

การเปรียบเทียบประสิทธิภาพช่วงความเชื่อมั่นสำหรับผลต่างระหว่าง ค่าสัดส่วนทวินามสองกลุ่มที่อิสระกัน

Efficiency Comparison of Confidence Intervals for the Difference Between Two Independent Binomial Proportions

นวรัตน์ รื่นแสง (Nawarat Ruensaeng) *

จุฑาภรณ์ สิ้นสมบุญทอง (Juthaphorn Sinsomboonthong)**

วันดี วณิชยศักดิ์พงศ์ (Wandee Wanishsakpong)***

บทคัดย่อ

งานวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพช่วงความเชื่อมั่น 95 % สำหรับผลต่างระหว่างค่าสัดส่วนทวินามสองกลุ่มที่อิสระกัน 5 วิธี ได้แก่ วิธี Asymptotic, CC วิธี Newcombe วิธี Agresti-Caffo วิธี KMS และวิธี Recentered โดยเกณฑ์ที่ใช้ในการเปรียบเทียบประสิทธิภาพ คือ การตรวจสอบค่าความน่าจะเป็นครอบคลุม และค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นซึ่งวิธีการประมาณแบบช่วงที่มีประสิทธิภาพดีจะต้องให้ค่าความน่าจะเป็นครอบคลุมเป็นไปตามที่กำหนดและมีค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นแคบที่สุดในการศึกษาครั้งนี้ จำลองข้อมูลที่มีการแจกแจงทวินามสองกลุ่มที่อิสระกันด้วยเทคนิคมอนติคาร์โล รวมสถานการณ์ที่ศึกษาทั้งหมด 30 สถานการณ์และในแต่ละสถานการณ์ทดลองซ้ำ 2,000 ครั้ง ผลการวิจัยสรุปได้ดังนี้ วิธี KMS มีแนวโน้มให้ประสิทธิภาพดีที่สุดเมื่อ $p_1 - p_2$ เท่ากับ 0.8 ที่ตัวอย่างขนาด (n_1, n_2) เท่ากับ (10, 10), (30, 30) และ (70, 70) สำหรับวิธี Recentered มีแนวโน้มให้ประสิทธิภาพดีที่สุดในเกือบทุกระดับของขนาดตัวอย่าง (n_1, n_2) เมื่อผลต่างระหว่างค่าสัดส่วนทวินาม $(p_1 - p_2)$ เท่ากับ 0, 0.2 และ 0.8 ยกเว้นกรณีขนาดตัวอย่าง (n_1, n_2) เท่ากับ (10, 10) ที่ผลต่างระหว่างค่าสัดส่วนทวินาม $(p_1 - p_2)$ เท่ากับ 0.2 และขนาดตัวอย่าง (n_1, n_2) เท่ากับ (10, 10), (30, 30) และ (70, 70) ที่ผลต่างระหว่างค่าสัดส่วนทวินาม

* นิสิตปริญญาโท ภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ บางเขน

(Student, Department of Statistics, Faculty of Science, Kasetsart University, Bangkok.)

Email; pompam_stargirl@hotmail.com)

** อาจารย์ที่ปรึกษาหลัก, ผู้ช่วยศาสตราจารย์ ดร.; ประจำภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ บางเขน

(Major Advisor, Lecturer, Department of Statistics, Faculty of Science, Kasetsart University, Bangkok.)

Email; fscijps@ku.ac.th)

*** อาจารย์ที่ปรึกษาร่วม, อาจารย์; ประจำภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ บางเขน

(Co-Advisor, Lecturer, Department of Statistics, Faculty of Science, Kasetsart University, Bangkok.)

Email; fsciwdw@ku.ac.th)

$(p_1 - p_2)$ เท่ากับ 0.8 นอกจากนี้วิธี Agresti-Caffo มีแนวโน้มให้ประสิทธิภาพดีที่สุดเมื่อผลต่างระหว่างค่าสัดส่วนทวินามเท่ากับ 0.4 และ 0.6 ในทุกระดับของขนาดตัวอย่าง

คำสำคัญ : การแจกแจงทวินาม ค่าสัดส่วนทวินาม ช่วงความเชื่อมั่น ค่าความน่าจะเป็นครอบคลุม

Abstract

The objective of this research was to compare the efficiencies of 95% confidence intervals of five methods—Asymptotic, CC, Newcombe, Agresti-Caffo, KMS and Recentered methods—for the difference between two independent binomial proportions. The criteria for efficiency comparison were the coverage probability examination and the average length of the confidence interval. The efficient coverage probability was not lower than the given confidence level and the efficient average length of the confidence interval was the shortest value. A simulation dataset was conducted in the form of two independent binomial distributions for 30 situations and repeated 2,000 times for each situation by using the Monte Carlo simulation technique. The conclusions of this research are as follows: KMS method tends to have a good performance when the difference between two independent binomial proportions $(p_1 - p_2)$ equals 0.8 for the sample size (n_1, n_2) equals (10, 10), (30, 30) and (70, 70). For the difference between two independent binomial proportions $(p_1 - p_2)$ equals 0, 0.2 and 0.8, the Recentered method tends to have a good performance for almost all levels of the sample size (n_1, n_2) , except the sample size (n_1, n_2) equals (10, 10) for the difference between two independent binomial proportions $(p_1 - p_2)$ equals 0.2 and sample size (n_1, n_2) equals (10, 10), (30, 30) and (70, 70) for the difference between two independent binomial proportions $(p_1 - p_2)$ equals 0.8. Furthermore, Agresti-Caffo method tends to have a good performance when the difference between two independent binomial proportions $(p_1 - p_2)$ equals 0.4 and 0.6 for all levels of the sample size (n_1, n_2) .

Keywords: Binomial distribution, Binomial Proportion, Confidence Interval, Coverage Probability

บทนำ

การประมาณค่าพารามิเตอร์แบ่งออกเป็น 2 ประเภท คือ การประมาณแบบค่าเดียว (Point Estimation) และการประมาณค่าแบบช่วง (Interval Estimation) โดยการประมาณพารามิเตอร์แบบค่าเดียวเป็นการประมาณพารามิเตอร์ที่สนใจด้วยค่าสถิติเพียงค่าเดียวซึ่งการประมาณแบบค่าเดียวมีโอกาสเกิดความผิดพลาดได้สูง เนื่องจากมีโอกาสค่อนข้างน้อยที่ตัวประมาณจะมีค่าเท่ากับค่าพารามิเตอร์ที่แท้จริง ตัวอย่างเช่น ถ้า $\hat{\theta}$ เป็นตัวประมาณของพารามิเตอร์ θ และสมมติว่า $\hat{\theta}$ มีการแจกแจงปกติ แล้วความน่าจะเป็นที่ $\hat{\theta}$ เท่ากับ θ หรือ $P(\hat{\theta} = \theta)$ จะมีค่าเท่ากับศูนย์เป็นต้น ส่วนการประมาณค่าแบบช่วงเป็นการประมาณพารามิเตอร์ที่สนใจโดยให้ค่าประมาณเป็นช่วงและคำนึงถึงความคลาดเคลื่อนจากการสุ่มตัวอย่างเพื่อช่วยยืนยันว่าช่วงที่สร้างนี้คลุมค่าพารามิเตอร์ไว้ด้วยความน่าจะเป็นตามที่กำหนด ดังนั้นการประมาณพารามิเตอร์แบบช่วงมีโอกาที่จะประมาณค่าได้ถูกต้องใกล้เคียงกับค่าพารามิเตอร์ที่แท้จริงมากกว่าการประมาณแบบค่าเดียว กล่าวคือการประมาณพารามิเตอร์ θ แบบช่วงที่ระดับความเชื่อมั่น $(1-\alpha)100\%$ สามารถเขียนได้ดังนี้ $L < \theta < U$ ซึ่งหมายความว่าความน่าจะเป็นที่ช่วงความเชื่อมั่นระหว่างขอบเขตล่าง (L) และขอบเขตบน (U) จะคลุมพารามิเตอร์ θ มีค่าเท่ากับ $1-\alpha$ เมื่อ $0 < \alpha < 1$ โดยสัญลักษณ์ $1-\alpha$ แทนสัมประสิทธิ์ความเชื่อมั่น ซึ่งเป็นโอกาสที่ช่วงที่ประมาณได้คลุมค่าพารามิเตอร์ที่แท้จริง ดังนั้นถ้าให้สัมประสิทธิ์ความเชื่อมั่น $1-\alpha$ มีค่ามาก จะทำให้โอกาสที่ช่วงที่ประมาณได้คลุมค่าพารามิเตอร์ที่แท้จริงมีค่ามาก (มหาวิทยาลัยเกษตรศาสตร์, 2560)

การศึกษาการประมาณผลต่างระหว่างค่าสัดส่วนทวินามสองกลุ่มที่อิสระกันได้นำไปใช้ประโยชน์ในงานวิจัยทางด้านต่าง ๆ เช่น ด้านการแพทย์อาจสนใจศึกษาการประมาณผลต่างระหว่างสัดส่วนความมีประสิทธิภาพของยา 2 ชนิดกับคนไข้ จากการให้ยาชนิดที่หนึ่งกับคนไข้กลุ่มหนึ่ง และให้ยาชนิดที่สองกับคนไข้กลุ่มที่สอง โดยเมื่อคนไข้รับประทานยาแล้วจะถูกวัดผลเป็น 2 กรณี คือ ยาใช้ได้ผล หรือ ไม่ได้ผล หรืออาจสนใจศึกษาผลของการเยี่ยมบ้านผู้ป่วยโรคความดันโลหิตสูงโดยทีมสหสาขาวิชาชีพโดยการแบ่งกลุ่มคนไข้เป็น 2 กลุ่ม โดยกลุ่มหนึ่งได้รับการเยี่ยมบ้านโดยทีมสหสาขาวิชาชีพ และได้รับการบริหารทางเภสัชกรรมตามปกติของโรงพยาบาล ขณะที่กลุ่มที่สองได้รับการบริหารทางเภสัชกรรมตามปกติของโรงพยาบาลเท่านั้น โดยเมื่อคนไข้ได้รับการเยี่ยมบ้านโดยทีมสหสาขาวิชาชีพแล้วจะถูกวัดผลเป็น 2 กรณี คือ สามารถควบคุมความดันโลหิตได้ตามเป้าหมาย หรือ ไม่สามารถควบคุมความดันโลหิตได้ตามเป้าหมาย (พุทธชาติ และคณะ, 2559) ในทางด้านการตลาดก็ได้้นำการประมาณผลต่างระหว่างค่าสัดส่วนทวินามสองกลุ่มที่อิสระกันไปใช้ เช่น การศึกษาการประมาณผลต่างระหว่างสัดส่วนรถยนต์ที่ขายได้ 2 รุ่นในแต่ละสัปดาห์ของโชว์รูมแห่งหนึ่ง ซึ่งจะมีลูกค้ามาดูรถยนต์ระหว่างสัปดาห์หลายคน โดยเมื่อลูกค้ามาดูรถยนต์แต่ละรุ่นแล้วจะวัดผลการตัดสินใจเป็น 2 กรณี คือ ซื้อ หรือ ไม่ซื้อรถยนต์ จะเห็นได้ว่าการประมาณผลต่างระหว่างค่าสัดส่วนทวินามสองกลุ่มที่อิสระกันมีประโยชน์อย่างยิ่งในงานวิจัยด้านต่าง ๆ และถ้าการประมาณมีความถูกต้องและแม่นยำเพียงพอจะทำให้เกิดประโยชน์สูงสุดต่อหน่วยงานที่นำผลการวิจัยไปใช้ ซึ่งการประมาณค่าแบบช่วงสำหรับผลต่างระหว่างค่าสัดส่วนทวินามสองกลุ่มที่อิสระกันมีหลายวิธีดังนี้ วิธี Asymptotic with continuity correction หรือวิธี Asymptotic, CC (Fleiss, 1981) ได้รับการพัฒนามาจากวิธี Asymptotic no Continuity Correction (Wald and Wolfowitz,

1939) โดยเพิ่มความต่อเนื่อง $\left(\frac{1}{n_1} + \frac{1}{n_2} \right) \frac{1}{2}$ ในการสร้างช่วงความเชื่อมั่นเพื่อให้ความต่อเนื่องของช่วงความเชื่อมั่นเข้าใกล้ค่าจริงมากขึ้น และเป็นวิธีที่ให้ประสิทธิภาพดีสำหรับทุกขนาดตัวอย่างเมื่อผลต่างระหว่างค่าสัดส่วนของสิ่งที่สนใจมากกว่าหรือเท่ากับศูนย์ ส่วนวิธี Score without continuity correction หรือเรียกอีกชื่อว่าวิธี Newcombe เป็นวิธีการประมาณแบบช่วงความเชื่อมั่นสำหรับค่าสัดส่วนทวินามสองประชากรที่อิสระกัน โดยพัฒนามาจากวิธี Score (Wilson, 1927) ซึ่งเป็นวิธีที่ดีอีกวิธีหนึ่งสำหรับทุกขนาดตัวอย่างเมื่อผลต่างระหว่างค่าสัดส่วนของสิ่งที่สนใจมากกว่าหรือเท่ากับศูนย์เช่นกัน (Newcombe, 1998) ในปี ค.ศ. 2000 Agresti และ Caffo ได้นำเสนอวิธีการประมาณแบบช่วงความเชื่อมั่นสำหรับค่าสัดส่วนทวินามสองประชากรที่อิสระกันขึ้นมาใหม่โดยการเพิ่มลักษณะสิ่งที่สนใจจำนวน 2 ค่า และสิ่งที่ไม่สนใจอีก 2 ค่า สำหรับแต่ละตัวอย่าง โดยนำแนวคิดของวิธีเบส์มาประยุกต์ใช้ร่วมกับวิธีของวาลด์ วิธี Agresti-Caffo เป็นวิธีที่ให้ประสิทธิภาพดีที่สุดเมื่อผลต่างระหว่างค่าสัดส่วนของสิ่งที่สนใจไม่เท่ากับศูนย์ (Prendergast and Staudte, 2014) ต่อมาในปี ค.ศ. 2005 Brown และ Li ได้เสนอช่วงความเชื่อมั่นแบบใหม่สำหรับการประมาณผลต่างระหว่างค่าสัดส่วนทวินามสองกลุ่มที่อิสระกันคือ วิธี Recentered และได้ทำการเปรียบเทียบประสิทธิภาพกับวิธี Wald วิธี Jeffrey วิธี Agresti วิธี Newcombe และวิธี Score ผลการวิจัยพบว่าวิธี Recentered เป็นวิธีที่มีประสิทธิภาพดีมากเมื่อผลต่างระหว่างค่าสัดส่วน ทวินามสองกลุ่มไม่เท่ากับศูนย์และตัวอย่างมีขนาดเล็ก นอกจากนี้ยังมีการวิจัยที่เกี่ยวข้องกับการศึกษาประสิทธิภาพของวิธีการประมาณค่าแบบช่วงสำหรับผลต่างระหว่างค่าสัดส่วนทวินามสองกลุ่มที่อิสระกันดังนี้ ในปี ค.ศ. 2014 Prendergast และ Staudte ได้เปรียบเทียบประสิทธิภาพวิธี Wald วิธี Agresti-Caffo วิธี Newcombe และวิธี KMS สำหรับประชากรขนาดเล็ก จากการศึกษาพบว่าวิธี Agresti-Caffo เป็นวิธีที่ให้ประสิทธิภาพดีที่สุดเมื่อผลต่างระหว่างค่าสัดส่วนของสิ่งที่สนใจไม่เท่ากับศูนย์ในขณะที่วิธี KMS (Kulinskaya et al., 2010) เป็นวิธีที่ให้ประสิทธิภาพดีเช่นกันเมื่อผลต่างระหว่างค่าสัดส่วนของสิ่งที่สนใจมีค่าไม่เท่ากับศูนย์ และในปี พ.ศ. 2553 สุภารัตน์ได้ศึกษาประสิทธิภาพของวิธีการประมาณช่วงความเชื่อมั่นสำหรับผลต่างระหว่างค่าสัดส่วนทวินามสองกลุ่มที่อิสระกัน 4 วิธี ได้แก่ วิธีวาลด์ วิธี Adding-4 วิธี T2 และวิธี Recentered จากงานวิจัยพบว่าวิธีวาลด์เหมาะสมกับตัวอย่างขนาดใหญ่ วิธี Adding-4 เหมาะในกรณีขนาดตัวอย่าง (n_1, n_2) เท่ากับ (10, 10) และ $p_1 - p_2$ มีค่าเท่ากับ 0.4 และ 0.6 วิธี T2 เหมาะสมในทุกกรณีที่ศึกษา และวิธี Recentered เหมาะสมเกือบทุกกรณีที่ศึกษาเช่นกัน ต่อมาในปี พ.ศ. 2558 กนกกาญจน์ และคณะ ได้ศึกษาเปรียบเทียบประสิทธิภาพวิธีการประมาณช่วงความเชื่อมั่นสำหรับผลต่างระหว่างค่าสัดส่วนทวินามประชากรสองกลุ่มที่อิสระกันเมื่อตัวอย่างขนาดเล็ก โดยได้เปรียบเทียบค่าสัมประสิทธิ์ความเชื่อมั่นและค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นจาก 5 วิธี ได้แก่ วิธีแบบฉบับ วิธี Newcombe วิธี Adding-4 วิธี Peskun และวิธี Recentered จากผลงานวิจัยพบว่าวิธี Newcombe เหมาะสมเมื่อ $p_1 - p_2$ มีค่าเท่ากับ 0.3 ถึง 0.5 วิธี Peskun เหมาะสมในทุกกรณีที่ศึกษา วิธี Adding-4 เหมาะสมที่สุดเมื่อ $p_1 - p_2$ มีค่าเท่ากับ 0.5 และวิธี Recentered เหมาะสมเกือบทุกกรณีที่ศึกษา

จากงานวิจัยที่กล่าวมาข้างต้นพบว่าวิธี Asymptotic, CC และวิธี Newcombe เป็นวิธีที่เหมาะสมสำหรับการประมาณช่วงความเชื่อมั่นเมื่อ $p_1 - p_2$ มีค่ามากกว่าหรือเท่ากับศูนย์ ส่วนวิธี Agresti-Caffo วิธี KMS และวิธี Recentered เป็นวิธีที่เหมาะสมสำหรับการประมาณช่วงความเชื่อมั่นเมื่อ $p_1 - p_2$ มีค่าไม่เท่ากับศูนย์ ซึ่งส่วนใหญ่ถูกพิจารณาในกรณีตัวอย่างขนาดเล็ก และผลการศึกษาพบว่าส่วนใหญ่วิธีเหล่านี้มีความเหมาะสมสำหรับการประมาณช่วงความเชื่อมั่นในสถานการณ์ใกล้เคียงกันสำหรับกรณีขนาดตัวอย่างเล็กเท่านั้น ซึ่งบางวิธีสามารถคำนวณได้ง่ายแต่บางวิธีอาจคำนวณได้ค่อนข้างยาก ดังนั้นผู้วิจัยจึงเลือกวิธีที่เหมาะสมในแต่ละงานวิจัยมาศึกษาเพิ่มเติมในกรณีตัวอย่างขนาดใหญ่ขึ้นโดยศึกษาเพิ่มเติมกรณีขนาดตัวอย่างเท่ากับ (50, 50), (70, 70) และ (100, 100) เพื่อพิจารณาวิธีที่เหมาะสมสำหรับการประมาณช่วงความเชื่อมั่นในกรณีขนาดตัวอย่างใหญ่นอกจากนี้หากวิธีเหล่านี้ให้ความเหมาะสมสำหรับการประมาณช่วงความเชื่อมั่นในสถานการณ์ใกล้เคียงกัน ผู้ที่สนใจในงานวิจัยนี้จะสามารถเลือกวิธีการประมาณที่เหมาะสมโดยพิจารณาจากความยากง่ายในการคำนวณแต่ละวิธีที่ศึกษาเพื่อหาวิธีที่เหมาะสมที่สุดสำหรับการประมาณช่วงความเชื่อมั่นได้ งานวิจัยนี้ผู้วิจัยจึงสนใจศึกษาเปรียบเทียบประสิทธิภาพช่วงความเชื่อมั่นสำหรับผลต่างระหว่างค่าสัดส่วนทวินามสองกลุ่มที่อิสระกัน 5 วิธี ได้แก่ วิธี Asymptotic, CC วิธี Newcombe วิธี Agresti-Caffo วิธี KMS และวิธี Recentered เพื่อหาวิธีที่เหมาะสมต่อการนำไปใช้ประมาณช่วงความเชื่อมั่นในแต่ละสถานการณ์ที่ศึกษา โดยเกณฑ์ที่ใช้ในการเปรียบเทียบประสิทธิภาพ คือ การตรวจสอบค่าความน่าจะเป็นครอบคลุม และค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นซึ่งวิธีการประมาณแบบช่วงที่มีประสิทธิภาพจะต้องให้ค่าความน่าจะเป็นครอบคลุมหรือค่าสัมประสิทธิ์ความเชื่อมั่นเป็นไปตามที่กำหนดและมีค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นแคบที่สุด

วิธีวิจัย

การเปรียบเทียบประสิทธิภาพช่วงความเชื่อมั่น 95 % สำหรับผลต่างระหว่างค่าสัดส่วนทวินามสองกลุ่มที่อิสระกัน 5 วิธี คือ วิธี Asymptotic, CC วิธี Newcombe วิธี Agresti-Caffo วิธี KMS และวิธี Recentered มีขั้นตอนการวิจัยดังต่อไปนี้

1. กำหนดขนาดตัวอย่าง (n_1, n_2) เท่ากับ (10, 10), (20, 20), (30, 30), (50, 50), (70, 70) และ (100, 100) และกำหนดผลต่างระหว่างค่าสัดส่วนทวินามสองกลุ่มที่อิสระกัน $(p_1 - p_2)$ เท่ากับ 0, 0.2, 0.4, 0.6 และ 0.8 โดย $p_2 = 0.1$

2. ใช้โปรแกรม SAS 9.4 ในการจำลองข้อมูลที่มีการแจกแจงทวินามสองกลุ่มที่อิสระกันด้วยพารามิเตอร์ (n_1, p_1) และ (n_2, p_2) ตามลำดับ โดยเทคนิคมอนติคาร์โล (Monte Carlo Simulation) รวมสถานการณ์ที่ศึกษาทั้งหมด 30 สถานการณ์และมีการทำซ้ำ 2,000 ครั้ง ในแต่ละสถานการณ์

3. กำหนดระดับความเชื่อมั่นที่ 95%

4. คำนวณช่วงความเชื่อมั่นสำหรับผลต่างระหว่างค่าสัดส่วนทวินามของสองประชากรที่อิสระกันที่ศึกษาในงานวิจัยนี้ มี 5 วิธี คือ

- 4.1) วิธี Asymptotic with Continuity Correction หรือวิธี Asymptotic, CC

Fleiss (1981) ได้นำเสนอวิธี Asymptotic with Continuity Correction ซึ่งเป็นวิธีการประมาณแบบช่วงความเชื่อมั่นสำหรับผลต่างระหว่างค่าสัดส่วนทวินามสองประชากรที่อิสระกันซึ่งพัฒนามาจาก

วิธี Asymptotic no Continuity Correction (Wald and Wolfowitz, 1939) โดยการปรับแก้ไขเพิ่มความต่อเนื่อง จึงได้ใช้ค่า $\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\frac{1}{2}$ มาปรับในการสร้างช่วงความเชื่อมั่นเพื่อความต่อเนื่องทำให้ช่วงความเชื่อมั่นเข้าใกล้ค่าจริงมากขึ้น

กำหนดให้ขอบเขตล่างและขอบเขตบนของช่วงความเชื่อมั่นวิธี Asymptotic, CC เขียนแทนด้วยสัญลักษณ์ L_A และ U_A ตามลำดับ จะได้ว่าช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของ p_1-p_2 โดยวิธี Asymptotic, CC (Fleiss, 1981) เป็นดังนี้

$$L_A \text{ คือ } (\hat{p}_1 - \hat{p}_2) - \left(Z_{\frac{\alpha}{2}} \sqrt{\left(\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2} \right) + \left(\frac{1}{n_1} + \frac{1}{n_2} \right) \frac{1}{2}} \right)$$

$$U_A \text{ คือ } (\hat{p}_1 - \hat{p}_2) + \left(Z_{\frac{\alpha}{2}} \sqrt{\left(\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2} \right) + \left(\frac{1}{n_1} + \frac{1}{n_2} \right) \frac{1}{2}} \right)$$

เมื่อ X_1 เป็นตัวแปรสุ่มแทนจำนวนสิ่งที่น่าสนใจในตัวอย่างกลุ่มที่ 1 ; X_2 เป็นตัวแปรสุ่มแทนจำนวนสิ่งที่น่าสนใจในตัวอย่างกลุ่มที่ 2 ; n_1 แทนขนาดตัวอย่างที่สุ่มจากประชากรที่ 1 ; n_2 แทนขนาดตัวอย่างที่สุ่มจากประชากรที่ 2 ; $\hat{p}_1$ แทนสัดส่วนของสิ่งที่น่าสนใจในตัวอย่างที่ 1 ซึ่ง $\hat{p}_1 = \frac{X_1}{n_1}$; $\hat{p}_2$ แทนสัดส่วนของสิ่งที่น่าสนใจในตัวอย่างที่ 2 ซึ่ง $\hat{p}_2 = \frac{X_2}{n_2}$; $Z_{\frac{\alpha}{2}}$ คือค่าควอนไทล์ที่ $1 - \frac{\alpha}{2}$ ของการแจกแจงปกติมาตรฐาน

4.2) วิธี Newcombe

Newcombe (1998) ได้นำเสนอวิธี Newcombe ซึ่งเป็นวิธีการประมาณแบบช่วงความเชื่อมั่นสำหรับค่าสัดส่วนทวินามสองประชากรที่อิสระกัน โดยพัฒนามาจากวิธี Score (Wilson, 1927) ซึ่งเป็นวิธีการที่นำมาประมาณแบบช่วงสำหรับค่าสัดส่วนทวินามสองประชากร

กำหนดให้ขอบเขตล่างและขอบเขตบนของช่วงความเชื่อมั่นวิธี Newcombe เขียนแทนด้วยสัญลักษณ์ L_N และ U_N ตามลำดับ จะได้ว่าช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของ p_1-p_2 โดยวิธี Newcombe (Newcombe, 1998) เป็นดังนี้

$$L_N \text{ คือ } (\hat{p}_1 - \hat{p}_2) - Z_{\frac{\alpha}{2}} \sqrt{\left(\frac{l_1(1-l_1)}{n_1} + \frac{u_2(1-u_2)}{n_2} \right)}$$

$$U_N \text{ คือ } (\hat{p}_1 - \hat{p}_2) + Z_{\frac{\alpha}{2}} \sqrt{\left(\frac{l_1(1-l_1)}{n_1} + \frac{u_2(1-u_2)}{n_2} \right)}$$

$$\begin{aligned} \text{เมื่อ } (l_1, u_1) \text{ คือ } & \left(\frac{\left(\hat{p}_1 + \frac{Z_{\frac{\alpha}{2}}^2}{2n_1} \right) \pm Z_{\frac{\alpha}{2}} \sqrt{\hat{p}_1(1-\hat{p}_1)/n_1 + Z_{\frac{\alpha}{2}}^2/4n_1^2}}{1 + \frac{Z_{\frac{\alpha}{2}}^2}{n_1}} \right) ; \\ (l_2, u_2) \text{ คือ } & \left(\frac{\left(\hat{p}_2 + \frac{Z_{\frac{\alpha}{2}}^2}{2n_2} \right) \pm Z_{\frac{\alpha}{2}} \sqrt{\hat{p}_2(1-\hat{p}_2)/n_2 + Z_{\frac{\alpha}{2}}^2/4n_2^2}}{1 + \frac{Z_{\frac{\alpha}{2}}^2}{n_2}} \right) ; \end{aligned}$$

X_1 เป็นตัวแปรสุ่มแทนจำนวนสิ่งที่น่าสนใจในตัวอย่างกลุ่มที่ 1 ; X_2 เป็นตัวแปรสุ่มแทนจำนวนสิ่งที่น่าสนใจในตัวอย่างกลุ่มที่ 2 ; n_1 แทนขนาดตัวอย่างที่สุ่มจากประชากรที่ 1 ; n_2 แทนขนาดตัวอย่างที่สุ่มจากประชากรที่ 2 ; $\hat{p}_1$ แทนสัดส่วนของสิ่งที่น่าสนใจในตัวอย่างที่ 1 ซึ่ง $\hat{p}_1 = \frac{X_1}{n_1}$; $\hat{p}_2$ แทนสัดส่วนของสิ่งที่น่าสนใจในตัวอย่างที่ 2 ซึ่ง

$\hat{p}_2 = \frac{X_2}{n_2}$; $Z_{\frac{\alpha}{2}}$ คือค่าควอนไทล์ที่ $1 - \frac{\alpha}{2}$ ของการแจกแจงปกติมาตรฐาน

4.3) วิธี Agresti-Caffo

Agresti and Caffo (2000) ได้นำเสนอวิธีการประมาณแบบช่วงความเชื่อมั่นสำหรับค่าสัดส่วนทวินามสองประชากรที่อิสระกันขึ้นมาใหม่โดยการเพิ่มลักษณะสิ่งที่น่าสนใจจำนวน 2 ค่า และสิ่งที่ไม่สนใจอีก 2 ค่าสำหรับแต่ละตัวอย่าง และนำแนวคิดของวิธีเบส์มาประยุกต์ใช้ร่วมกับการประมาณโดยวิธีของวาลด์ จะได้

$$\tilde{p}_i = \frac{(X_i + 1)}{(n_i + 2)} \text{ แทนสัดส่วนของสิ่งที่น่าสนใจลำดับที่ } i \text{ เมื่อ } i = 1, 2, \dots, n$$

โดยกำหนดให้ X_i เป็นตัวแปรสุ่มที่มีการแจกแจงทวินามที่มีพารามิเตอร์คือ n_i และ p_i โดยที่ตัวแปรสุ่ม P_i มีการแจกแจงก่อน (Prior Distribution) เป็นการแจกแจงบีตา และการแจกแจงแบบมีเงื่อนไขของ P_i เมื่อกำหนดค่าตัวแปรสุ่ม X_i เท่ากับ x_i มีการแจกแจงภายหลัง (Posterior Distribution) ซึ่งจะได้ $P_i | x_i$ มีการแจกแจงบีตา ที่มีพารามิเตอร์คือ $x_i + 1$ และ $n_i - x_i + 1$ ดังนั้นตัวประมาณแบบเบส์ (Bayes Estimator) ของพารามิเตอร์ p ภายใต้ฟังก์ชันการสูญเสียความคลาดเคลื่อนกำลังสองโดยมีค่าเฉลี่ยเป็น $E(P_i | x_i)$ ซึ่งเขียนแทนด้วยสัญลักษณ์ $\tilde{p}_i$

จะได้
$$\tilde{p}_i = E(P_i | x_i) = \frac{(X_i + 1)}{(n_i + 2)}$$

และความแปรปรวนของ $P_i | x_i$ เป็นดังนี้
$$\text{Var}(P_i | x_i) = \frac{\tilde{p}_i (1 - \tilde{p}_i)}{n_i + 3}$$

ดังนั้นจะได้
$$\tilde{p}_1 = \frac{(X_1 + 1)}{(n_1 + 2)} \quad \text{แทนสัดส่วนของสิ่งที่สนใจในตัวอย่างที่ 1}$$

และ
$$\tilde{p}_2 = \frac{(X_2 + 1)}{(n_2 + 2)} \quad \text{แทนสัดส่วนของสิ่งที่สนใจในตัวอย่างที่ 2}$$

กำหนดให้ขอบเขตล่างและขอบเขตบนของช่วงความเชื่อมั่นวิธี Agresti – Caffo เขียนแทนด้วยสัญลักษณ์ L_{AC} และ U_{AC} ตามลำดับ จะได้ว่าช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของ $p_1 - p_2$ โดยวิธี Agresti – Caffo (Agresti and Caffo, 2000) เป็นดังนี้

$$L_{AC} \text{ คือ } (\tilde{p}_1 - \tilde{p}_2) - \left[Z_{\frac{\alpha}{2}} \sqrt{\left(\frac{\tilde{p}_1(1-\tilde{p}_1)}{n_1+3} + \frac{\tilde{p}_2(1-\tilde{p}_2)}{n_2+3} \right)} \right]$$

$$U_{AC} \text{ คือ } (\tilde{p}_1 - \tilde{p}_2) + \left[Z_{\frac{\alpha}{2}} \sqrt{\left(\frac{\tilde{p}_1(1-\tilde{p}_1)}{n_1+3} + \frac{\tilde{p}_2(1-\tilde{p}_2)}{n_2+3} \right)} \right]$$

เมื่อ X_1 เป็นตัวแปรสุ่มแทนจำนวนสิ่งที่สนใจในตัวอย่างกลุ่มที่ 1 ; X_2 เป็นตัวแปรสุ่มแทนจำนวนสิ่งที่สนใจในตัวอย่างกลุ่มที่ 2 ; n_1 แทนขนาดตัวอย่างที่สุ่มจากประชากรที่ 1 ; n_2 แทนขนาดตัวอย่างที่สุ่มจากประชากรที่ 2 ; $\tilde{p}_1$ แทนสัดส่วนของสิ่งที่สนใจในตัวอย่างที่ 1 ซึ่ง $\tilde{p}_1 = \frac{(X_1 + 1)}{(n_1 + 2)}$; $\tilde{p}_2$ แทนสัดส่วนของสิ่งที่สนใจในตัวอย่างที่ 2

ซึ่ง $\tilde{p}_2 = \frac{(X_2 + 1)}{(n_2 + 2)}$; $Z_{\frac{\alpha}{2}}$ คือค่าควอนไทล์ที่ $1 - \frac{\alpha}{2}$ ของการแจกแจงปกติมาตรฐาน

4.4) วิธี KMS

Kulinskaya et al. (2010) ได้เสนอวิธีการประมาณแบบช่วงความเชื่อมั่นสำหรับค่าสัดส่วน ทวินามสองประชากรที่อิสระกันคือ วิธี KMS ได้แสดงการหาการประมาณความแปรปรวนที่ทำให้คงที่ของค่าผลต่างระหว่างค่าสัดส่วนทวินามสองกลุ่มที่อิสระกัน ในที่นี้พารามิเตอร์ที่ถูกเพิ่มคือ $\psi = Ap_1 + (1-A)p_2$ เมื่อ $0 \leq A \leq 1$ จากนั้นจึงหาค่าความแปรปรวนที่ถูกเปลี่ยนให้คงที่ (Variance Stabilizing Transformation) หรือเรียกว่า VST ของผลต่างระหว่างค่าสัดส่วนทวินามสองกลุ่มที่อิสระกัน ($\Delta_0 = p_1 - p_2$) ซึ่งจะทำให้ได้ความแปรปรวนคงที่และได้ตัวสถิติดังต่อไปนี้

$$T_A(\hat{\Delta}, \hat{\psi}, \Delta_0) = \sqrt{\frac{2N \frac{n_2}{N} \frac{n_1}{N}}{u}} \left(\arcsin\left(\frac{u\hat{\Delta} + v}{w}\right) - \arcsin\left(\frac{u\Delta_0 + v}{w}\right) \right)$$

กำหนดให้ขอบเขตล่างและขอบเขตบนของช่วงความเชื่อมั่นวิธี KMS เขียนแทนด้วยสัญลักษณ์ L_{KMS} และ U_{KMS} ตามลำดับ จะได้ว่าช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของ p_1-p_2 โดยวิธี KMS (Kulinskaya et al., 2010) เป็นดังนี้

$$L_{KMS} \text{ คือ } \frac{\hat{w}}{u} \sin \left(\arcsin \left(\frac{u\hat{\Delta}+\hat{v}}{\hat{w}} \right) - Z_{\frac{\alpha}{2}} \sqrt{\frac{u}{2N \frac{n_2}{N} \frac{n_1}{N}}} \right) - \frac{\hat{v}}{u}$$

$$U_{KMS} \text{ คือ } \frac{\hat{w}}{u} \sin \left(\arcsin \left(\frac{u\hat{\Delta}+\hat{v}}{\hat{w}} \right) + Z_{\frac{\alpha}{2}} \sqrt{\frac{u}{2N \frac{n_2}{N} \frac{n_1}{N}}} \right) - \frac{\hat{v}}{u}$$

$$\text{เมื่อ } 0 \leq A \leq 1, N=n_1+n_2 \quad ; \quad u=2 \left\{ \left(1-A\right)^2 \frac{n_2}{N} + A^2 \frac{n_1}{N} \right\} \quad ; \quad \hat{\Delta} = \frac{\left(X_1 + \frac{1}{2}\right)}{n_1+1} - \frac{\left(X_2 + \frac{1}{2}\right)}{n_2+1} \quad ;$$

$$\hat{\psi} = A \left(\frac{\left(X_1 + \frac{1}{2}\right)}{n_1+1} \right) + (1-A) \left(\frac{\left(X_2 + \frac{1}{2}\right)}{n_2+1} \right) \quad ; \quad \hat{v} = (1-2\hat{\psi}) \left(A - \frac{n_2}{N} \right) \quad ; \quad \hat{w} = \sqrt{2u\hat{\psi}(1-\hat{\psi}) + \hat{v}^2} \quad ;$$

X_1 เป็นตัวแปรสุ่มแทนจำนวนสิ่งที่น่าสนใจในตัวอย่างกลุ่มที่ 1 ; X_2 เป็นตัวแปรสุ่มแทนจำนวนสิ่งที่น่าสนใจในตัวอย่างกลุ่มที่ 2 ; n_1 แทนขนาดตัวอย่างที่สุ่มจากประชากรที่ 1 ; n_2 แทนขนาดตัวอย่างที่สุ่มจากประชากรที่ 2 ; $\hat{p}_1$ แทนสัดส่วนของสิ่งที่น่าสนใจในตัวอย่างที่ 1 ซึ่ง $\hat{p}_1 = \frac{X_1}{n_1}$; $\hat{p}_2$ แทนสัดส่วนของสิ่งที่น่าสนใจในตัวอย่างที่ 2 ซึ่ง

$$\hat{p}_2 = \frac{X_2}{n_2} \quad ; \quad Z_{\frac{\alpha}{2}} \text{ คือค่าควอนไทล์ที่ } 1-\frac{\alpha}{2} \text{ ของการแจกแจงปกติมาตรฐาน}$$

4.5) วิธี Recentered

Brown และ Li (2005) ได้สร้างช่วงความเชื่อมั่นสำหรับผลต่างระหว่างค่าสัดส่วนทวินามของสองกลุ่มที่อิสระกันแบบใหม่ โดยมีการสร้างพารามิเตอร์ใหม่ขึ้นมาคือ p ภายใต้ข้อสมมติ $p_1 - p_2 \geq 0$

$$\text{จะได้ } p = \frac{n_2 p_1 + n_1 p_2}{n_1 + n_2} \quad \text{ตัวประมาณของ } p \text{ คือ } \hat{p} = \frac{n_2 \hat{p}_1 + n_1 \hat{p}_2}{n_1 + n_2}$$

$$\text{จะได้ } p_1 \text{ และ } p_2 \text{ เท่ากับ } p_1 = p + \frac{n_1(p_1 - p_2)}{n_1 + n_2} \text{ และ } p_2 = p - \frac{n_2(p_1 - p_2)}{n_1 + n_2}$$

$$\text{จะได้ความแปรปรวนของ } \hat{p}_1 - \hat{p}_2 \text{ ดังนี้ } \text{Var}(\hat{p}_1 - \hat{p}_2) = \left(\frac{1}{n_1} + \frac{1}{n_2} \right) p(1-p) - \frac{(p_1 - p_2)^2}{n_1 + n_2}$$

ขอบเขตของพารามิเตอร์ p พิจารณาได้จากข้อกำหนด $p_1 - p_2 \geq 0$ และ $0 \leq p_1, p_2 \leq 1$ เมื่อแทนค่า

$p_1 = 1$ ใน p_1 จะได้ $p = 1 - \frac{(p_1 - p_2)n_1}{n_1 + n_2}$ และแทน $p_2 = 0$ ใน p_2 จะได้ $p = \frac{(p_1 - p_2)n_2}{n_1 + n_2}$ กล่าวคือ

พารามิเตอร์ที่กำหนดขึ้นใหม่นี้ มีค่าน้อยกว่า $1 - \frac{(p_1 - p_2)n_1}{n_1 + n_2}$ และมากกว่า $\frac{(p_1 - p_2)n_2}{n_1 + n_2}$ ซึ่ง $\hat{p}$ จะมี

ขอบเขตเช่นเดียวกับ p จากขอบเขตของ $\hat{p}$ ข้างต้นสามารถสร้างตัวประมาณใหม่แทน $\hat{p}$ คือ

$$\tilde{p} = \begin{cases} \frac{(\hat{p}_1 - \hat{p}_2)n_2}{n_1 + n_2} & ; \hat{p} < \frac{(\hat{p}_1 - \hat{p}_2)n_2}{n_1 + n_2} \\ \hat{p} & ; \frac{(\hat{p}_1 - \hat{p}_2)n_2}{n_1 + n_2} \leq \hat{p} \leq 1 - \frac{(\hat{p}_1 - \hat{p}_2)n_1}{n_1 + n_2} \\ 1 - \frac{(\hat{p}_1 - \hat{p}_2)n_1}{n_1 + n_2} & ; \hat{p} > 1 - \frac{(\hat{p}_1 - \hat{p}_2)n_1}{n_1 + n_2} \end{cases}$$

แทนค่า p ด้วย $\tilde{p}$ จะได้ความแปรปรวนของ $\hat{p}_1 - \hat{p}_2$ ดังต่อไปนี้

$$\text{Var}(\hat{p}_1 - \hat{p}_2) = \left(\left(\frac{1}{n_1} + \frac{1}{n_2} \right) \tilde{p}(1 - \tilde{p}) \right) - \frac{(p_1 - p_2)^2}{(n_1 + n_2)}$$

การสร้างช่วงความเชื่อมั่นสำหรับ $p_1 - p_2$ กรณีที่ไม่ทราบค่าความแปรปรวนของ $p_1 - p_2$ จึงใช้ตัวสถิติที่มีการแจกแจงแบบ t ด้วยองศาเสรีเท่ากับ $n_1 + n_2 - 2$ พิจารณาจาก

$$t = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\text{Var}(\hat{p}_1 - \hat{p}_2)}}$$

เมื่อแทนค่า $\text{Var}(\hat{p}_1 - \hat{p}_2)$ ในตัวสถิติที่มีการแจกแจงแบบ t จะได้ $p_1 - p_2$ ดังต่อไปนี้

$$p_1 - p_2 = \frac{\hat{p}_1 - \hat{p}_2}{1 + \frac{t^2}{n_1 + n_2}} \pm \frac{t \sqrt{\left(1 + \frac{t^2}{n_1 + n_2} \right) \left(\frac{1}{n_1} + \frac{1}{n_2} \right) \tilde{p}(1 - \tilde{p}) - \frac{(\hat{p}_1 - \hat{p}_2)^2}{n_1 + n_2}}}{1 + \frac{t^2}{n_1 + n_2}}$$

กำหนดให้ขอบเขตล่างและขอบเขตบนของช่วงความเชื่อมั่นวิธี KMS เขียนแทนด้วยสัญลักษณ์ L_R และ U_R ตามลำดับ จะได้ว่าช่วงความเชื่อมั่น $(1 - \alpha)100\%$ ของ $p_1 - p_2$ โดยวิธี Recentered (Brown and Li, 2005) เป็นดังนี้

$$L_R \text{ คือ } \frac{\hat{p}_1 - \hat{p}_2}{1 + \frac{t^2}{n_1 + n_2}} - \frac{t \sqrt{\left(1 + \frac{t^2}{n_1 + n_2} \right) \left(\frac{1}{n_1} + \frac{1}{n_2} \right) \tilde{p}(1 - \tilde{p}) - \frac{(\hat{p}_1 - \hat{p}_2)^2}{n_1 + n_2}}}{1 + \frac{t^2}{n_1 + n_2}}$$

$$U_R \text{ คือ } \frac{\hat{p}_1 - \hat{p}_2}{1 + \frac{t^2}{n_1 + n_2}} + \frac{t \sqrt{\left(1 + \frac{t^2}{n_1 + n_2}\right) \left(\frac{1}{n_1} + \frac{1}{n_2}\right) \tilde{p}(1 - \tilde{p}) - \frac{(\hat{p}_1 - \hat{p}_2)^2}{n_1 + n_2}}}{1 + \frac{t^2}{n_1 + n_2}}$$

เมื่อ X_2 เป็นตัวแปรสุ่มแทนจำนวนสิ่งที่น่าสนใจในตัวอย่างกลุ่มที่ 2 ; n_1 แทนขนาดตัวอย่างที่สุ่มจากประชากรที่ 1 ; n_2 แทนขนาดตัวอย่างที่สุ่มจากประชากรที่ 2 ; $\hat{p}_1$ แทนสัดส่วนของสิ่งที่น่าสนใจในตัวอย่างที่ 1 ; $\hat{p}_2$ แทนสัดส่วนของสิ่งที่น่าสนใจในตัวอย่างที่ 2 ; $t = t_{\frac{\alpha}{2}, n_1 + n_2 - 2}$ คือค่าควอนไทล์ที่ $1 - \frac{\alpha}{2}$ ของการแจกแจงที ที่องศาเสรีเท่ากับ

$n_1 + n_2 - 2$

5. เกณฑ์ที่ใช้ในการเปรียบเทียบประสิทธิภาพมี 2 เกณฑ์ คือ การตรวจสอบค่าความน่าจะเป็นครอบคลุมและการเปรียบเทียบค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นดังนี้

5.1 การตรวจสอบค่าความน่าจะเป็นครอบคลุม มีขั้นตอนดังนี้

5.1.1 การคำนวณค่าประมาณความน่าจะเป็นครอบคลุม ($\hat{c}$) สำหรับช่วงความเชื่อมั่น 95 % ทั้ง 5 วิธีสามารถคำนวณได้จากสูตรดังนี้

$$\hat{c} = \frac{\text{จำนวนครั้งที่ช่วงความเชื่อมั่นครอบคลุมค่าพารามิเตอร์ } p_1 - p_2}{M}$$

เมื่อ M คือจำนวนรอบในการทำซ้ำ ในที่นี้ $M = 2,000$ รอบ

5.1.2 การตรวจสอบค่าความน่าจะเป็นครอบคลุมสำหรับช่วงความเชื่อมั่น 95% ทั้ง 5 วิธี ทำได้โดยการทดสอบสมมติฐานทางสถิติดังนี้

กำหนดสมมติฐานเพื่อการทดสอบคือ

H_0 : ค่าความน่าจะเป็นครอบคลุมมีค่าไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด ($c \geq c_0$)

H_1 : ค่าความน่าจะเป็นครอบคลุมมีค่าต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด ($c < c_0$)

ตัวสถิติทดสอบคือ $Z = \frac{\hat{c} - c_0}{\sqrt{\frac{c_0(1-c_0)}{M}}}$ เมื่อ M คือจำนวนรอบในการทำซ้ำซึ่งเท่ากับ 2,000 รอบ

ซึ่งปฏิเสธสมมติฐานว่างเมื่อ $\frac{\hat{c} - c_0}{\sqrt{\frac{c_0(1-c_0)}{M}}} < -Z_\alpha$

หรือ $\hat{c} < c_0 - Z_\alpha \sqrt{\frac{c_0(1-c_0)}{M}}$ เมื่อกำหนดระดับนัยสำคัญสำหรับการทดสอบสมมติฐานข้างต้น

เท่ากับ 0.05 และค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด (c_0) คือ 0.95 จะได้ว่า

$$\hat{C} < 0.95 - 1.645 \sqrt{\frac{0.95(1-0.95)}{2,000}}$$

หรือ $\hat{C} < 0.9419$

ดังนั้นวิธีการประมาณที่ให้ค่าความน่าจะเป็นครอบคลุมไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด (0.95) เมื่อสอดคล้องกับสมการที่ (1) ดังนี้

$$\hat{C} \geq 0.9419 \quad (1)$$

5.2 การคำนวณค่าความยาวเฉลี่ยของช่วงความเชื่อมั่น (Average Length: AL)

เมื่อตรวจสอบแล้วพบว่าวิธีการประมาณใดให้ค่าความน่าจะเป็นครอบคลุมไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนดจึงจะนำมาคำนวณค่าความยาวเฉลี่ยของช่วงความเชื่อมั่น (AL) ซึ่งสามารถคำนวณได้โดยนำค่าความยาวของช่วงความเชื่อมั่นทุก ๆ ช่วงมาหาค่าเฉลี่ย จะได้

$$AL = \frac{\sum_{i=1}^{2,000} (U_i - L_i)}{2,000}$$

โดย L_i และ U_i เป็นขีดจำกัดล่างและบนตามลำดับของช่วงความเชื่อมั่นในรอบที่ i เมื่อ $i = 1, 2, \dots, 2,000$

6. สรุปผลการวิจัยในแต่ละสถานการณ์

ผลการวิจัย

การเปรียบเทียบประสิทธิภาพช่วงความเชื่อมั่น 95 % สำหรับผลต่างระหว่างค่าสัดส่วนทวินามสองกลุ่มที่อิสระกัน 5 วิธี ได้แก่ วิธี Asymptotic, CC วิธี Newcombe วิธี Agresti-Caffo วิธี KMS และวิธี Recentered โดยการพิจารณาว่าวิธีการประมาณค่าแบบช่วงวิธีใดเหมาะสมที่สุดในแต่ละสถานการณ์ที่กำหนด ซึ่งเกณฑ์ที่ใช้ในการเปรียบเทียบประสิทธิภาพ คือ การตรวจสอบค่าความน่าจะเป็นครอบคลุม และพิจารณาค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นซึ่งวิธีการประมาณแบบช่วงที่มีประสิทธิภาพดีจะต้องให้ค่าความน่าจะเป็นครอบคลุมไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนดและให้ค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นแคบที่สุด ซึ่งค่าประมาณสัมประสิทธิ์ความเชื่อมั่นและค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นทั้ง 5 วิธี แสดงดังตารางที่ 1 และ 2 ตามลำดับ

จากตารางที่ 1 แสดงค่าประมาณความน่าจะเป็นครอบคลุม ($\hat{C}$) ของช่วงความเชื่อมั่น 95% สำหรับผลต่างระหว่างค่าสัดส่วนทวินามสองกลุ่มที่อิสระกัน จากผลการตรวจสอบความน่าจะเป็นครอบคลุมในการประมาณค่าพารามิเตอร์ $p_1 - p_2$ โดยใช้เกณฑ์ตามที่กำหนดในสมการที่ (2) ซึ่งวิธีการประมาณใดให้ค่าประมาณความน่าจะเป็นครอบคลุมไม่ต่ำกว่า 0.9419 แสดงว่าวิธีนั้นให้ค่าความน่าจะเป็นครอบคลุมไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด ซึ่งผลการวิจัยพบว่าวิธี Newcombe วิธี Agresti-Caffo และวิธี Recentered ส่วนใหญ่มีแนวโน้มให้ค่าความน่าจะเป็นครอบคลุมไม่น้อยกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนดในทุกสถานการณ์ ส่วนวิธี Asymptotic, CC มีแนวโน้มให้ค่าความน่าจะเป็นครอบคลุมไม่น้อยกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนดในเกือบทุกสถานการณ์ ยกเว้นกรณี $p_1 - p_2$ เท่ากับ 0.6 ที่ขนาดตัวอย่าง (n_1, n_2) เท่ากับ (10, 10) และเมื่อ $p_1 - p_2$ เท่ากับ 0.8 ที่ขนาดตัวอย่าง (n_1, n_2) เท่ากับ (10, 10) และ (20, 20) จะให้ค่าความน่าจะเป็น

ครอบคลุมน้อยกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด และวิธี KMS มีแนวโน้มให้ค่าความน่าจะเป็นครอบคลุมไม่น้อยกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนดในเกือบทุกสถานการณ์ ยกเว้นกรณี p_1-p_2 เท่ากับ 0.8 ที่ขนาดตัวอย่าง (n_1, n_2) เท่ากับ (20, 20) และ (50, 50) จะให้ค่าความน่าจะเป็นครอบคลุมน้อยกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด ดังนั้นจะไม่พิจารณาเปรียบเทียบความยาวเฉลี่ยของช่วงความเชื่อมั่นสำหรับวิธี Asymptotic, CC และวิธี KMS ในกรณีดังกล่าวนี้

และเมื่อพิจารณาค่าในตารางที่ 2 พบว่าความยาวเฉลี่ยของช่วงความเชื่อมั่นทุกวิธีการประมาณจะมีแนวโน้มแคบลงเมื่อขนาดตัวอย่างเพิ่มขึ้นในทุกระดับของค่า p_1-p_2 นอกจากนี้เมื่อ p_1-p_2 มีแนวโน้มเพิ่มขึ้นในช่วงไม่เกิน 0.4 พบว่าทุกวิธีจะมีความยาวเฉลี่ยของช่วงความเชื่อมั่นเพิ่มขึ้นในทุกระดับของขนาดตัวอย่าง (n_1, n_2) และเมื่อ p_1-p_2 มีแนวโน้มเพิ่มขึ้นในช่วงมากกว่า 0.4 พบว่าส่วนใหญ่ทุกวิธีจะมีความยาวเฉลี่ยของช่วงความเชื่อมั่นลดลงในทุกระดับของขนาดตัวอย่าง (n_1, n_2) ยกเว้นวิธี Newcombe เมื่อ p_1-p_2 เพิ่มขึ้น พบว่าความยาวเฉลี่ยของช่วงความเชื่อมั่นมีแนวโน้มเพิ่มขึ้นเมื่อขนาดตัวอย่าง (n_1, n_2) เท่ากับ (10, 10) ในทุกระดับของค่า p_1-p_2 และเมื่อพิจารณาประสิทธิภาพของแต่ละวิธีการประมาณพบว่าวิธี Agresti-Caffo มีแนวโน้มให้ค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นแคบที่สุดในกรณีที่ p_1-p_2 เท่ากับ 0.2 ที่ตัวอย่างขนาด (n_1, n_2) เท่ากับ (10, 10) และ p_1-p_2 เท่ากับ 0.4, 0.6 ในทุกขนาดตัวอย่าง (n_1, n_2) ส่วนวิธี KMS มีแนวโน้มให้ค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นแคบที่สุดในกรณีที่ p_1-p_2 เท่ากับ 0.8 ที่ตัวอย่างขนาด (n_1, n_2) เท่ากับ (10, 10), (30, 30) และ (70, 70) นอกจากนี้วิธี Recentered มีแนวโน้มให้ค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นน้อยที่สุดในเกือบทุกระดับของขนาดตัวอย่าง (n_1, n_2) เมื่อผลต่างระหว่างค่าสัดส่วนทวินาม (p_1-p_2) เท่ากับ 0, 0.2 และ 0.8 ยกเว้นในกรณีที่ p_1-p_2 เท่ากับ 0.2 ที่ตัวอย่างขนาด (n_1, n_2) เท่ากับ (10, 10), p_1-p_2 เท่ากับ 0.4 และ 0.6 ในทุกขนาดตัวอย่าง (n_1, n_2) และ p_1-p_2 เท่ากับ 0.8 ที่ขนาดตัวอย่าง (n_1, n_2) เท่ากับ (10, 10), (30, 30) และ (70, 70)

ตารางที่ 1 ค่าประมาณความน่าจะเป็นครอบคลุมของช่วงความเชื่อมั่น 95% ของผลต่างระหว่างค่าสัดส่วน
ทวินามสองกลุ่มที่อิสระกันทั้ง 5 วิธี

ผลต่างระหว่างค่า สัดส่วนทวินาม ($p_1 - p_2$)	ขนาด ตัวอย่าง (n_1, n_2)	วิธีการประมาณ				
		Asymptotic, CC	Newcombe	Agresti- Caffo	KMS	Recentered
0	(10, 10)	0.9995	0.9860	0.9935	0.9710	0.9970
	(20, 20)	0.9885	0.9825	0.9845	0.9540	0.9785
	(30, 30)	0.9865	0.9810	0.9760	0.9590	0.9740
	(50, 50)	0.9785	0.9695	0.9645	0.9520	0.9745
	(70, 70)	0.9765	0.9635	0.9630	0.9510	0.9750
	(100, 100)	0.9740	0.9600	0.9595	0.9505	0.9720
0.2	(10, 10)	0.9700	0.9640	0.9710	0.9730	0.9760
	(20, 20)	0.9770	0.9550	0.9535	0.9630	0.9655
	(30, 30)	0.9765	0.9610	0.9530	0.9640	0.9700
	(50, 50)	0.9770	0.9585	0.9570	0.9630	0.9695
	(70, 70)	0.9730	0.9570	0.9560	0.9590	0.9725
	(100, 100)	0.9725	0.9585	0.9555	0.9595	0.9725
0.4	(10, 10)	0.9675	0.9665	0.9505	0.9590	0.9740
	(20, 20)	0.9760	0.9655	0.9480	0.9655	0.9720
	(30, 30)	0.9745	0.9670	0.9485	0.9640	0.9730
	(50, 50)	0.9735	0.9670	0.9565	0.9635	0.9715
	(70, 70)	0.9720	0.9685	0.9565	0.9575	0.9715
	(100, 100)	0.9690	0.9675	0.9525	0.9555	0.9710
0.6	(10, 10)	0.9370 [#]	0.9720	0.9425	0.9590	0.9680
	(20, 20)	0.9660	0.9830	0.9485	0.9575	0.9660
	(30, 30)	0.9660	0.9800	0.9485	0.9505	0.9690
	(50, 50)	0.9695	0.9785	0.9440	0.9455	0.9655
	(70, 70)	0.9625	0.9745	0.9485	0.9535	0.9700
	(100, 100)	0.9630	0.9700	0.9435	0.9435	0.9650
0.8	(10, 10)	0.8810 [#]	0.9950	0.9495	0.9470	0.9495
	(20, 20)	0.9195 [#]	0.9955	0.9530	0.9385 [#]	0.9530
	(30, 30)	0.9490	0.9930	0.9505	0.9500	0.9650
	(50, 50)	0.9695	0.9865	0.9465	0.9315 [#]	0.9545
	(70, 70)	0.9675	0.9840	0.9475	0.9450	0.9680
	(100, 100)	0.9625	0.9750	0.9505	0.9505	0.9685

หมายเหตุ : # คือค่าประมาณความน่าจะเป็นครอบคลุมน้อยกว่า 0.9419 แสดงว่าวิธีนี้จะให้ค่าความน่าจะเป็นครอบคลุมน้อยกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด ดังนั้นจะไม่พิจารณาเปรียบเทียบความยาวเฉลี่ยของช่วงความเชื่อมั่นสำหรับวิธีนี้

ตารางที่ 2 ค่าความยาวเฉลี่ยของช่วงความเชื่อมั่น 95% ของผลต่างระหว่างค่าสัดส่วนทวินามสองกลุ่มที่อิสระกันทั้ง 5 วิธี

ผลต่างระหว่างค่า สัดส่วนทวินาม ($p_1 - p_2$)	ขนาด ตัวอย่าง (n_1, n_2)	วิธีการประมาณ				
		Asymptotic, CC	Newcombe	Agresti-Caffo	KMS	Recentered
0	(10, 10)	0.6569	0.6206	0.5550	0.5342	0.4453*
	(20, 20)	0.4513	0.4196	0.3863	0.3759	0.3459*
	(30, 30)	0.3601	0.3338	0.3129	0.3068	0.2906*
	(50, 50)	0.2704	0.2508	0.2399	0.2363	0.2290*
	(70, 70)	0.2248	0.2088	0.2019	0.1998	0.1953*
	(100, 100)	0.1848	0.1725	0.1683	0.1669	0.1643*
0.2	(10, 10)	0.8429	0.7031	0.6336*	0.6437	0.6353
	(20, 20)	0.5657	0.4973	0.4637	0.4676	0.4621*
	(30, 30)	0.4513	0.4046	0.3832	0.3854	0.3821*
	(50, 50)	0.3401	0.3115	0.2994	0.3005	0.2986*
	(70, 70)	0.2832	0.2624	0.2541	0.2548	0.2537*
	(100, 100)	0.2335	0.2188	0.2132	0.2137	0.2129*
0.4	(10, 10)	0.8876	0.7854	0.6567*	0.6786	0.6885
	(20, 20)	0.5978	0.5568	0.4856*	0.4939	0.4957
	(30, 30)	0.4775	0.4529	0.4034*	0.4084	0.4087
	(50, 50)	0.3600	0.3471	0.3165*	0.3189	0.3187
	(70, 70)	0.2999	0.2914	0.2691*	0.2707	0.2705
	(100, 100)	0.2475	0.2419	0.2262*	0.2270	0.2269
0.6	(10, 10)	-	0.8352	0.6304*	0.6412	0.6607
	(20, 20)	0.5680	0.5818	0.4601*	0.4664	0.4701
	(30, 30)	0.4528	0.4676	0.3814*	0.3849	0.3860
	(50, 50)	0.3408	0.3530	0.2989*	0.3000	0.3001
	(70, 70)	0.2838	0.2935	0.2535*	0.2546	0.2545
	(100, 100)	0.2338	0.2414	0.2129*	0.2135	0.2132
0.8	(10, 10)	-	0.8364	0.5479	0.5255*	0.5479
	(20, 20)	-	0.5551	0.3817	-	0.3750*
	(30, 30)	0.3664	0.4337	0.3118	0.3046*	0.3053
	(50, 50)	0.2720	0.3172	0.2379	-	0.2344*
	(70, 70)	0.2258	0.2587	0.2006	0.1983*	0.1987
	(100, 100)	0.1857	0.2092	0.1679	0.1664	0.1660*

หมายเหตุ : * คือค่าความยาวเฉลี่ยของช่วงเชื่อมั่นน้อยที่สุดในแต่ละสถานการณ์

สรุปผล

วิธีการประมาณช่วงความเชื่อมั่น 95 % สำหรับผลต่างระหว่างค่าสัดส่วนทวินามสองกลุ่มที่อิสระกัน ซึ่งให้ค่าสัมประสิทธิ์ความเชื่อมั่นไม่น้อยกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด และให้ค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นน้อยที่สุดในสถานการณ์ต่าง ๆ แสดงดังตารางที่ 3

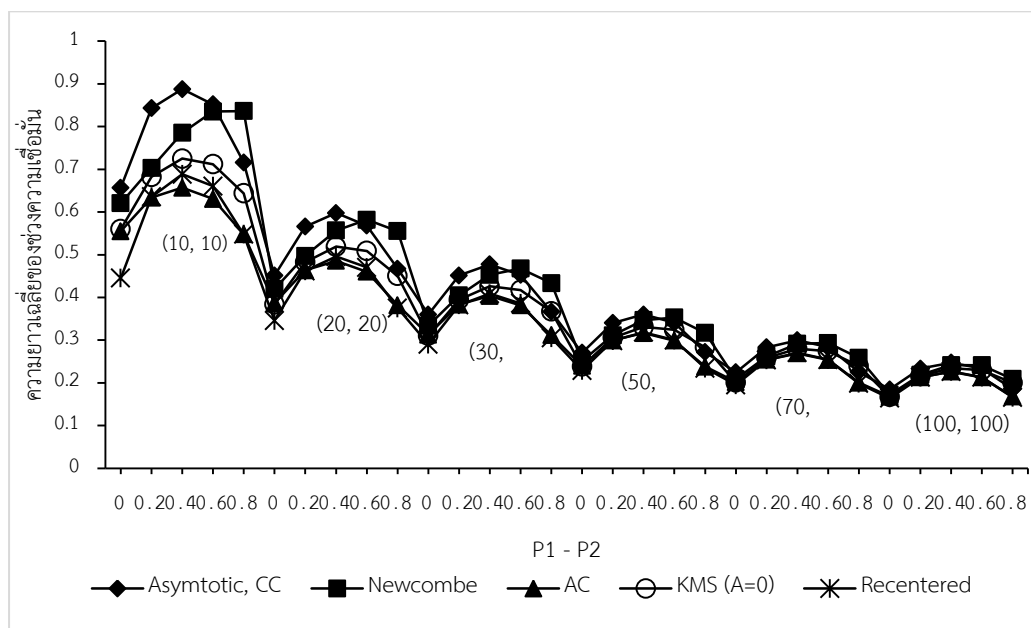
จากตารางที่ 3 พบว่าเมื่อผลต่างระหว่างค่าสัดส่วนทวินามเท่ากับ 0 , 0.2 และ 0.8 วิธี Recentered มีแนวโน้มให้ประสิทธิภาพดีที่สุดในทุกอันดับของขนาดตัวอย่าง (n_1, n_2) ยกเว้นกรณีขนาดตัวอย่าง (n_1, n_2) เท่ากับ (10, 10) และผลต่างระหว่างค่าสัดส่วนทวินามสองกลุ่มที่อิสระกัน $(p_1 - p_2)$ เท่ากับ 0.2 วิธี Agresti-Caffo จะให้ประสิทธิภาพดีที่สุด และถ้าผลต่างระหว่างค่าสัดส่วนทวินามสองกลุ่มที่อิสระกัน $(p_1 - p_2)$ เท่ากับ 0.8 วิธี KMS จะให้ประสิทธิภาพดีที่สุดเมื่อขนาดตัวอย่าง (n_1, n_2) เท่ากับ (10, 10), (30, 30) และ (70, 70) นอกจากนี้วิธี Agresti-Caffo มีแนวโน้มให้ประสิทธิภาพดีที่สุดในทุกอันดับของขนาดตัวอย่าง (n_1, n_2) เมื่อผลต่างระหว่างค่าสัดส่วนทวินามสองกลุ่มที่อิสระกัน $(p_1 - p_2)$ เท่ากับ 0.4 และ 0.6

ตารางที่ 3 วิธีการประมาณช่วงความเชื่อมั่น 95% สำหรับผลต่างระหว่างค่าสัดส่วนทวินามที่สองกลุ่มที่อิสระกัน ที่มีประสิทธิภาพดีที่สุดในแต่ละสถานการณ์

ผลต่างระหว่าง ค่าสัดส่วนทวิ นาม $(p_1 - p_2)$	ขนาดตัวอย่าง (n_1, n_2)					
	(10, 10)	(20, 20)	(30, 30)	(50, 50)	(70, 70)	(100, 100)
0	R	R	R	R	R	R
0.2	AC	R	R	R	R	R
0.4	AC	AC	AC	AC	AC	AC
0.6	AC	AC	AC	AC	AC	AC
0.8	KMS	R	KMS	R	KMS	R

หมายเหตุ : กำหนดให้ AC แทน วิธี Agresti-Caffo, KMS แทน วิธี KMS และ R แทน วิธี Recentered

นอกจากนี้จากภาพที่ 1 เมื่อพิจารณาความยาวเฉลี่ยของช่วงความเชื่อมั่นทั้ง 5 วิธี คือวิธี Asymptotic, CC วิธี Newcombe วิธี Agresti-Caffo วิธี KMS ในกรณี $A = 0$ และวิธี Recentered พบว่าเมื่อขนาดตัวอย่าง (n_1, n_2) เพิ่มขึ้น ค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นจะมีแนวโน้มลดลงในทุกอันดับของค่า $p_1 - p_2$



ภาพที่ 1 ความยาวเฉลี่ยของช่วงความเชื่อมั่นของผลต่างระหว่างค่าสัดส่วนทวินามสองกลุ่มที่อิสระกันของ 5 วิธี ในทุกขนาดตัวอย่าง

วิจารณ์

จากการเปรียบเทียบประสิทธิภาพช่วงความเชื่อมั่น 95 % สำหรับผลต่างระหว่างค่าสัดส่วนทวินามสองกลุ่มที่อิสระกัน 5 วิธี ได้แก่ วิธี Asymptotic, CC วิธี Newcombe วิธี Agresti-Caffo วิธี KMS และวิธี Recentered โดยเกณฑ์ที่ใช้ในการเปรียบเทียบประสิทธิภาพ คือ การตรวจสอบค่าความน่าจะเป็นครอบคลุมและค่าความยาวเฉลี่ยของช่วงความเชื่อมั่น พบว่าวิธี KMS ให้ค่าความน่าจะเป็นครอบคลุมที่ดี เมื่อขนาดตัวอย่าง (n_1, n_2) เพิ่มขึ้น ค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นจะมีแนวโน้มลดลงในทุกระดับของค่า $p_1 - p_2$ ดังนั้นวิธี KMS จึงเหมาะสมในกรณีตัวอย่างขนาดใหญ่เมื่อ $p_1 - p_2$ มีค่าเข้าใกล้ 1 และวิธี Agresti-Caffo พบว่าให้ค่าความน่าจะเป็นครอบคลุมที่ดีในทุกระดับค่า $p_1 - p_2$ แต่จะให้ค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นแคบที่สุดในกรณีที่ $p_1 - p_2$ เท่ากับ 0.2 ที่ตัวอย่างขนาด (n_1, n_2) เท่ากับ (10, 10) และ $p_1 - p_2$ เท่ากับ 0.4, 0.6 ในทุกขนาดตัวอย่าง (n_1, n_2) ซึ่งทั้ง 2 วิธีที่กล่าวมานี้สอดคล้องกับงานวิจัยของ Prendergast และ Staudte (2014) ส่วนวิธี Recentered แม้ว่าจะไม่สามารถให้ค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นแคบที่สุดในทุกสถานการณ์ แต่ยังให้ค่าความน่าจะเป็นครอบคลุมไม่น้อยกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนดในทุกสถานการณ์ และวิธี Recentered จะให้ค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นแคบที่สุดในกรณีตัวอย่างทั้งขนาดเล็กและใหญ่ เมื่อค่าของ $p_1 - p_2$ เท่ากับ 0 และ 0.2 ซึ่งสอดคล้องกับงานวิจัยของสุธารัตน์ (2553) และงานวิจัยของกนกกาญจน์ และคณะ (2558) นอกจากนี้จากวิธีที่กล่าวมาข้างต้นวิธีการประมาณช่วงความเชื่อมั่นสำหรับผลต่างระหว่างค่าสัดส่วนทวินามสองกลุ่มที่อิสระกันต่างมีความเหมาะสมในสถานการณ์ที่ต่างกัน ซึ่งอาจเป็นประโยชน์ต่อการวิจัยทางการแพทย์ เช่น การควบคุมสัดส่วนของผู้ป่วยให้อยู่ในเกณฑ์เป้าหมายเพื่อจะแก้ปัญหาทางคุณภาพชีวิต (พุทธชาติ และคณะ, 2559) หรือทางด้านการตลาด เช่น การควบคุมคุณภาพการบริการที่อาจมีผลต่อสัดส่วนการเลือกซื้อสินค้าของลูกค้าที่ตัดสินใจซื้อ เป็นต้น

เอกสารอ้างอิง

ภาษาไทย

- กนกกาญจน์ ศศิวิมลฤทธิ์, อภิญา หิรัญวงษ์ และบุญอ้อม โฉมที. (2558). การเปรียบเทียบวิธีการประมาณช่วงความเชื่อมั่นสำหรับผลต่างระหว่างสัดส่วนประชากรสองกลุ่มที่เป็นอิสระกันเมื่อตัวอย่างขนาดเล็ก. วารสารวิทยาศาสตร์และเทคโนโลยี, 23(2): 183-193.
- พุทธชาติ มากชุมนุม, นลินี พูลทรัพย์ และทิพาพร พงษ์เมษา. (2559). “ผลของการเยี่ยมบ้านผู้ป่วยโรคความดันโลหิตสูงโดยทีมสหสาขาวิชาชีพ”. Veridian E-Journal, Science and Technology Silpakorn University ปีที่ 3, ฉบับที่ 1 (มกราคม-กุมภาพันธ์ 2559): 13-33.
- สุดารัตน์ นิจสุนกิจ. (2553). การประมาณช่วงความเชื่อมั่นสำหรับผลต่างระหว่างสัดส่วนสองกลุ่มที่เป็นอิสระกัน: การเปรียบเทียบ 4 วิธี. วิทยานิพนธ์ปริญญาโท, มหาวิทยาลัยเกษตรศาสตร์.
- มหาวิทยาลัยเกษตรศาสตร์, คณะวิทยาศาสตร์, ภาควิชาสถิติ. 2560. หลักสถิติ. ครั้งที่ 1. ห้างหุ้นส่วนจำกัด วี.เจ.พรินต์ติ้ง, กรุงเทพฯ.

ภาษาต่างประเทศ

- Agresti, A., & Caffo, B. (2000). Simple and Effective Confidence Intervals for Proportions and Differences of Proportions Result From Adding Two Successes and Two Failures. The American Statistician. 54(4): 280–288.
- Brown, L. & Li, X. (2005). Confidence Intervals for Two Sample Binomial Distribution, Journal of Statistical Planning and Inference. 130: 359-375.
- Fleiss, J. L. (1981). Statistical Methods for Rates and Proportions, 2nd edn. Wiley. New York.
- Kulinskaya, E., Morgenthaler, S., and Staudte, R. G. (2010). Variance Stabilizing the Difference of Two Binomial Proportions. The American Statistician. 64(4): 350-356.
- Newcombe, R. (1998). Interval Estimation for the Difference between Independent Proportions: Comparison of Eleven Methods. Statistics in Medicine. 17: 873–890.
- Prendergast, L.A. & Staudte, R.G. (2014). Better Than You Think: Interval Estimators of the Difference of Binomial Proportions. Journal of Statistical Planning and Inference. 148: 38-48.
- Wald, A. & Wolfowitz, J. (1939). Confidence Limits for Continuous Distribution Functions. The Annals of Mathematical Statistics. 10: 105–118.
- Wilson, E. B. (1927). Probable Inference, the Law of Succession, and Statistical Inference. Journal of the American Statistical Association. 22: 209-212.