมีขนาดใหญ่ (n>30) ดังตารางที่ 4 แต่ Browne (2002) ใน Winkler et al. (2002b) ได้กล่าวว่า การประมาณค่าด้วย Rule of Three จะให้ค่าประมาณที่ไม่เหมาะสม เมื่อ n<20 จะให้ค่าประมาณที่เกินจริงมากกว่าร้อยละ 7.8 และเมื่อขนาดตัวอย่างมากกว่า 80 จะให้ค่าประมาณที่

เกินจริงน้อยกว่าหรือเท่ากับร้อยละ 2 และเสนอแนะให้ใช้ $\frac{3}{(n+1.7)}$ แทน Rule of Three ซึ่งจะให้ค่าประมาณต่ำกว่าค่าจริง (underestimate) น้อยกว่าร้อยละ 1 เมื่อ n>3 และให้ค่าเกินจริงร้อยละ 1.1 เมื่อ n=3

ตารางที่ 4 การประมาณค่าด้วย Rule of Three เมื่อไม่เกิดผลลัพธ์ในขนาดตัวอย่างแต่ละระดับที่ระดับความเชื่อมั่น 95%

Rate of	Rules out any Long-Run Rate(%) Higher than	
	Exactly	Rule of three
0/10	26	30
0/20	14	15
0/30	10	10
0/50	6	6
0/100	3	3
0/1,000	0.3	0.3

*exactly คำนวณจาก $(1-p)^n=0.05$ (Hanley และ Lippman-Hand, 1983)

3.2 การประมาณค่าแบบเบย์เซียน (Bayesian estimate)

เมื่อกำหนดให้ p เป็นพารามิเตอร์ที่สนใจ (อโนทัย ตรีวานิช, 2539)

g(p) เป็นฟังก์ชันความหนาแน่นของ p ซึ่งเป็นตัวแทนที่จะอธิบายถึงความเชื่อที่มีอยู่เกี่ยวกับคุณลักษณะของ p ก่อนที่จะทำการสุ่มตัวอย่างมาสังเกต โดยมีคุณสมบัติว่า สำหรับจำนวน a, b ใด ๆ (a<b) จะได้

เมื่อ p เป็นตัวแปรสุ่มชนิดต่อเนื่อง

$$P(a \le p \le b) = \int_a^b g(p)\,dp$$

เมื่อ p เป็นตัวแปรสุ่มชนิดไม่ต่อเนื่อง

$$P(a \le p \le b) = \sum_{\theta=a}^{b} g(p)$$

และ g(p) จะถูกเรียกว่า ฟังก์ชันความน่าจะเป็นเบื้องต้น (prior probability function)

หลังจากนั้นสุ่มตัวอย่าง X มาขนาด n ซึ่งมีฟังก์ชันความหนาแน่นร่วมเป็น f(x, p) เพื่อทำการสังเกต (สำหรับแนวคิดแบบเบย์เซียน; f(x,p) นี้จะเป็นฟังก์ชันความหนาแน่นแบบเงื่อนไขของ X เมื่อกำหนด p ดังนั้น จึงใช้สัญลักษณ์ f(x|p) แทน) โดยทฤษฎีของ Bayes จะได้ฟังก์ชันความน่าจะเป็นแบบมีเงื่อนไขของ p เมื่อกำหนด x; g(p|x) ให้อยู่ในรูป

กรณีตัวแปรสุ่มชนิดต่อเนื่อง

$$g(p \mid x) = \frac{f(x \mid p)g(p)}{\int f(x \mid p)g(p)dp}$$

กรณีตัวแปรสุ่มชนิดไม่ต่อเนื่อง

$$g(p \mid x) = \frac{f(x \mid p)g(p)}{\sum f(x \mid p)g(p)dp}$$

g(p|x) จะเป็นฟังก์ชันความน่าจะเป็นที่อธิบายความเชื่อเกี่ยวกับ p ภายหลังจากทำการสังเกตข้อมูล X และ g(p|x) จะถูกเรียกว่า ฟังก์ชันความน่าจะเป็นภายหลัง (posterior probability function)

การประมาณค่าสัดส่วนแบบเบย์เซียน

กำหนดให้ x เป็นตัวแปรสุ่มแบบทวินาม ซึ่งมีฟังก์ชันความน่าจะเป็นเบื้องต้นแบบบีต้า ที่มีพารามิเตอร์ a และ b

ซึ่งได้ฟังก์ชัน

$$f(x \mid p) = \binom{n}{x} p^{x}(1-p)^{n-x}$$

$$g(p) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} p^{a-1}(1-p)^{b-1}$$

ดังนั้น

$$f(x \mid p)g(p) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} p^{a-1}(1-p)^{b-1} \binom{n}{x} p^{x}(1-p)^{n-x}$$

$$= \binom{n}{x} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} p^{x+a-1}(1-p)^{n-x+b-1}$$

สำหรับ $0 < p < 1$ และ x=0,1,2,...,n

และจาก

$$\int_0^1 x^{a-1}(1-x)^{b-1}dx = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

ดังนั้นจะได้

$$\int f(x \mid p)g(p)dp = \binom{n}{x} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \cdot \frac{\Gamma(a+x)\Gamma(n-x+b)}{\Gamma(n+a+b)}$$

จะได้ฟังก์ชันความน่าจะเป็นภายหลังสำหรับ p ดังนี้

$$g(p \mid x) = \frac{f(x \mid p)g(p)}{\int_0^1 f(x \mid p)g(p)}$$

$$= \frac{\Gamma(n+a+b)}{\Gamma(x+a)\Gamma(n-x+b)} p^{x+a-1}(1-p)^{n-x+b-1}$$

$$\sim \text{Beta}(x+a, n-x+b)$$

การประมาณค่าช่วงเชื่อถือแบบเบย์เซียน ดังนี้

$$\int_0^{P_u} g(p \mid x)dp = (1-\alpha)100\% \qquad (28)$$

การเลือกฟังก์ชันความน่าจะเป็นเบื้องต้น

การประมาณค่าแบบเบย์เซียนจะต้องอาศัยข้อมูลความน่าจะเป็นเบื้องต้น เพื่อนำมาประมาณค่าความน่าจะเป็นภายหลัง เมื่อทราบความน่าจะเป็นเบื้องต้นแล้วจะต้องเลือกฟังก์ชันความน่าจะเป็นเบื้องต้นของพารามิเตอร์นั้นๆ ที่เหมาะสม ซึ่งวิธีเลือกฟังก์ชันความน่าจะเป็นเบื้องต้นที่นิยมแบบหนึ่งคือการใช้ Conjugate prior

การใช้ Conjugate prior สำหรับพารามิเตอร์ใดๆ ก็คือ การเลือกฟังก์ชันความน่าจะเป็นเบื้องต้น g(p) ที่นำมาปรับด้วยค่า likelihood ของตัวอย่างที่ศึกษา แล้วจะได้รูปแบบฟังก์ชันความน่าจะเป็นภายหลัง g(p|x) เช่นเดียวกับฟังก์ชันความน่าจะเป็นเบื้องต้น g(p) ดังตาราง 5 ถ้ากำหนดให้ตัวอย่างสุ่มชุดหนึ่งสุ่มจากประชากรที่มีฟังก์ชันความน่าจะเป็นแบบทวินามซึ่งมีพารามิเตอร์ p ถ้าฟังก์ชันความน่าจะเป็นเบื้องต้นของ p คือ ฟังก์ชันความน่าจะเป็นแบบบีต้า และ likelihood มีฟังก์ชันความน่าจะเป็นแบบทวินาม แล้วฟังก์ชันความน่าจะเป็นภายหลังของ p ก็จะมีรูปแบบการแจกแจงแบบบีต้าด้วยเช่นกัน

ในการศึกษาต้องการค่าฟังก์ชันความน่าจะเป็นภายหลังแบบบีต้าที่สามารถใช้ประมาณค่าพารามิเตอร์ p ที่มีค่าอยู่ระหว่าง 0 ถึง 1 เพื่อให้สอดรับกับผลการประมาณค่าสัดส่วนของการเกิดผลลัพธ์ในงานรับรองคุณภาพโรงพยาบาล ดังนั้น ฟังก์ชันความน่าจะเป็นเบื้องต้นจะต้องเป็นฟังก์ชันความน่าจะเป็นแบบบีต้า และ likelihood จะมีฟังก์ชันความน่าจะเป็นแบบทวินาม

ตารางที่ 5 ฟังก์ชันความน่าจะเป็นเบื้องต้นของฟังก์ชันความน่าจะเป็นแบบต่างๆ

Prior	**Likelihood**	**Posterior**
Beta (a, b)	Binomial (n, p)	Beta (a+x, n−x+b)
Gamma (a, b)	Poisson (λ)	Gamma (a+x, b+n)
Gamma (a, b)	Exponential (λ)	Gamma (a+n, b+x)

การประมาณค่าแบบเบย์เซียนสามารถแบ่งตามลักษณะของความน่าจะเป็นเบื้องต้นได้ 2 วิธี ดังนี้

3.2.1 กรณีทราบความน่าจะเป็นเบื้องต้น (Informative prior)

การกำหนดความน่าจะเป็นเบื้องต้นทำได้ 2 วิธี คือ 1. การกำหนดจากทฤษฎีหรือความเชื่อของการเกิดผลลัพธ์โดยอาศัยประสบการณ์ เช่น จากประสบการณ์แพทย์คาดว่าโอกาสของการเกิดผลข้างเคียงจากการรักษาเท่ากับ 0.002 ในผู้ป่วย 1,000 ราย เป็นต้น 2. ความน่าจะเป็นเบื้องต้นที่นำมาจากการศึกษาที่ผ่านมา มีวิธีการพิจารณาเลือกค่าพารามิเตอร์สำหรับการคำนวณ ดังนี้

Dixon et al. (2005) ได้ศึกษาวิธีการเพิ่มความแม่นยำในการประมาณค่าผลลัพธ์ที่มีโอกาสเกิดน้อย โดยได้เสนอแนะวิธีการประมาณค่าแบบเบย์เซียนกรณีทราบความน่าจะเป็นเบื้องต้น โดยกล่าวว่าการประมาณค่าแบบเบย์เซียนจะมีความน่าเชื่อถือ และมีความแม่นยำก็ต่อเมื่อเลือกความน่าจะเป็นเบื้องต้นที่เหมาะสม โดยสามารถพิจารณาได้จากค่าส่วนเบี่ยงเบนมาตรฐาน โดยส่วนเบี่ยงเบนมาตรฐานที่มีค่าน้อย จะส่งผลให้มีความแม่นยำในการประมาณค่าสูง ในทางกลับกันหากส่วนเบี่ยงเบนมาตรฐานมีค่ามาก จะส่งผลให้ความแม่นยำต่ำ ซึ่งบ่งบอกถึงการมีข้อมูลที่ไม่เพียงพอของความน่าจะเป็นเบื้องต้น หรือความน่าจะเป็นเบื้องต้นไม่เหมาะสม ถ้าหากมีการศึกษาเกี่ยวกับเรื่องเดียวกันมากกว่า 1 การศึกษา นักวิจัยสามารถเลือกความน่าจะเป็นเบื้องต้นได้จากการพิจารณาค่าส่วนเบี่ยงเบนมาตรฐานของการศึกษา

ในกรณีที่กำหนดความน่าจะเป็นเบื้องต้นจากการศึกษาที่ผ่านมา จะสามารถกำหนดค่าพารามิเตอร์ของฟังก์ชันความน่าจะเป็นเบื้องต้นแบบบีต้าได้จากความสัมพันธ์ระหว่างค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐานของการแจกแจงทวินาม และการแจกแจงแบบบีต้า โดยสามารถคำนวณค่าพารามิเตอร์ a และ b จากการแก้สมการที่ 29 และ 30 (Bolstad, 2004) ดังนี้

เมื่อ p คือ ค่าเฉลี่ยของค่าสัดส่วน (สมการ 3)

$$p = \frac{a}{a+b} \qquad (29)$$

ส่วนเบี่ยงเบนมาตรฐานของความน่าจะเป็นแบบบีต้า คือ

$$\sqrt{\frac{ab}{(a+b)^2(a+b+1)}}$$

เมื่อ $\frac{b}{a+b} = 1-p$

ส่วนเบี่ยงเบนมาตรฐานของค่าสัดส่วน คือ $\sigma = \sqrt{\frac{p(1-p)}{a+b+1}}$ (30)

นอกจากนี้การกำหนดค่าพารามิเตอร์ของฟังก์ชันความน่าจะเป็นเบื้องต้นแบบบีต้าสามารถกำหนดได้ด้วยการพิจารณาจากรูปแบบการแจกแจงเบื้องต้นที่สามารถสะท้อนความเชื่อของบุคคลต่อค่าพารามิเตอร์ได้ เช่น หากมีความเชื่อเกี่ยวกับความน่าจะเป็นของพารามิเตอร์ที่มีโอกาสเกิดน้อย ลักษณะการแจกแจงเบื้องต้นจะต้องมีการแจกแจงแบบเบ้ขวา โดยมีพารามิเตอร์ $a<b$ เป็นต้น หรือกรณีทราบขนาดตัวอย่าง (n) และค่าเฉลี่ยของค่าสัดส่วน (p) สามารถที่จะคำนวณค่าพารามิเตอร์ของบีต้าได้จาก (Lee, 1989)

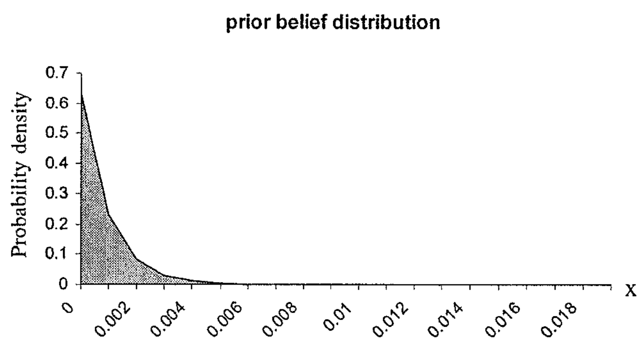
$$p = \frac{a}{(a+b)} \qquad (31)$$

$$n = a+b \qquad (32)$$

เช่น p = 0.02 n = 120 จะได้ a = 2.4 และ b= 117.6 เป็นต้น

จากการศึกษาของ Winkler และคณะ (2002a) เสนอแนะว่าเมื่อผลลัพธ์เป็นศูนย์อาจจะใช้วิธีการกำหนดค่าพารามิเตอร์ของบีต้าได้จากความสัมพันธ์ระหว่างค่าเฉลี่ยของค่าสัดส่วน กับ $5^{th}/95^{th}$ percentiles ของค่าสัดส่วน นอกจากนี้ยังสามารถกำหนดค่าพารามิเตอร์ของการแจกแจงแบบบีต้าได้จากการกำหนดค่าฐานนิยมของค่าสัดส่วน และ $5^{th}/95^{th}$ percentiles ของค่าสัดส่วน เช่น mode=0.01 95^{th} percentile=0.015 จะได้ a=17.74 b=1658.26 เป็นต้น ซึ่งในปัจจุบันได้มีผู้ที่ศึกษา และเขียนคำสั่งสำเร็จรูปในการคำนวณค่าพารามิเตอร์ของบีต้าที่เปิดให้ใช้ฟรี เช่น http://www.ausvet.com.au/pprev/content.php?page=BetaParams เป็นต้น

สำหรับการศึกษาในครั้งนี้ทำการศึกษาเพื่อการรับรองคุณภาพโรงพยาบาล ในกรณีที่ผลลัพธ์ในตัวอย่างเท่ากับศูนย์ ซึ่งผู้วิจัยคาดว่าเมื่อผลลัพธ์ของตัวชี้วัดมีโอกาสเกิดน้อย ค่าฐานนิยมของผลลัพธ์ของตัวชี้วัดจะมีค่าเข้าใกล้หรือเท่ากับ 0 ดังนั้น จากหน้า 12 เมื่อค่าฐานนิยมเท่ากับ 0 ฟังก์ชันความน่าจะเป็นเบื้องต้นแบบบีต้าจะมีค่าพารามิเตอร์ a=1 และ b>1 ดังตัวอย่างในภาพที่ 5



ภาพที่ 5 ตัวอย่างลักษณะการแจกแจงแบบบีต้าที่มีค่าฐานนิยมเท่ากับ 0 (a=1, b=999)

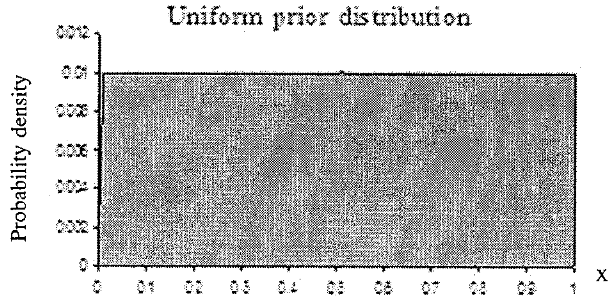
3.2.2 กรณีไม่ทราบความน่าจะเป็นเบื้องต้น (Non-informative prior)

บ่อยครั้งที่เราไม่แน่ใจในข้อมูล หรือความเชื่อเกี่ยวกับโอกาสของการเกิดผลลัพธ์ที่สนใจ ก่อนที่จะทำการศึกษาจากข้อมูลโดยการสุ่มตัวอย่าง หรือบางครั้งเราไม่สามารถหาข้อมูลความน่าจะเป็นของการเกิดผลลัพธ์ที่สนใจได้ เนื่องจากยังไม่มีผู้ที่ทำการศึกษามาก่อน อย่างไรก็ตาม การอนุมานแบบเบย์เซียนจำเป็นต้องกระทำโดยอาศัยความน่าจะเป็นเบื้องต้น ดังนั้น นักวิจัยจึงต้องกำหนดค่าความน่าจะเป็นเบื้องต้นซึ่งเรียกการอนุมานแบบเบย์เซียนนี้ว่า non-informative prior ในที่นี้ขอเสนอ prior ที่ถูกใช้บ่อยในกรณี non-informative prior 2 วิธี ได้แก่ Uniform priors และ Jeffrey's priors (Sahu, 2002; Winkler et al., 2002a)

3.2.2.1 Uniform priors

จากการเลือกฟังก์ชันความน่าจะเป็นเบื้องต้นของ p โดยใช้การ conjugate prior จะได้ฟังก์ชันความน่าจะเป็นเบื้องต้นแบบบีต้า โดยที่การแจกแจงแบบยูนิฟอร์มจะเป็นการ

แจกแจงกรณีพิเศษของการแจกแจงแบบบีต้า เมื่อ a และ b เท่ากับ 1 และมีลักษณะการแจกแจงดังภาพที่ 6 ดังนั้น เมื่อ P~B(1,1) เป็นความน่าจะเป็นเบื้องต้นแล้ว จะได้ความน่าจะเป็นภายหลังเป็น P~B(x+1, n-x+1) และจะได้ค่าเฉลี่ยของค่าสัดส่วนภายหลังเท่ากับ $\frac{1+x}{2+n}$ (Zhu และ Lu, 2004)



ภาพที่ 6 การแจกแจงเบื้องต้นแบบ Uniform

3.2.2.2 Jeffrey's priors

Jeffrey's priors เป็นความน่าจะเป็นเบื้องต้นที่ถูกใช้บ่อยในการประมาณค่าแบบเบย์เซียน กรณีไม่ทราบความน่าจะเป็นเบื้องต้น โดยที่ P~B(0.5,0.5) และมีลักษณะการแจกแจงแบบ U-shape ที่มีค่าเฉลี่ยเท่ากับ 0.5 และค่าฐานนิยมเท่ากับ 0 และ 1 ดังภาพที่ 7

ฟังก์ชันความน่าจะเป็นเบื้องต้นแบบ Jeffrey's ดังนี้

$$g(p) = \{I(p)\}^{1/2} \tag{33}$$

เมื่อ $I(p)$ คือ Fisher information

$$I(p) = -E\left[\frac{\partial^2}{\partial p^2} f(x \mid p)\right]$$

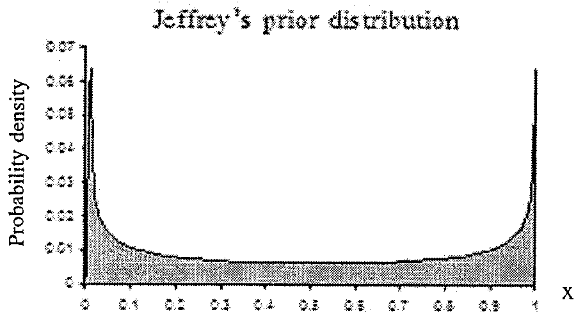
ถ้ากำหนดให้ตัวอย่างสุ่มชุดหนึ่งสุ่มจากประชากรที่มีฟังก์ชันความน่าจะเป็นแบบทวินามซึ่งมีพารามิเตอร์ p จะได้ non-informative prior ดังนี้

$$f(p) \propto p^{-\frac{1}{2}}(1-p)^{-\frac{1}{2}}$$

ซึ่งจะได้ Posterior distribution ดังนี้

$$f(p \mid x) \propto p^{x-\frac{1}{2}}(1-p)^{n-x-\frac{1}{2}} \tag{34}$$

สำหรับรายละเอียดวิธีการของ Jeffrey's prior ไม่ได้นำเสนอไว้ ผู้สนใจสามารถศึกษาได้ใน Box and Tiao, 1973 (Section 1.3)



ภาพที่ 7 การแจกแจงเบื้องต้นแบบ Jreffrey's

เมื่อผลลัพธ์ในตัวอย่างไม่เป็นศูนย์ การประมาณค่าแบบเบย์เซียนกรณีไม่ทราบความน่าจะเป็นเบื้องต้นจะไม่มีผลกระทบต่อผลการประมาณค่า ทั้งนี้เนื่องจากข้อมูลที่ได้จากการสังเกตจะเป็นสิ่งที่นำมาปรับความน่าจะเป็นภายหลังให้เป็นค่าประมาณที่มีความน่าเชื่อถือมากกว่าการอาศัยข้อมูลความน่าจะเป็นเบื้องต้นเพียงอย่างเดียว (Gelman et al., 1995; Ellison, 1996 อ้างใน Dixon et al., 2005) ในทางตรงกันข้ามเมื่อผลลัพธ์ในตัวอย่างเป็นศูนย์ การกำหนดความน่าจะเป็นเบื้องต้นจะมีอิทธิพลต่อความน่าจะเป็นภายหลัง ซึ่งการใช้ความน่าจะเป็นเบื้องต้นแบบ Uniform และ Jeffrey's จะให้ค่าประมาณที่สูงกว่าการประมาณค่าแบบเบย์เซียน กรณีทราบความน่าจะเป็นเบื้องต้น (Basu et al., 1996; Bailey et al., 2002) นอกจากนี้เมื่อไม่พบผลลัพธ์ในตัวอย่างลักษณะการแจกแจงเบื้องต้นแบบบีต้าควรมีลักษณะการแจกแจงแบบ Unimodel ดังภาพที่ 5 แต่จากภาพที่ 7 จะเห็นได้ว่าลักษณะการแจกแจงเบื้องต้นของ Jeffrey's มีลักษณะการแจกแจงแบบ Bimodel ที่มีค่าฐานนิยมเท่ากับ 0 และ 1 ซึ่งในกรณีที่ไม่พบผลลัพธ์ในตัวอย่าง ค่าฐานนิยมของการแจกแจงเบื้องต้นควรเท่ากับ 0 เท่านั้น จึงถือได้ว่าการแจกแจงเบื้องต้นแบบ Jeffrey's เป็นการแจกแจงเบื้องต้นที่ไม่เหมาะสมสำหรับกรณีผลลัพธ์ในตัวอย่างเป็นศูนย์

และจากการศึกษาที่ผ่านมาพบว่าการประมาณค่าแบบเบย์เซียนกรณีทราบความน่าจะเป็นเบื้องต้นจะให้ค่าประมาณที่ดีกว่าการประมาณค่าด้วยวิธีฟรีเควนทิส (Basu et al., 1996; Smith et al., 2000) และให้ค่าประมาณที่ดีกว่าการประมาณค่าด้วย Rule of Three (Smith et al., 2000; Bailey et al., 2002)

สรุป ในการกำหนดความน่าจะเป็นเบื้องต้นของการประมาณค่าช่วงเชื่อถือเบย์เซียนเพื่อการรับรองคุณภาพโรงพยาบาล โดยเชื่อว่าแต่ละโรงพยาบาลมีการพัฒนาคุณภาพงานการรักษาพยาบาลจนได้ระดับมาตรฐานจึงจะขอการรับรองคุณภาพโรงพยาบาล ดังนั้น ความน่าจะเป็นเบื้องต้นของผลการตรวจจึงมีความเชื่อได้ว่าเป็นไปตามเกณฑ์ที่กำหนด การศึกษาในครั้งนี้จึงใช้เกณฑ์คุณภาพเป็นค่าเฉลี่ยของการแจกแจงความน่าจะเป็นเบื้องต้นในการประมาณค่าขีดจำกัดบนช่วงเชื่อถือเบย์เซียน เมื่อผลลัพธ์ในตัวอย่างเป็นศูนย์