อัลกอริทึมใหม่โดยใช้ควอนไทล์ของภาวะน่าจะเป็นเพื่อสร้างช่วงความเชื่อมั่นสำหรับความแตกต่างระหว่างสองสัดส่วน

A Novel Algorithm Using Quantile Likelihood to Construct the Confidence Intervals for the Difference Between Two Proportions

โชติกา วุฒิสาร (Chotikar Vuthisarn)* เบญจมาศ ตุลยนิติกุล (Benjamas Tulyanitikul)**
ดร.พัทธ์ชนก ศรีสุรเดชชัย (Dr.Patchanok Srisuradetchai)**

บทคัดย่อ

งานวิจัยนี้ผู้วิจัยได้เสนออัลกอริทึมใหม่แทนการวิเคราะห์เชิงสถิติทางตรงเพื่อสร้างช่วงความเชื่อมั่นสำหรับความแตกต่างระหว่างสองสัดส่วน นั่นคือ อัลกอริทึมควอนไทล์ของภาวะน่าจะเป็น (QL) โดยนำหลักการของช่วงความเชื่อมั่นโดยวิธีบูตสแตรปภายใต้ตัวประมาณมัธยฐานไม่เอนเอียง (MUE) และอัลกอริทึมการเลือกตัวอย่างแบบเลือกซ้ำที่สำคัญ (SIR) มาปรับใช้ แล้วนำไปเปรียบเทียบกับวิธีวัลด์ วิธีไฮบริด วิธีเบส์ วิธี MUE และ อัลกอริทึม SIR ผลการศึกษาพบว่าทั้ง 5 วิธีให้ผลดีกว่าวิธีวัลด์ในทุกลักษณะประชากร และช่วงที่ได้จากอัลกอริทึม QL ให้ค่าความน่าจะเป็นคุ้มรวมใกล้เคียงกับวิธี SIR วิธีเบส์ วิธี MUE และวิธีไฮบริด ในหลายสถานการณ์ นอกจากนี้อัลกอริทึม QL นี้เกิดจากการใช้พารามิเตอร์โดยตรง ดังนั้น ช่วงความเชื่อมั่นของ $p_1 - p_2$ ที่ได้ จะอยู่ในช่วง $[-1,1]$ เสมอ และสามารถหาช่วงความเชื่อมั่นในกรณีที่ $\hat{p}_1$ และ $\hat{p}_2$ เท่ากับศูนย์ได้ โดยสรุปแล้ว ช่วงความเชื่อมั่นโดยอัลกอริทึมควอนไทล์ของภาวะน่าจะเป็นสามารถสร้างช่วงความเชื่อมั่นได้ดีใกล้เคียงช่วงที่ดีอย่างวิธีไฮบริด วิธีเบส์ และวิธี MUE ได้ในหลายกรณีที่ศึกษา

ABSTRACT

This research proposes a novel algorithm instead of direct statistical analysis for construction the confidence intervals for the difference between two proportions which is named Quantile likelihood (QL). This algorithm, using the concept of fully specified bootstrap based on the median unbiased estimator and SIR algorithm, is applied to construct the confidence intervals and then compare to those obtained from Wald, hybrid, Bayes, MUE and SIR algorithm confidence intervals. The result shows that theses 5 method are preferable than the Wald method in all population characteristics and confidence intervals obtained from QL algorithm give the similar coverage probability to those of SIR algorithm, MUE and hybrid method in several situations. Moreover, QL algorithm directly used sampled of the parameters. Therefore, confidence intervals of $p_1 - p_2$ are definitely in $[-1,1]$ and can construct a confidence interval of $p_1 - p_2$ in case that $\hat{p}_1$ and $\hat{p}_2$ are equal to zero. In short, QL algorithm can produce as good a confidence interval similar to hybrid Bayes and MUE method which are the efficient method.

คำสำคัญ: อัลกอริทึมควอนไทล์ของภาวะน่าจะเป็น อัลกอริทึมการเลือกสุ่มการเลือกตัวอย่างแบบเลือกซ้ำที่สำคัญ ความน่าจะเป็นคุ้มรวม

Keywords: Quantile Likelihood algorithm, Sampling Importance Resampling algorithm, coverage probability

** นักศึกษา หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาคณิตศาสตร์และสถิติ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์*

*** ผู้ช่วยศาสตราจารย์ สาขาวิชาคณิตศาสตร์และสถิติ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์*

บทนำ

การประมาณค่าแบบช่วงสำหรับสัดส่วนประชากร (Proportion) เป็นที่สนใจศึกษาและเปรียบเทียบอย่างกว้างขวาง โดยพารามิเตอร์ที่สนใจ คือ p เป็นพารามิเตอร์สัดส่วนของการแจกแจงทวินาม แทน ความน่าจะเป็นของการเกิดเหตุการณ์ที่สนใจ ที่มีฟังก์ชันมวลความน่าจะเป็น $f(y)={}^{n}C_{y}p^{y}(1-p)^{n-y}$, $y=0,1,\ldots,n$ โดยที่ตัวแปรสุ่ม Y แทน จำนวนครั้งของการเกิดเหตุการณ์ที่สนใจจากการทดลองสุ่มทั้งหมด n ครั้ง หรือ $Y: B(n,p)$

สำหรับช่วงความเชื่อมั่นของสัดส่วนสองกลุ่มประชากรโดยใช้แนวคิดของการแจกแจงเชิงเส้นกำกับ หรือ วิธีวัลด์ (Wald) ซึ่งเป็นที่รู้จักและนิยมนำมาใช้กันอย่างแพร่หลาย เมื่อกำหนดให้ $Y_1: B(n_1,p_1)$ และ $Y_2: B(n_2,p_2)$ จะได้ว่าช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของ p_1-p_2 คือ $\hat{p}_1-\hat{p}_2\pm Z_{1-\frac{\alpha}{2}}\sqrt{\hat{p}_1(1-\hat{p}_1)/n_1+\hat{p}_2(1-\hat{p}_2)/n_2}$ โดยที่ $\hat{p}_i=y_i/n_i$; $i=1,2$ และ $Z_{1-\alpha/2}$ แทน ควอนไทล์ที่ $1-\alpha/2$ ของการแจกแจงปรกติมาตรฐาน แต่อย่างไรก็ตามจากการทบทวนวรรณกรรมพบว่าช่วงความเชื่อมั่นที่ได้จากวิธีวัลด์มีแนวโน้มที่ให้ค่าความน่าจะเป็นคุ้มรวม (Coverage Probability: CP) ต่ำกว่าสัมประสิทธิ์ความเชื่อมั่น (Confidence Coefficient) ที่กำหนด เมื่อตัวอย่างขนาดเล็ก หรือ กรณีขนาดตัวอย่างเล็กและค่า p เข้าใกล้ 0 หรือ 1 จึงมีนักวิจัยเสนอการสร้างช่วงความเชื่อมั่นประเภทนี้ในวิธีต่าง ๆ หลายวิธี เช่น ช่วงความเชื่อมั่นแบบวิธีไฮบริด (Hybrid) ของ Newcombe R G. (1998) ช่วงความเชื่อมั่นโดยวิธีแบบเบส์ที่ใช้การแจกแจงก่อน (Prior Distribution) คือ การแจกแจงแบบเอกรูป (Uniform Distribution) เสนอโดย Carlin B P, Louis T A. (2009) ช่วงความเชื่อมั่นโดยวิธีบูตสแตรปภายใต้ตัวประมาณมัธยฐานไม่เอนเอียง (MUE) ของ Yan et al. (2009) และช่วงความเชื่อมั่นโดยอัลกอริทึม SIR นำเสนอโดย โชติกา วุฒิสาร รมิดา ศรีเหรา และ พัทธ์ชนก ศรีสุรเดชชัย (รอตีพิมพ์) ที่ใช้หลักการอัลกอริทึมการเลือกตัวอย่างแบบเลือกซ้ำที่สำคัญ (Sampling Importance Resampling: SIR) ที่เสนอโดย Rubin D B. (1987, 1988) ซึ่งเป็นอัลกอริทึมที่มีจุดประสงค์หลักในการสร้างตัวอย่างที่มีการแจกแจงภายหลัง (Posterior Distribution) และนำตัวอย่างที่ได้ไปหาช่วงความมั่นสำหรับความแตกต่างระหว่างสองสัดส่วน เป็นต้น

ในงานวิจัยนี้ผู้วิจัยได้เสนอช่วงความเชื่อมั่นสำหรับความแตกต่างระหว่างสองสัดส่วนโดยอัลกอริทึมควอนไทล์ของภาวะน่าจะเป็น (Quantile Likelihood: QL) ซึ่งเป็นการนำหลักการของวิธี MUE ของ Yan et al. (2009) และวิธีอัลกอริทึม SIR มาปรับใช้ในการสร้างช่วงความเชื่อมั่นสำหรับความแตกต่างระหว่างสองสัดส่วน และเนื่องจากวิธีอัลกอริทึมควอนไทล์ของภาวะน่าจะเป็นเกิดจากพารามิเตอร์โดยตรง ดังนั้นช่วงความเชื่อมั่นของ p_1-p_2 โดยอัลกอริทึมนี้จะได้ขีดจำกัดบนและขีดจำกัดล่างอยู่ในช่วง $[-1,1]$ เสมอ ในขณะที่ช่วงความเชื่อมั่นสำหรับความแตกต่างระหว่างสองสัดส่วนที่ได้จากวิธีแบบวัลด์ วิธีแบบไฮบริด และ วิธีแบบเบส์ ไม่ได้จำกัดอยู่ในช่วง $[-1,1]$ นอกจากนี้วิธีอัลกอริทึมนี้ยังสามารถคำนวณหาช่วงความเชื่อมั่นได้ในกรณีที่ค่าตัวแปรสุ่ม $Y_1: B(n_1,p_1)$ และ $Y_2: B(n_2,p_2)$ มีค่าเป็นศูนย์ ซึ่งช่วงความเชื่อมั่นวิธีวัลด์จะไม่สามารถคำนวณได้

วัตถุประสงค์การวิจัย

เพื่อพัฒนาอัลกอริทึมใหม่สำหรับช่วงความเชื่อมั่นสำหรับความแตกต่างระหว่างสองสัดส่วน ศึกษาและเปรียบเทียบประสิทธิของการประมาณช่วงความเชื่อมั่นของ p_1-p_2 ทั้งหมด 6 วิธีได้แก่ ช่วงความเชื่อมั่นโดยวิธีวัลด์ (Wald: WL) วิธีไฮบริด(Hybrid: HYB) วิธีเบส์ (Bayes : BAY) วิธี MUE วิธีอัลกอริทึม SIR และวิธีอัลกอริทึม QL

วิธีการวิจัย

กำหนดให้ $Y_1 : B(n_1, p_1)$ และ $Y_2 : B(n_2, p_2)$ โดยที่ $\hat{p}_i$ คือ สัดส่วนตัวอย่างที่มีลักษณะที่สนใจจากประชากรกลุ่มที่ i และ n_i คือ ขนาดตัวอย่างของกลุ่มที่ i เมื่อ $i = 1,2$ ในการศึกษาจะมีขั้นตอนหลัก ๆ เป็นดังนี้

1) ทบทวนวรรณกรรมที่เกี่ยวข้องกับช่วงความเชื่อมั่นสำหรับความแตกต่างระหว่างสองสัดส่วน
2) ศึกษาแนวคิดของช่วงความเชื่อมั่นสำหรับความแตกต่างระหว่างสองสัดส่วนโดยวิธี MUE และวิธีอัลกอริทึม SIR พร้อมทั้งนำเสนออัลกอริทึมใหม่สำหรับสร้างช่วงความเชื่อมั่นสำหรับความแตกต่างระหว่างสองสัดส่วนในงานวิจัยนี้ ซึ่งก็คือ ช่วงความเชื่อมั่นสำหรับความแตกต่างระหว่างสองสัดส่วนโดยอัลกอริทึมควอนไทล์ของภาวะน่าจะเป็น (Quantile Likelihood: QL)
3) ศึกษาเชิงการจำลองโดยแบ่งเป็น 4 สถานการณ์ ดังนี้

สถานการณ์ที่ 1: ผลต่าง $p_1 - p_2$ มีค่าน้อยมากและประชากรทั้งสองสมมาตร คือ $p_1 = p_2 = 0.5$

สถานการณ์ที่ 2: ผลต่าง $p_1 - p_2$ มีค่าน้อยมากและประชากรทั้งสองเบ้ คือ $p_1 = p_2 = 0.05$

สถานการณ์ที่ 3: ผลต่าง $p_1 - p_2$ มีค่าในระดับกลาง คือ $p_1 = 0.75, p_2 = 0.25$

สถานการณ์ที่ 4: ผลต่าง $p_1 - p_2$ มีค่าในระดับสูง คือ $p_1 = 0.95, p_2 = 0.05$

ในแต่ละสถานการณ์กำหนดขนาดตัวอย่าง n_i ที่ใช้กรณีศึกษาของทั้งสองประชากรมีค่าเป็น 10, 30, 50 และ 100

4) จำลองค่า y_1 และ y_2 จาก $B(n_1, p_1)$ และ $B(n_2, p_2)$ แล้วคำนวณช่วงความเชื่อมั่น 95% ของความแตกต่างระหว่างสองสัดส่วนตามสูตร ดังนี้

4.1 วิธีวัลด์: $\hat{p}_1 - \hat{p}_2 \pm Z_{1-\frac{\alpha}{2}} \sqrt{\hat{p}_1(1-\hat{p}_1)/n_1 + \hat{p}_2(1-\hat{p}_2)/n_2}$

4.2 วิธีไฮบริด: $LCL = (\hat{p}_1 - \hat{p}_2) - Z_{\frac{\alpha}{2}} \sqrt{\frac{l_1(1-l_1)}{n_1} + \frac{u_2(1-u_2)}{n_2}}$ และ $UCL = (\hat{p}_1 - \hat{p}_2) + Z_{\frac{\alpha}{2}} \sqrt{\frac{u_1(1-u_1)}{n_1} + \frac{l_2(1-l_2)}{n_2}}$

โดยที่ ค่า l_i และ u_i คือ ค่า p_i ที่ได้จากการแก้สมการ $Z_{\frac{\alpha}{2}} = |\hat{p}_i - p_i| / \sqrt{\frac{p_i(1-p_i)}{n_i}}$ และ $u_i > l_i,\ i = 1,2$

4.3 วิธีเบส์: $(\hat{p}_1 - \hat{p}_2) \pm Z_{\alpha/2}\sqrt{V(\hat{p}_1) + V(\hat{p}_2)}$ โดยที่ $\hat{p}_i = \frac{y_i + 1}{n_i + 2}$ และ $V(\hat{p}_i) = \frac{\hat{p}_i(1-\hat{p}_i)}{n_i + 3}$ $, i = 1,\ 2$

4.4 วิธี MUE : เป็นวิธีประมาณค่าพารามิเตอร์สัดส่วน p_1 และ p_2 ที่ทำให้ค่าสังเกต y_1 และ y_2 เป็นค่ามัธยฐานตามลำดับ โดยใช้แนวคิดบูตสแตรปแบบอิงพารามิเตอร์ (Parametric Bootstrap) โดยมีขั้นตอนดังนี้

กำหนดให้ δ แทน ความแตกต่างของสัดส่วน จะได้ว่า $\delta = p_1 - p_2 \in [-1,1]$

และ $\tilde{p}_g$ คือ ตัวประมาณมัธยฐานไม่เอนเอียง ของ p_g จะได้ว่า

$$\tilde{p}_g = \frac{(\tilde{p}_g^L + \tilde{p}_g^U)}{2} \quad ; g = 1,2$$

โดยที่ $\Pr(Y_g \ge y_g \mid p_g = \tilde{p}_g^L) \ge 0.5$ และ $\Pr(Y_g \ge y_g \mid p_g = \tilde{p}_g^U) \ge 0.5$

จากความสัมพันธ์ระหว่างการแจกแจงสะสมเบต้าและทวินาม (Cumulative Beta and Binomial Distribution) จะได้ว่า $\tilde{p}_g^L = F^{-1}(0.5 \mid \alpha = y_g, \beta = n_g - y_g + 1)$ และ $\tilde{p}_g^U = F^{-1}(0.5 \mid \alpha = y_g + 1, \beta = n_g - y_g)$

เมื่อ $\tilde{p}_g^L$ คือ ค่าควอนไทล์ที่ 0.5 ของการแจกแจงเบต้าที่ $\alpha = y_g$ และ $\beta = n_g - y_g + 1$ และ

$\tilde{p}_g^U$ คือ ค่าควอนไทล์ที่ 0.5 ของการแจกแจงเบต้าที่ $\alpha = y_g + 1$ และ $\beta = n_g - y_g$

เราจะได้ $\tilde{p}_1$ และ $\tilde{p}_2$ เป็นตัวประมาณมัธยฐานไม่เอนเอียงสำหรับ p_1 และ p_2 ตามลำดับ จากนั้นแทนในฟังก์ชันมวลความน่าจะเป็นร่วม (Joint Probability Mass Function) ซึ่งก็คือ

$$\Pr\left(y_1, y_2 \mid \hat{p}_1, \hat{p}_2\right) = \binom{n_1}{y_1}\binom{n_2}{y_2}\hat{p}_1^{y_1}\left(1-\hat{p}_1\right)^{n_1-y_1}\hat{p}_2^{y_2}\left(1-\hat{p}_2\right)^{n_2-y_2} \qquad (*)$$

กำหนดให้ y_g^* คือ ตัวอย่างบูตสแตรปแบบอิงพารามิเตอร์ (Parametric Bootstrap Sample) $Y_1^* : B(y_1, \hat{p}_1)$ และ $Y_2^* : B(y_2, \hat{p}_2)$ โดยที่ y_1^*, y_2^* จะมีจำนวนทั้งหมด n_1+1 และ n_2+1 ตามลำดับ และแทนในสมการที่ (*) ซึ่งเราจะสามารถหา $\hat{\delta} = \hat{p}_1 - \hat{p}_2$ ได้ทั้งหมด $(n_1+1)(n_2+1)$ จำนวน จากนั้นเรียง $\hat{\delta}$ จากน้อยไปมาก แล้วคำนวณความน่าจะเป็นสะสม (Cumulative Probability) จากนั้นสร้างช่วงความเชื่อมั่น $100(1-\alpha)\%$ สำหรับ $\hat{\delta}$ โดยการประมาณค่าในช่วงเชิงเส้น (Linear Interpolation) คือ $\hat{\delta}_q = \dfrac{\hat{\delta}_L(\pi_U - q) + \hat{\delta}_U(q - \pi_L)}{(\pi_U - \pi_L)}$ เมื่อ π_U คือ ควอนไทล์ใกล้สุดที่มากกว่า q และ π_L คือ ควอนไทล์ใกล้สุดที่น้อยกว่า q โดยที่ $q = \alpha/2,\ 1-\alpha/2$

ดังนั้น ช่วงความเชื่อมั่น $100(1-\alpha)\%$ สำหรับ $\hat{\delta}$ คือ

$$LCL = \hat{\delta}_{\alpha/2} = \frac{\hat{\delta}_L(\pi_U - \alpha/2) + \hat{\delta}_U(\alpha/2 - \pi_L)}{(\pi_U - \pi_L)} \text{ และ } UCL = \hat{\delta}_{1-\alpha/2} = \frac{\hat{\delta}_L(\pi_U - (1-\alpha/2)) + \hat{\delta}_U((1-\alpha/2) - \pi_L)}{(\pi_U - \pi_L)} \qquad (**)$$

4.5 วิธีอัลกอริทึม SIR : ขั้นตอนในการหาช่วงความเชื่อมั่นสำหรับความแตกต่างระหว่างสองสัดส่วนดังนี้

กำหนดค่าสังเกต y_1 และ y_2 จากการทดลองทวินาม $B(n_1, p_1)$ และ $B(n_2, p_2)$ โดยที่ n_1 และ n_2 เป็นค่าที่ทราบและ p_1 และ p_2 เป็นพารามิเตอร์ที่สนใจที่ไม่ทราบ และให้ การแจกแจงก่อนของ p_1 และ p_2 มีการแจกแจงเอกรูปบนช่วง (0,1)

4.5.1 สุ่มตัวอย่าง $\{(p_1, p_2)_1, \ldots, (p_1, p_2)_m\}$ จากการแจกแจงก่อน โดยที่ p_i และ p_j ใด ๆ เป็นอิสระต่อกัน

4.5.2 คำนวณน้ำหนักให้แต่ละ $(p_1, p_2)_i$ จาก $w((p_1, p_2)_i) = L((p_1, p_2)_i) / \sum_{i=1}^{m} L((p_1, p_2)_i)$

โดยที่ $L(p_1, p_2) = p_1^{y_1}(1-p_1)^{n_1-y_1} p_2^{y_2}(1-p_2)^{n_2-y_2}$

4.5.3 สุ่มตัวอย่าง $\{(p_1, p_2)_1^*, \ldots, (p_1, p_2)_n^*\}$ (ขนาด n) แบบคืนที่จาก $\{(p_1, p_2)_1, \ldots, (p_1, p_2)_m\}$ ด้วยความน่าจะเป็น $w((p_1, p_2)_i)$ ในขั้นที่ 2)

4.5.4 คำนวณ $\{(p_1 - p_2)_1^*, \ldots, (p_1 - p_2)_n^*\}$ แล้วหาควอนไทล์ที่ $\alpha/2$ และ $1-\alpha/2$ ซึ่งคือ ขีดจำกัดล่างและบนของช่วงความเชื่อมั่น $(1-\alpha)100\%$

ในขั้นตอนวิธีข้างต้นนี้ ผู้วิจัยได้ทดลองหาค่า (m, n) ที่เหมาะสมโดยลองผิดลองถูก (Trial and Error) พบว่า $m = 5{,}000$ และ $n = 3{,}000$ จะให้ช่วงความเชื่อมั่นที่มีความน่าจะคุ้มรวมสูงแต่ใช้เวลาไม่มาก (น้อยกว่า 10 วินาที)

5) เปรียบเทียบค่าประมาณความน่าจะเป็นคุ้มรวม ความยาวช่วง และเปอร์เซ็นต์ของขีดจำกัดบนและล่างของช่วงที่สูงกว่า +1 และต่ำกว่า –1 ที่ได้จาก 5 วิธีที่กล่าวในข้างต้น และวิธีอัลกอริทึม QL ที่นำเสนอขึ้นในงานวิจัยครั้งนี้ โดยใช้การจำลองมอนติคาร์โลซึ่งมีจำนวนทำซ้ำ 5,000 ครั้ง

6) ทดสอบว่าช่วงที่ได้จะ CP เท่ากับ .95 หรือไม่ โดยพิจารณาจากทดสอบสมมติฐาน $H_0 : \text{CP} \geq 0.95$ ซึ่งเป็นไปตามนิยามที่ให้ไว้ใน Rohde (2014) ในที่นี้กำหนดระดับนัยสำคัญเท่ากับ .01 หากค่า CP จากการจำลองมอนติคาร์โลมากกว่าหรือเท่ากับ $0.95 - z_{.99}\sqrt{0.95(0.05)/5000} \approx 0.94$ จะไม่สามารถปฏิเสธสมมติฐานว่างได้และสรุปว่าช่วงความเชื่อมั่นนี้มี CP เท่ากับ .95 ที่ระดับนัยสำคัญ .01

ผลการวิจัย

ผลการวิจัยจะแบ่งออกเป็น 2 ส่วน คือ ช่วงความเชื่อมั่นโดยอัลกอริทึม QL ที่นำเสนอ และผลการศึกษาเชิงจำลอง

1. ขั้นตอนวิธีที่ใช้หาช่วงความเชื่อมั่น

นำเข้า: ค่าสังเกต y_1 และ y_2 จากการทดลองทวินาม $B(n_1, p_1)$ และ $B(n_2, p_2)$ โดยที่ n_1 และ n_2 เป็นค่าที่ทราบ

อัลกอริทึม:

1) คำนวณภาวะน่าจะเป็น

$$L_1 = L(p_1) = \binom{n_1}{y_1} p_1^{y_1} \left(1 - p_1^{y_1}\right)^{n_1 - y_1} \quad \text{และ} \quad L_2 = L(p_2) = \binom{n_2}{y_2} p_2^{y_2} \left(1 - p_2^{y_2}\right)^{n_2 - y_2}$$

2) คำนวณภาวะน่าจะเป็นร่วม

$$L(p_1, p_2) = L_1 L_2 = \binom{n_1}{y_1}\binom{n_2}{y_2} p_1^{y_1} \left(1 - p_1^{y_1}\right)^{n_1 - y_1} p_2^{y_2} \left(1 - p_2^{y_2}\right)^{n_2 - y_2} \qquad (***)$$

3) ในขั้นตอนนี้เป็นการนำแนวคิดของวิธีบูตสแตรปภายใต้ตัวประมาณมัธยฐานไม่เอนเอียง (MUE) จะสังเกตได้ว่า วิธีบูตสแตรปภายใต้ตัวประมาณมัธยฐานไม่เอนเอียงจะใช้วิธีการเปลี่ยนค่าของ $Y_1^* : B(y_1, \hat{p}_1)$ และ $Y_2^* : B(y_2, \hat{p}_2)$ ทุกค่าที่เป็นไปได้ทั้งหมด แต่สำหรับอัลกอริทึมควอนไทล์ของภาวะน่าจะเป็นนี้ เราจะเปลี่ยน $p_1, p_2 \in [0,1]$ โดยเพิ่มขึ้นคราวละ 0.005 แทน ซึ่งจะนำคู่อันดับ (p_1, p_2) ทุกคู่ที่เป็นไปได้แทนในสมการ (***) เพื่อคำนวณภาวะน่าจะเป็นร่วม $L(p_1, p_2)$ พร้อมทั้งคำนวณความแตกต่างระหว่างสองสัดส่วน $\delta = p_1 - p_2$

4) เนื่องจากค่าที่ได้จากการคำนวณในสมการ (***) นั้น คือค่าภาวะน่าจะเป็นร่วมของ p_1 และ p_2 เราจึงคำนวณน้ำหนักเพื่อให้ได้ความน่าจะเป็นซึ่งวิธีดังกล่าวเป็นการนำแนวคิดจากอัลกอริทึมการเลือกตัวอย่างแบบเลือกซ้ำที่สำคัญ (SIR) โดยมีสูตรการคำนวณดังนี้

$$L^* = \frac{L(p_1, p_2)}{\sum L(p_1, p_2)}$$

ตัวอย่างที่ 1 การคำนวณ กำหนดให้ y_1 และ y_2 มีการแจกแจงทวินาม $B(n_1 = 50, p_1 = 0.75)$ และ $B(n_2 = 100, p_2 = 0.25)$

ตารางที่ 1 แสดงการคำนวณความน่าจะเป็น (L^*)

ตัวอย่าง	p_1	p_2	$\delta = p_1 - p_2$	$L(p_1, p_2) = L_1 L_2$	$L^* = \frac{L(p_1, p_2)}{\sum L(p_1, p_2)}$
1	0	0	0	0	0
2	0	0.005	-0.005	0	0
M			M		M
200	0	0.995	-0.995	0	0
201	0	1	-1	0	0
202	0.005	0	0.005	0	0
203	0.005	0.005	0	7.18E-98	9.25E-99
M			M		M
401	0.005	0.995	-0.99	1.32E-226	1.70E-227
402	0.005	1	-0.995	0	0
M			M		M
20200	0.5	0.495	0.005	2.01E-11	2.59E-12
20201	0.5	0.5	0	1.16E-11	1.49E-12
20202	0.5	0.505	-0.005	6.57E-12	8.46E-13
M			M		M
20300	0.5	0.995	-0.495	4.34E-161	5.59E-162
20301	0.5	1	-0.5	0	0
M			M		M
40000	0.995	0	0.995	0	0
40001	0.995	0.005	0.99	6.82E-52	8.78E-53
M			M		M
40399	1	0.99	0.01	0	0
40400	1	0.995	0.005	0	0
40401	1	1	0	0	0

5) คำนวณความน่าจะเป็นที่จะได้ผลต่าง $p_1 - p_2$ คือ $P(P_1 - P_2 = \delta) = P(\delta)$ โดยจะรวมค่าน้ำหนักความน่าจะเป็นที่ได้จากขั้นที่ 4 ทุกค่าที่ค่าความแตกต่างเท่ากับ δ ที่เป็นไปได้ทั้งหมดเช่น

$$\delta = 0.6 = \begin{cases} p_1 = 0.6, & p_2 = 0 \\ p_1 = 0.605, & p_2 = 0.005 \\ \quad \mathrm{M} & \quad \mathrm{M} \\ p_1 = 1, & P_2 = 0.4 \end{cases}$$

ดังนั้น $P(\delta = 0.6) = \sum L^*(p_1, p_2 | \delta = 0.6) = 0.005594171$ ซึ่งจะได้ดังตารางที่ 2

ตารางที่ 2 แสดงการคำนวณน้ำหนักความน่าจะเป็นของค่าความแตกต่างที่เป็นได้ทั้งหมด $P(\delta)$

δ	$P(\delta)$
-1	0
-0.995	0
$\vdots$	$\vdots$
0.305	0.002924285
0.31	0.003309058
0.315	0.003731738
$\vdots$	$\vdots$
0.6	0.005594171
0.605	0.004887296
0.61	0.004241349
$\vdots$	$\vdots$
0.995	0
1	0

6) คำนวณช่วงความเชื่อมั่น $100(1-\alpha)\%$ ของ $p_1 - p_2$ โดยคำนวณความน่าจะเป็นสะสม (Cumulative Probability) แล้วใช้การประมาณค่าในช่วงเชิงเส้น (Linear Interpolation) ดังในสมการ $(**)$ เพื่อสร้างช่วงความเชื่อมั่นซึ่งจะได้ดังตารางที่ 3

ตารางที่ 3 แสดงการคำนวณช่วงความเชื่อมั่นสำหรับความแตกต่างระหว่างสองสัดส่วน

δ	$P(\delta)$	(π)
-1	0	0
-0.995	0	0
$\vdots$	$\vdots$	$\vdots$
0.305	0.002924285	0.021523
0.31	0.003309058	0.024832
$\vdots$	$\vdots$	$\vdots$
0.6	0.005594171	0.969513
0.605	0.004887296	0.9744
$\vdots$	$\vdots$	$\vdots$
0.995	0	1
1	0	1

หมายเหตุ π หมายถึง ความน่าจะเป็นสะสม (Cumulative Probability)

ดังนั้น ช่วงความเชื่อมั่น 95% ของ $p_1 - p_2$ คือ

$$LCL = \tilde{\delta}_{\alpha/2} = \frac{0.31(0.028564-0.025)+0.315(0.025-0.024832)}{(0.028564-0.024832)} = 0.3102$$

$$UCL = \tilde{\delta}_{1-\alpha/2} = \frac{0.605(0.97841-0.975)+0.2628(0.975-0.9744)}{(0.97841-0.9744)} = 0.6057$$

นำออก: ขีดจำกัดล่างและบนของช่วงความเชื่อมั่น $(1-\alpha)100\%$

2. ผลการศึกษาเชิงจำลอง

ในการเปรียบเทียบค่าประมาณความน่าจะเป็นคุ้มรวม ความยาวของช่วงและเปอร์เซ็นต์ที่ขีดจำกัดล่างและบนออกนอกช่วง $[-1,+1]$ โดยได้ทำการทดลองมอนติคาร์โลในทั้ง 4 สถานการณ์ สรุปได้ดังนี้

ตารางที่ 4 ค่าประมาณความน่าจะเป็นคุ้มรวม (CP) ความยาว (Length) ของช่วงความเชื่อมั่น 95% สำหรับสถานการณ์ที่ 1 : $p_1 = 0.5,\ p_2 = 0.5$

n_1	n_2	วิธี	CP	Length	n_1	n_2	วิธี	CP	Length	n_1	n_2	วิธี	CP	Length
10	10	WL	<u>0.9070</u>	0.8295	30	30	WL	0.9537	0.4974	50	50	WL	0.9445	0.3880
		HYB	0.9586	**0.7228**			HYB	0.9540	**0.4698**			HYB	0.9445	**0.3743**
		BAY	0.9586	0.7407			BAY	0.9540	0.4753			BAY	0.9445	0.3772
		MUE	<u>0.8899</u>	0.7851			MUE	<u>0.9386</u>	0.4948			MUE	0.9414	0.3940
		SIR	0.9585	0.7345			SIR	0.9542	0.4731			SIR	0.9501	0.3755
		QL	0.9573	0.7359			QL	0.9489	0.4739			QL	0.9431	0.3764
	30	WL	<u>0.9209</u>	0.6843		50	WL	0.9452	0.4462		100	WL	0.9467	0.3367
		HYB	0.9476	**0.5921**			HYB	0.9503	**0.4211**			HYB	0.9469	**0.3247**
		BAY	0.9459	0.6226			BAY	0.9503	0.4291			BAY	0.9469	0.3288
		MUE	0.9437	0.6649			MUE	0.9479	0.4419			MUE	0.9488	0.3365
		SIR	0.9553	0.6158			SIR	0.9538	0.4270			SIR	0.9538	0.3273
		QL	0.9522	0.6165			QL	0.9502	0.4279			QL	0.9511	0.3281
	50	WL	<u>0.9141</u>	0.6477		100	WL	<u>0.9364</u>	0.4020	100	100	WL	0.9423	0.2758
		HYB	0.9430	**0.5602**			HYB	0.9452	**0.3794**			HYB	0.9423	**0.2707**
		BAY	0.9420	0.5880			BAY	0.9429	0.3870			BAY	0.9423	0.2718
		MUE	0.9408	0.6238			MUE	0.9434	0.3975			MUE	0.9445	0.2788
		SIR	0.9541	0.5798			SIR	0.9474	0.3846			SIR	0.9472	**0.2707**
		QL	0.9610	0.5807			QL	0.9475	0.3855			QL	0.9508	0.2717
	100	WL	<u>0.9077</u>	0.6176										
		HYB	0.9514	**0.5343**										
		BAY	<u>0.9345</u>	0.5578										
		MUE	<u>0.9346</u>	0.5930										
		SIR	0.9581	0.5477										
		QL	0.9556	0.5487										

หมายเหตุ : ขีดเส้นใต้ หมายถึง CP < .94 ที่ระดับนัยสำคัญ .01
ตัวหนา หมายถึง ความยาวของช่วงต่ำสุด

ในสถานการณ์ที่ผลต่าง $p_1 - p_2 = 0$ (มีค่าน้อยมาก) และประชากรทั้งสองสมมาตร กล่าวคือ p_1 และ p_2 เท่ากับ 0.5 จากตารางที่ 4 พบว่า ช่วงความเชื่อมั่นแบบวัลด์ให้ความน่าจะเป็นคุ้มรวมต่ำกว่าสัมประสิทธิ์ที่กำหนดเพียงวิธีเดียว ในกรณีที่ขนาดตัวอย่างของอย่างน้อย 1 กลุ่มมีขนาดเล็ก $(n_i = 10)$ นอกจากนี้มีบางกรณีที่ตัวอย่างขนาด

เล็กและทั้งสองกลุ่มมีขนาดเท่ากัน หรือ แตกต่างกันมาก อย่างเช่น $(n_1, n_2) = (10,10)$ และ $(10,100)$ พบว่าวิธี MUE จะให้ค่าความน่าจะเป็นคุ้มรวมต่ำกว่าสัมประสิทธิ์ที่กำหนดเช่นกัน ดังที่ขีดเส้นใต้ไว้ นั่นคือ ค่าความน่าจะเป็นคุ้มรวม (CP) ยังมีไม่เท่ากับ 0.95 เมื่อเทียบกับ .94 ที่ระดับนัยสำคัญ .01 ตามที่ระบุในขั้นตอนการศึกษา หรือยังไม่ผ่านเกณฑ์ที่กำหนด โดยภาพรวมสำหรับสถานการณ์นี้หาก $\min(n_1, n_2) \geq 30$ ส่วนใหญ่ทั้ง 6 วิธีจะให้ช่วงที่มีความน่าจะเป็นคุ้มรวมใกล้เคียง .95 ซึ่งวิธีวัลด์จะมีแนวโน้มที่จะให้ช่วงที่มีความน่าจะเป็นคุ้มรวมต่ำที่สุดและมีความยาวของช่วงมากที่สุด ส่วนวิธีอัลกอริทึม QL ให้ช่วงที่มีความน่าจะเป็นคุ้มรวมและความยาวของช่วงใกล้เคียงวิธีไฮบริดที่ให้ความยาวของช่วงต่ำสุด นอกจากนี้ทั้ง 6 วิธีให้ช่วงที่อยู่ใน $[-1, +1]$ ในทุกกรณีที่ศึกษา

ตารางที่ 5 ค่าประมาณความน่าจะเป็นคุ้มรวม (CP) ความยาว (Length) ของช่วงความเชื่อมั่น 95%
สำหรับสถานการณ์ที่ 2 : $p_1 = 0.05,\ p_2 = 0.05$

n_1	n_2	วิธี	CP	Length	n_1	n_2	วิธี	CP	Length	n_1	n_2	วิธี	CP	Length
10	10	WL	0.6310	0.2850	30	30	WL	0.9245	0.2052	50	50	WL	0.9428	**0.1651**
		HYB	0.9986	0.5910			HYB	0.9930	0.2875			HYB	0.9799	0.2063
		BAY	0.9986	0.4968			BAY	0.9930	**0.2533**			BAY	0.9799	0.1875
		MUE	0.9929	**0.4089**			MUE	0.9320	0.2125			MUE	0.9121	0.1666
		SIR	0.9986	0.5176			SIR	0.9848	0.2595			SIR	0.9693	0.1903
		QL	0.9991	0.5184			QL	0.9824	0.2608			QL	0.9695	0.1911
	30	WL	0.8383	0.2516		50	WL	0.9223	0.1867		100	WL	0.9160	0.1435
		HYB	0.9940	0.4700			HYB	0.9889	0.2566			HYB	0.9757	0.1776
		BAY	0.9940	0.3935			BAY	0.9889	0.2230			BAY	0.9774	0.1599
		MUE	0.9932	**0.3188**			MUE	0.9601	**0.1928**			MUE	0.9558	**0.1452**
		SIR	0.9770	0.3991			SIR	0.9732	0.2267			SIR	0.9566	0.1603
		QL	0.9780	0.4002			QL	0.9795	0.2272			QL	0.9661	0.1613
	50	WL	0.8092	0.2337		100	WL	0.8335	0.1662	100	100	WL	0.9470	**0.1186**
		HYB	0.9837	0.4340			HYB	0.9860	0.2260			HYB	0.9641	0.1345
		BAY	0.9934	0.3738			BAY	0.9883	0.1993			BAY	0.9659	0.1268
		MUE	0.9660	**0.2976**			MUE	0.9726	**0.1711**			MUE	0.9255	0.1192
		SIR	0.9587	0.3719			SIR	0.9603	0.1991			SIR	0.9507	0.1267
		QL	0.9704	0.3728			QL	0.9740	0.2003			QL	0.9560	0.1289
	100	WL	0.5558	0.2162										
		HYB	0.9688	0.3997										
		BAY	0.9915	0.3614										
		MUE	0.9479	**0.2801**										
		SIR	0.9428	0.3525										
		QL	0.9523	0.3537										

หมายเหตุ : ขีดเส้นใต้ หมายถึง CP < .94 ที่ระดับนัยสำคัญ .01
ตัวหนา หมายถึง ความยาวของช่วงต่ำสุด

สำหรับสถานการณ์ผลต่าง $p_1 - p_2 = 0$ (มีค่าน้อยมาก) แต่ประชากรในทั้งสองเบ้เมื่อพิจารณาที่ค่าความเบ้ในแต่ละขนาดตัวอย่างคือ $n_i = 10, 30, 50$ และ 100 ซึ่งมีความเบ้เท่ากับ 1.3059, 0.7540, 0.5840 และ 0.4130 ตามลำดับ ($p_1 = p_2 = .05$) จากการจำลองสรุปดังในตารางที่ 5 ผลการศึกษาพบกว่า ช่วงที่ได้จากวิธีแบบวัลด์ ในกรณีที่ $\min(n_1, n_2) \leq 30$ ซึ่งประชากรมีลักษณะเบ้ค่อนข้างสูง และวิธี MUE ในกรณีที่ขนาดตัวอย่างทั้งสองกลุ่มมีขนาดใหญ่ $(n_i = 50,100)$ ให้ค่าความน่าจะเป็นคุ้มรวมต่ำกว่า .95 ส่วนวิธีที่เหลือส่วนใหญ่ให้ค่าความน่าจะเป็นคุ้มรวม

เท่ากับ .95 ถึงแม้ตัวอย่างของทั้ง 2 กลุ่มมีขนาดเล็ก โดยวิธีเบส์มีแนวโน้มให้ค่าความน่าจะเป็นคุ้มรวมสูงที่สุดในหลายกรณีที่ศึกษา นอกจากนี้ในทุกกรณีที่ศึกษาทั้ง 6 วิธีจะให้ช่วงที่อยู่ใน $[-1,+1]$

ตารางที่ 6 ค่าประมาณความน่าจะเป็นคุ้มรวม (CP) ความยาว (Length) ของช่วงความเชื่อมั่น 95%

สำหรับสถานการณ์ที่ 3: $p_1 = 0.75,\ p_2 = 0.25$

n_1	n_2	วิธี	CP	Length
10	10	WL	<u>0.8937</u>	0.7112
		HYB	0.9570	**0.6575**
		BAY	0.9556	0.6738
		MUE	0.9382	0.6681
		SIR	0.9554	0.6650
		QL	0.9572	0.6658
	30	WL	<u>0.8923</u>	0.5848
		HYB	<u>0.9341</u>	0.5369
		BAY	0.9505	0.5605
		MUE	0.9435	0.5667
		SIR	0.9474	**0.5534**
		QL	0.9513	0.5550
	50	WL	<u>0.8903</u>	0.5502
		HYB	0.9428	**0.5070**
		BAY	0.9506	0.5288
		MUE	0.9435	0.5667
		SIR	0.9546	0.5210
		QL	0.9513	0.5550
	100	WL	<u>0.9135</u>	0.5271
		HYB	0.9512	**0.4851**
		BAY	0.9508	0.5050
		MUE	<u>0.9343</u>	0.5053
		SIR	0.9582	0.4953
		QL	0.9588	0.4952

n_1	n_2	วิธี	CP	Length
30	30	WL	<u>0.9388</u>	0.4297
		HYB	0.9450	**0.4143**
		BAY	0.9454	0.4194
		MUE	0.9458	0.4233
		SIR	0.9468	0.4165
		QL	0.9516	0.4173
	50	WL	<u>0.9294</u>	0.3849
		HYB	0.9474	**0.3708**
		BAY	0.9448	0.3770
		MUE	0.9458	0.3814
		SIR	0.9472	0.3745
		QL	0.9456	0.3759
	100	WL	<u>0.9259</u>	0.3467
		HYB	0.9463	**0.3340**
		BAY	0.9443	0.3399
		MUE	0.9441	0.3432
		SIR	0.9496	0.3371
		QL	0.9494	0.3381

n_1	n_2	วิธี	CP	Length
50	50	WL	0.9470	0.3354
		HYB	0.9498	**0.3276**
		BAY	0.9514	0.3303
		MUE	0.9550	0.3327
		SIR	0.9471	0.3283
		QL	0.9451	0.3292
	100	WL	0.9418	0.2911
		HYB	0.9486	**0.2843**
		BAY	0.9476	0.2874
		MUE	0.9530	0.2895
		SIR	0.9483	0.2856
		QL	0.9492	0.2866
100	100	WL	0.9421	0.2386
		HYB	0.9505	0.2357
		BAY	0.9498	0.2367
		MUE	0.9428	0.2377
		SIR	0.9487	**0.2352**
		QL	0.9525	0.2364

หมายเหตุ : ขีดเส้นใต้ หมายถึง CP < .94 ที่ระดับนัยสำคัญ .01

ตัวหนา หมายถึง ความยาวของช่วงต่ำสุด

สำหรับสถานการณ์ที่ผลต่าง $p_1 - p_2 = 0.5$ และประชากรทั้งสองเบ้ ($p_1 = 0.75,\ p_2 = 0.25$) ผลการศึกษาแสดงตารางที่ 6 วิธีวัลด์ยังต้องใช้ตัวอย่างขนาดใหญ่ในทั้ง 2 กลุ่มหรือ $\min(n_1, n_2) \geq 50$ จึงจะให้ CP เท่ากับ .95 และวิธีไฮบริดมีแนวโน้มที่ให้ความยาวของช่วงต่ำสุด ส่วนอีก 4 วิธีที่เหลือให้ค่าความน่าจะเป็นคุ้มรวมเท่ากับ .95 ตามที่ต้องการ ถึงแม้ตัวอย่างในทั้ง 2 กลุ่มจะมีขนาดเล็กก็ตามค่าความน่าจะเป็นคุ้มรวมที่ได้มีค่าใกล้เคียงกันเช่นเดียวกับความยาวของช่วงความเชื่อมั่นและทั้ง 6 วิธีให้ช่วงที่อยู่ใน $[-1,+1]$ ในทุกกรณีที่ศึกษา

เนื่องจากช่วงความเชื่อมั่นสำหรับความแตกต่างระหว่างสองสัดส่วนที่ได้จากทั้ง 6 วิธี ในสถานการณ์ที่ 1-3 ให้ช่วงความเชื่อมั่นที่ได้ขีดจำกัดบนและล่างที่อยู่ในช่วง $[-1,+1]$ ทุกกรณี ดังนั้นในตารางที่ 4 ถึง ตารางที่ 6 จึงนำเสนอเพียงค่าความน่าจะเป็นคุ้มรวมและความยาวของช่วงเท่านั้น

ตารางที่ 7 ค่าประมาณความน่าจะเป็นคุ้มรวม (CP) ความยาว (Length) และเปอร์เซ็นต์ของขีดจำกัดบนที่มากกว่าหนึ่ง (% > 1) ของช่วงความเชื่อมั่น 95% สำหรับสถานการณ์ที่ 4: $p_1 = 0.95,\ p_2 = 0.05$

n_1	n_2	วิธี	CP	Length	>1	n_1	n_2	วิธี	CP	Length	>1	n_1	n_2	วิธี	CP	Length	>1
10	10	WL	0.6400	0.2842	61.51	30	30	WL	0.8000	0.2051	59.66	50	50	WL	0.8745	0.1655	24.91
		HYB	0.9275	0.4804	35.72			HYB	0.9216	0.2506	4.82			HYB	0.9338	0.1864	0.47
		BAY	0.9275	0.4966	73.62			BAY	0.9694	0.2532	19.67			BAY	0.9662	0.1879	3.49
		MUE	0.9259	0.3557	0			MUE	**0.9703**	0.2041	0			MUE	0.9324	0.1608	0
		SIR	0.8108	0.4804	0			SIR	0.9038	0.2466	0			SIR	0.9149	0.1833	0
		QL	0.7377	0.4833	0			QL	0.9203	0.2488	0			QL	0.9326	0.1850	0
	30	WL	0.6632	0.2505	76.86		50	WL	0.8531	0.1875	40.5		100	WL	0.9188	0.1435	0
		HYB	0.8899	0.3911	12.96			HYB	0.9167	0.2251	1.67			HYB	0.9354	0.1612	0
		BAY	0.9692	0.3933	55.28			BAY	0.9693	0.2235	8.52			BAY	0.9651	0.1600	0
		MUE	**0.9402**	0.2915	0			MUE	**0.9588**	0.1838	0			MUE	**0.9519**	**0.1418**	0
		SIR	0.8648	0.3824	0			SIR	0.9014	0.2179	0			SIR	0.9232	0.1559	0
		QL	0.8538	0.3856	0			QL	0.9027	0.2189	0			QL	0.9191	0.1583	0
	50	WL	0.6730	0.2339	75.72		100	WL	0.8609	0.1659	23.23	100	100	WL	0.9219	0.1188	0.96
		HYB	0.9183	0.3693	4.47			HYB	0.9320	0.2012	0.16			HYB	0.9449	0.1263	0
		BAY	0.9787	0.3738	54.34			BAY	0.9726	0.1992	4.14			BAY	0.9651	0.1270	0.04
		MUE	**0.9523**	0.2775	0			MUE	**0.9739**	0.1674	0			MUE	**0.9506**	**0.1191**	0
		SIR	0.8805	0.3629	0			SIR	0.9171	0.1940	0			SIR	0.9248	0.1236	0
		QL	0.8945	0.3642	0			QL	0.9193	0.1971	0			QL	0.9199	0.1261	0
	100	WL	0.6211	0.2106	49.49												
		HYB	0.9307	0.3507	0.55												
		BAY	0.9856	0.3592	54.81												
		MUE	**0.9683**	0.2709	0												
		SIR	0.9051	0.3468	0												
		QL	0.9113	0.3514	0												

หมายเหตุ : ตัวหนา หมายถึง CP < .94 ที่ระดับนัยสำคัญ .01 ความยาวของช่วงต่ำสุด และเปอร์เซ็นต์ของขีดจำกัดบนที่มากกว่าหนึ่ง (%) เท่ากับ ศูนย์

ในสถานการณ์สุดท้าย ผลจากการจำลองได้สรุปดังในตารางที่ 7 ในกรณีพารามิเตอร์ที่แท้จริง $p_1 - p_2 = .9$ เข้าใกล้ 1 จะสังเกตว่าสถานการณ์นี้ทำให้ขีดจำกัดบนที่ได้มากกว่าหนึ่ง หลายวิธี ได้แก่ วิธีวัลด์ วิธีไฮบริด และ วิธีแบบเบส์ สำหรับวิธีเบส์ เช่น กรณี $(n_1 = 10, n_2 = 30)$ แม้ว่าจะให้ค่าความน่าจะเป็นคุ้มรวมมากกว่า .95 อย่างไรก็ตามขีดจำกัดบนมากกว่าหนึ่งมากถึง 55% ส่วนวิธีวัลด์ยังคงเป็นวิธีที่มีประสิทธิภาพต่ำกว่าวิธีอื่น ๆ ในขณะที่ขีดจำกัดบนของช่วงจากวิธี QL และวิธี SIR มีค่าไม่เกิน 1 แต่ยังมีค่าความน่าจะเป็นคุ้มรวมไม่เท่ากับ .95 ถึงแม้ขนาดตัวอย่างของทั้ง 2 กลุ่มจะมีขนาดใหญ่ คือ $n_1 = 100, n_2 = 100$ โดยสรุป วิธี MUE มีแนวโน้มให้ค่าความน่าจะเป็นคุ้มรวมสูงสุด ความยาวของช่วงต่ำสุด และมีเปอร์เซ็นต์ของขีดจำกัดบนที่มากกว่าหนึ่ง $(\% > 1)$ มีค่าเป็นศูนย์ ซึ่งถือว่าวิธีนี้เหมาะสมกับสถานการณ์นี้

อภิปรายและสรุปผลการวิจัย

ช่วงความเชื่อมั่นสำหรับความแตกต่างระหว่างสองสัดส่วนโดยอัลกอริทึมควอนไทล์ของภาวะน่าจะเป็นที่ประยุกต์จากหลักการของช่วงความเชื่อมั่นโดยวิธีบูตสแตรปภายใต้ตัวประมาณมัธยฐานไม่เอนเอียง (MUE) และอัลกอริทึมการเลือกตัวอย่างแบบเลือกซ้ำที่สำคัญ (SIR) ในการหาช่วงความเชื่อมั่นสำหรับความแตกต่างระหว่างสอง

สัดส่วน ผลการศึกษาเชิงจำลอง พบว่า โดยภาพรวมช่วงความเชื่อมั่นโดยอัลกอริทึมควอนไทล์ของภาวะน่าจะเป็นมีค่าความน่าจะเป็นคุ้มรวมไม่แตกต่างไปจากวิธีไฮบริดซึ่งเป็นวิธีที่พบว่ามีประสิทธิภาพสูงในแง่ของความน่าจะเป็นคุ้มรวมและความไม่ซับซ้อนในการคำนวณจากการศึกษาเปรียบเทียบ 11 วิธี โดย Newcombe (1998) นอกจากนี้ทั้ง 6 วิธีให้ความยาวของช่วงใกล้เคียงกัน และมีเปอร์เซ็นต์การออกนอกช่วง $[-1,1]$ มีค่าเท่ากับศูนย์ในสถานการณ์ที่ 1, 2 และ 3 สำหรับสถานการณ์ที่ 4 พบว่าวิธี MUE ให้ผลดีและเหมาะสมที่สุด เมื่อพิจารณาที่ค่าความน่าจะเป็นคุ้มรวม ความยาวของช่วงต่ำสุด และมีเปอร์เซ็นต์ของขีดจำกัดบนที่มากกว่าหนึ่ง $(\% > 1)$ มีค่าเป็นศูนย์ นอกจากนี้ลักษณะของประชากรที่มีต่อความยาวช่วงพบว่า หากทั้ง p_1 และ p_2 ใกล้ 0.5 (สมมาตร) ความยาวของช่วงความเชื่อมั่นมีแนวโน้มที่จะสูงกว่าลักษณะของประชากรที่เบ้ เช่น $p_1 = .05$ และ $p_2 = .05$ โดยสรุป วิธีอัลกอริทึม QL สามารถสร้างช่วงความเชื่อมั่นได้ดีใกล้เคียงช่วงที่ดีอย่างวิธีไฮบริด วิธีเบส์ และวิธี MUE ได้ในหลายกรณีที่ศึกษา

กิตติกรรมประกาศ

ผู้วิจัยขอขอบพระคุณ ผศ.เบญจมาศ ตุลยนิติกุล และ ผศ.ดร.พัทธ์ชนก ศรีสุรเดชชัย เป็นอย่างสูงที่ให้คำปรึกษาและคำแนะนำ คอยช่วยเหลือและให้กำลังใจเป็นอย่างดี ขอบพระคุณครอบครัวและเพื่อน ๆ ที่สนับสนุนและเป็นกำลังใจให้มาโดยตลอด

เอกสารอ้างอิง

โชติกา วุฒิสาร รมิดา ศรีเหรา และ พัทธ์ชนก ศรีสุรเดชชัย. ช่วงความเชื่อมั่นสำหรับความแตกต่างของสองสัดส่วนโดยขั้นตอนวิธีการเลือกตัวอย่างแบบเลือกซ้ำที่สำคัญ. วารสารวิชาการพระจอมเกล้าพระนครเหนือ 2561; 28(4) (รอตีพิมพ์)

Newcombe R G. Interval Estimation for the Difference between Independent Proportions: Comparison of Eleven Methods. Statistics in Medicine 1998; 17: 873–890.

Rohde C A. Introductory Statistical Inference with the Likelihood Function. 1st ed. London: Springer: 2014.

Rubin D B. A noniterative sampling/ importance resampling alternative to the data augmentation algorithm for creating a few imputations when fractions of missing information are modest: the SIR algorithm. Discussion of Tanner and Wong. Journal of the American Statistical Association 1987; 85: 543–546.

Rubin D B. Using the SIR algorithm to simulate posterior distributions. In Bernardo J M, DeGroot M H, Lindley D V, and Smith A F. Bayesian Statistics, Clarendon, Oxford; 1988.

Yan L, Newcombe R G, Stuart L, Carter R E. Fully specified bootstrap confidence intervals for the difference of independent binomial proportion bases on the median unbiased estimator. Biostatistics 2009; 28: 2876-2890.