บทที่ 1

บทนำ

ที่มาและความสำคัญของปัญหา

ในการศึกษาวิจัยทางการแพทย์ส่วนใหญ่มักจะเกี่ยวข้องกับการเปรียบเทียบระหว่าง 2 วิธีการปฏิบัติ เช่น การเปรียบเทียบระหว่างวิธีการรักษาแบบใหม่ กับ วิธีการรักษาแบบมาตรฐาน โดยปกติแล้วความแตกต่างของวิธีการรักษานี้ไม่ควรจะวัดด้วยความแตกต่างของค่าเฉลี่ยของตัวแปรตาม* เพียงตัวเดียว การวัดค่าตัวแปรตามหลาย ๆ ตัวน่าจะใช้วัดประสิทธิภาพ (efficiency) ของวิธีการรักษาได้ดีกว่า เช่น การศึกษาประสิทธิภาพของขี้ผึ้งในการรักษาตาตุ่มที่เกิดการบิดไปด้านข้างอย่างเฉียบพลัน (acute lateral distortion of the ankle) ระหว่างกลุ่มที่ได้รับขี้ผึ้งกับกลุ่มควบคุมนั้น การประเมินความก้าวหน้าของการรักษาจะวัดจากความเจ็บที่เกิดจากการเคลื่อนไหว (pain upon movement) ความเจ็บเมื่ออยู่นิ่ง ๆ (pain of rest) ซึ่งวัดด้วยสเกล visual analoge จาก 0 ถึง 100 และวัดการเพิ่มขึ้นของโมเมนตัม เป็นต้น ดังนั้นวิธีการทางสถิติที่จะนำมาใช้กับสถานการณ์นี้ควรจะเป็นวิธีที่คำนึงถึงความหลากหลายของตัวแปร กล่าวคือ ตัวแปรทุกตัวที่เราพิจารณาควรจะมีอิทธิพลต่อกระบวนการตัดสินใจในการทดสอบสมมุติฐานเกี่ยวกับความแตกต่างของค่าเฉลี่ยระหว่าง 2 กลุ่มวิธีการปฏิบัติ (treatment group) ดังกล่าว และควรจะคำนึงถึงความไม่เป็นอิสระกันระหว่างตัวแปรเหล่านี้ด้วย นอกจากนี้เราอาจจะคาดเดาได้ล่วงหน้าก่อนการวิเคราะห์ว่า ความแตกต่างระหว่างวิธีการปฏิบัตินั้นเป็นไปในทางเดียวกันสำหรับทุกตัวแปร กล่าวคือ ผลต่างระหว่างวิธีการปฏิบัติของตัวแปรทุกตัวเป็นบวกหรือลบเหมือนกันทุกตัวแปร เช่น วิธีการรักษาแบบใหม่ให้ผลในทางที่ดีขึ้นสำหรับทุกตัวแปร ซึ่งเป็นสิ่งที่เราต้องการ ดังนั้นวิธีการทางสถิติที่เหมาะสมก็ควรจะมีความไวต่อสมมุติฐานทางเลือกอื่นที่แสดงถึงการเปลี่ยนแปลงที่ดีขึ้นของทุกตัวแปรดังกล่าว

หลังจากการปฏิเสธสมมุติฐานว่างที่ว่าไม่มีความแตกต่างระหว่างสองวิธีการปฏิบัติสำหรับทุกตัวแปร (global hypothesis**) แล้วเรามักจะสนใจต่อไปว่าตัวแปรใดบ้างที่สนับสนุนความแตกต่างอย่าง

---

* เป็นตัวแปรเชิงปริมาณ ซึ่งมีตัวแปรอิสระเป็นตัวแปรเชิงคุณภาพ 1 ตัวแปร คือ ประชากร มี 2 ระดับ คือ ประชากรกลุ่มที่ 1 และ 2 และในการวิจัยนี้คำว่า ตัวแปร จะหมายถึงตัวแปรตามเท่านั้น

** $H_0^k : \mu_{1k} - \mu_{2k} = 0$ ทุกค่าของ k , $k = 1,2,\ldots,K$

มีนัยสำคัญนี้ ซึ่งเป็นที่มาของปัญหาการเปรียบเทียบพหุคูณ (Multiple Comparison) สำหรับการวิจัยทางการแพทย์ที่กล่าวมาข้างต้นซึ่งมีจำนวนตัวแปรที่สนใจจำกัด (K ตัว) และสามารถกำหนดตัวแปรได้แน่นอนนั้น ถ้าต้องการผลสรุปของการทดสอบสมมุติฐาน H_0 ที่ถูกต้องสำหรับทุกกรณีแล้ว การเปรียบเทียบพหุคูณที่ใช้จะต้องควบคุมอัตราความผิดพลาดของการทดลอง (Experimentalwise หรือ Familywise Error Rate : FWE *) ได้อย่างเข้มงวด (strong sense) ซึ่งมีนิยามดังนี้ ให้ H = $\{ H_0^1 , H_0^2 , \ldots , H_0^K \}$ เป็นเซตของสมมุติฐานว่างของตัวแปร K ตัวที่สนใจ การเปรียบเทียบพหุคูณจะควบคุม FWE อย่างเข้มงวด ถ้าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 (Type I Error) ในการทดสอบอย่างน้อย 1 การทดสอบ มีค่าน้อยกว่าหรือเท่ากับ α โดยไม่คำนึงถึงว่าจริง ๆ แล้วมีสมมุติฐานใดบ้างใน H ที่เป็นจริง สำหรับการเปรียบเทียบพหุคูณที่รู้จักกันอย่างแพร่หลาย คือ วิธี Least Square Deviation (LSD) ของฟิชเชอร์ (Fisher) ซึ่งเป็นวิธีการทดสอบสมมุติฐานแบบ 2 ขั้นตอน (Two - step procedure) กล่าวคือ เราจะทำการทดสอบพหุคูณโดยใช้ตัวสถิติ t (Mutiple t - test) สำหรับแต่ละตัวแปรที่ระดับนัยสำคัญ α ก็ต่อเมื่อเราปฏิเสธสมมุติฐาน H_0 ด้วยตัวสถิติ F ที่ระดับนัยสำคัญ α จะเห็นได้ว่าวิธีนี้ควบคุม FWE ไม่เข้มงวด (weak sense) เนื่องจากจะควบคุมได้เฉพาะกรณีที่ทุกค่าใน H เป็นจริง (ภายใต้ H_0) เท่านั้น สำหรับวิธีการพื้นฐานที่รู้จักกันอย่างแพร่หลายอีกวิธีการหนึ่ง คือ วิธีบอนเฟอร์โรนี (Bonferroni procedure) ซึ่งเป็นวิธีทดสอบเชิงเดี่ยว (Single - step procedure) ที่ทำการทดสอบแต่ละสมมุติฐานใน H ด้วยตัวสถิติ t ณ ระดับนัยสำคัญ $\alpha' = \alpha / K$ โดยไม่มีการทดสอบ F ก่อน และเมื่ออาศัยอสมการของบอนเฟอร์โรนี (ซึ่งกล่าวไว้ว่าความน่าจะเป็นของผลผนวก (union) ของเหตุการณ์มีค่าน้อยกว่าผลบวกของความน่าจะเป็นของแต่ละเหตุการณ์) จะได้ว่า วิธีนี้เป็นวิธีที่ควบคุม FWE ได้อย่างเข้มงวด (Hochberg and Tamhane , 1987)

สำหรับวิธีการเปรียบเทียบพหุคูณต่าง ๆ ที่ใช้ในการเปรียบเทียบค่าเฉลี่ยระหว่างสองกลุ่มวิธีการปฏิบัติของตัวแปรต่าง ๆ เพื่อดูว่ามีตัวแปรใดบ้างที่มีความแตกต่างอย่างมีนัยสำคัญซึ่งจะศึกษาในการวิจัยนี้จะคำนึงถึงสิ่งที่สำคัญ 3 ประการดังได้กล่าวแล้วข้างต้น คือ ความหลากหลายของตัวแปร ความหลากหลายของการทดสอบ และความแตกต่างในทิศทางเดียวกัน** ของตัวแปรต่าง ๆ (นั่นคือ จะมุ่งเน้นที่การทดสอบสมมุติฐานข้างเดียว (One-sided Tests)) นอกจากนี้ตัวแปรต่าง ๆ ที่สนใจอาจจะมีสหสัมพันธ์กัน ดังนั้นวิธีการที่ใช้ควรจะคำนึงถึงสหสัมพันธ์ระหว่างตัวแปรต่าง ๆ ด้วย ซึ่งเราสามารถแบ่งการวิเคราะห์ออกเป็น 3 กลุ่มใหญ่ ๆ คือ

---

* ความน่าจะเป็นของการปฏิเสธสมมุติฐานว่างที่เป็นจริงอย่างน้อย 1 สมมุติฐานจากสมมุติฐานว่างทั้งหมด

** ผลต่างของค่าเฉลี่ยระหว่างวิธีการปฏิบัติของทุกตัวแปรเป็นบวกหรือลบเหมือนกันทั้งหมด

1. วิธีการวิเคราะห์ทีละตัวแปร (Single Endpoint Analysis) เป็นวิธีที่ทดสอบสมมุติฐาน H_0^k ($k = 1,2,...,K$) โดยปรับค่า P (P - values) สำหรับแต่ละการทดสอบเพื่อทำให้ควบคุม FWE ได้อย่างเข้มงวด วิธีการมาตรฐานที่อยู่ในกลุ่มนี้คือ วิธีบอนเฟอร์โรนี ที่ได้กล่าวแล้ว แต่ในปี ค.ศ. 1979 โฮล์ม (Holm) ได้เสนอวิธีการที่พัฒนาจากวิธีบอนเฟอร์โรนี คือ วิธีบอนเฟอร์โรนี-โฮล์ม (Bonferroni - Holm Procedure) ซึ่งเป็นวิธีที่ควบคุม FWE ได้อย่างเข้มงวดเช่นเดียวกับวิธีบอนเฟอร์โรนี แต่มีอำนาจการทดสอบมากกว่า ถึงอย่างไรก็ตามวิธีนี้ก็ยังขาดคุณสมบัติที่ต้องการ เนื่องจากไม่ได้คำนึงถึงสหสัมพันธ์ระหว่างตัวแปรต่าง ๆ ด้วยเหตุนี้ในปี ค.ศ. 1991 เจมส์ (James) จึงได้เสนอวิธีการปรับค่าสำหรับโครงสร้างที่ตัวแปรมีสหสัมพันธ์กัน

2. วิธีการวิเคราะห์ทุกตัวแปรพร้อมกัน (Global Endpoint Analysis) เป็นวิธีการที่ใช้ตัวสถิติทดสอบสำหรับตัวแปรพหุ ตัวสถิติที่ใช้กันอย่างแพร่หลายในกลุ่มนี้ คือ ตัวสถิติ T^2 ของ โฮเทลลิ่ง (Hotelling 's T^2) แต่ตัวสถิตินี้ไม่ได้คิดค้นมาสำหรับกรณีที่เราอาจจะคาดเดาได้ล่วงหน้าว่า ตัวแปรแต่ละตัวมีความแตกต่างระหว่างกลุ่มวิธีการปฏิบัติในทิศทางเดียวกัน โอบรีน (O' Brien , 1984) จึงได้เสนอวิธีการทดสอบซึ่งใช้ตัวสถิติที่เหมาะกับกรณีนี้มากกว่า T^2 และมีอำนาจการทดสอบมากกว่ามาก คือวิธี OLS (Ordinary Least Square) และวิธี GLS (Generalised Least Square) ซึ่งอยู่บนพื้นฐานของการหาผลรวมที่ไม่ถ่วงน้ำหนักและถ่วงน้ำหนักของตัวสถิติทดสอบของแต่ละตัวแปรตามลำดับ โดยเฉพาะวิธี GLS นั้นค่าถ่วงน้ำหนักของแต่ละตัวแปรจะขึ้นอยู่กับสหสัมพันธ์ของตัวแปรนั้นกับตัวแปรอื่น ๆ ถ้ามีสหสัมพันธ์น้อยก็จะได้ค่าถ่วงน้ำหนักมากกว่า

3. วิธีการวิเคราะห์แบบบูตสแตรป (Bootstrap Analysis) เป็นวิธีการที่เสนอโดย เวสต์ฟอล และยัง (Westfall and Young) ในปี ค.ศ. 1989 และวิธีที่ปรับปรุงใหม่ ในปี ค.ศ.1993 เป็นวิธีที่ใช้เทคนิคบูตสแตรปในการปรับค่า P ซึ่งจะควบคุม FWE ได้อย่างเข้มงวด ถ้าจำนวนครั้งของการสุ่มซ้ำมากพอ (อย่างน้อย 1,000 ครั้ง) ทั้งยังคำนึงถึงความไม่เป็นอิสระกันระหว่างตัวแปรและมีความไวต่อสมมุติฐานทางเลือกอื่นข้างเดียว

เราสามารถประยุกต์วิธีแบบปิด (Closure Method) ซึ่งเสนอโดย มาร์คัส เพอริทส์ และแกบริล (Marcus Peritz and Gabriel ,1976) กับวิธีการในกลุ่มที่ 2 แล้วสร้างเป็นวิธีการทดสอบแบบปิด (Closed Testing Procedure) ซึ่งเป็นวิธีการเปรียบเทียบพหุคูณแบบขั้นลดลง (Step - down Procedure) โดยมีวิธีการทดสอบคือ ให้ $H = \{ H_0^J = \bigcap_{k \in J} H_0^k , J \subseteq \{1,2,...,K\} \}$ เราจะปฏิเสธสมมุติฐาน H_0^k ก็ต่อเมื่อเราปฏิเสธทุกสมมุติฐานร่วม H_0^J ใน H ที่มีมิติสูงกว่าซึ่งมี H_0^k ร่วมอยู่ด้วย นั่นคือเราจะเริ่มจากการทดสอบสมมุติฐานที่มีมิติสูงสุดก่อน คือ $H_0^{\{1,2,...,K\}}$ ถ้ายอมรับก็จะหยุดการ

ทดสอบแล้วยอมรับสมมุติฐานที่มีมิติต่ำกว่าทั้งหมด ถ้าปฏิเสธก็จะต้องทำการทดสอบสมมุติฐานที่มีมิติต่ำกว่าต่อไปเรื่อย ๆ ซึ่งวิธีการนี้จะควบคุม FWE ที่ระดับนัยสำคัญ α ได้อย่างเข้มงวด ถ้าทุก ๆ การทดสอบร่วม และ การทดสอบ H_0^* เป็นการทดสอบที่ระดับนัยสำคัญ α และเมื่อเราใช้ตัวสถิติ OLS และ GLS ในการทดสอบสมมุติฐานต่าง ๆ ในแต่ละขั้น ก็จะได้วิธีการเปรียบเทียบพหุคูณแบบขั้นลดลงที่เรียกว่าวิธีการทดสอบแบบปิด OLS (OLS Closed Test Procedure) และวิธีการทดสอบแบบปิด GLS (GLS Closed Test Procedure) ตามลำดับ ส่วนวิธีการในกลุ่มที่ 1 และ 3 ได้แก่ วิธีบอนเฟอร์โรนี - โฮล์ม วิธีเจมส์ - โฮล์ม (James - Holm Procedure) และวิธีปรับปรุงใหม่ของเวสต์ฟอลและยังหรือวิธีเวสต์ฟอล - ยัง (Westfall - Young Procedure) นั้น เป็นวิธีการทดสอบแบบขั้นลดลงที่ได้จากการประยุกต์วิธีการแบบปิดกับวิธีบอนเฟอร์โรนี วิธีดั้งเดิมของเจมส์ และวิธีดั้งเดิมของเวสต์ฟอลและยัง ตามลำดับ ซึ่งใช้การปรับค่า P แทนการทดสอบสมมุติฐานร่วม H_0^J

วัตถุประสงค์ของการวิจัย

การวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของวิธีการเปรียบเทียบพหุคูณแบบขั้นลดลงซึ่งใช้ในการทดสอบความแตกต่างของค่าเฉลี่ยระหว่าง 2 กลุ่มประชากรบนพื้นฐานของตัวแปรตามพหุคูณ* ดังนี้

1. วิธีการวิเคราะห์ทีละตัวแปร
    - 1.1 วิธีบอนเฟอร์โรนี - โฮล์ม (BON)
    - 1.2 วิธีเจมส์ - โฮล์ม (JAM)
2. วิธีการวิเคราะห์ทุกตัวแปรพร้อมกัน
    - 2.1 วิธีการทดสอบแบบปิด OLS (OLS)
    - 2.2 วิธีการทดสอบแบบปิด GLS (GLS)
3. วิธีการวิเคราะห์แบบบูตสแตรป
    - 3.1 วิธีเวสต์ฟอล-ยัง (WFY)

---

* คือ การพิจารณาว่าความแตกต่างระหว่าง 2 วิธีการปฏิบัติมีนัยสำคัญหรือไม่และถ้ามีนัยสำคัญแล้วตัวแปรใดบ้างที่มีความแตกต่างอย่างมีนัยสำคัญ

สมมุติฐานของการวิจัย

ในการทดสอบความแตกต่างระหว่างค่าเฉลี่ยของ 2 ประชากรบนพื้นฐานของตัวแปรตามพหุคูณ ภายใต้สถานการณ์ส่วนใหญ่วิธีเวลส์ฟอล-ยังควรจะมีประสิทธิภาพดีที่สุด เนื่องจากเป็นวิธีที่สร้างขึ้นมาเพื่อการวิจัยทางการแพทย์โดยเฉพาะ และมีคุณสมบัติตามที่ต้องการครบถ้วน ทั้งการคำนึงถึงความไม่เป็นอิสระกันระหว่างตัวแปรต่าง ๆ และมีความไวต่อสมมุติฐานทางเลือกอื่นข้างเดียว แต่กรณีที่โครงสร้างของสหสัมพันธ์เป็นแบบเท่ากัน และผลต่างของค่าเฉลี่ยของทุกตัวแปรเท่ากันด้วย ซึ่งเป็นกรณีที่แสดงถึงการที่ตัวแปรต่าง ๆ มีผลทำให้เกิดความแตกต่างเท่ากันนั้น วิธีการทดสอบแบบปิด OLS และวิธีการทดสอบแบบปิด GLS ซึ่งให้น้ำหนักสำหรับตัวสถิติของแต่ละตัวแปรเท่ากัน จะมีประสิทธิภาพใกล้เคียงกับวิธีเวลส์ฟอล-ยัง

ข้อตกลงเบื้องต้น

การวิจัยนี้พิจารณาการเปรียบเทียบความแตกต่างของค่าเฉลี่ยของ 2 กลุ่มวิธีการปฏิบัติหรือ 2 กลุ่มประชากร ซึ่งแต่ละกลุ่มจะวัดค่าตัวแปร K ตัว จากหน่วยทดลอง n_1 และ n_2 หน่วยสำหรับวิธีการปฏิบัติหรือประชากรกลุ่มที่ 1 และ 2 ตามลำดับ

ให้ y_{ijk} = ค่าที่ได้จากประชากรกลุ่มที่ i หน่วยทดลองที่ j และตัวแปรที่ k

($i = 1,2$, $j = 1,2,\dots,n_i$, $k = 1,2,\dots,K$)

ข้อมูลที่ได้จากประชากรกลุ่มที่ i (ในรูปเมทริกซ์) คือ

$$\left(y_{ijk}\right) = \begin{bmatrix} y_{i11} & y_{i12} & \cdots & y_{i1K} \\ y_{i21} & y_{i22} & \cdots & y_{i2K} \\ \vdots & & & \vdots \\ y_{in_i1} & y_{in_i2} & \cdots & y_{in_iK} \end{bmatrix} = \begin{bmatrix} \underset{\sim}{y}_{i1}' \\ \underset{\sim}{y}_{i2}' \\ \vdots \\ \underset{\sim}{y}_{in_i}' \end{bmatrix}$$

สำหรับกลุ่มประชากรกลุ่มที่ i ($i = 1,2$) $\underset{\sim}{y}_{i1}', \underset{\sim}{y}_{i2}', \dots, \underset{\sim}{y}_{in_i}'$ เป็นตัวอย่างสุ่มขนาด n_i จากการแจกแจงปกติ K ตัวแปรที่มีเวกเตอร์ของค่าเฉลี่ยเท่ากับ $\underset{\sim}{\mu}_i$ และเมทริกซ์ความแปรปรวนร่วมเท่ากับ Σ

เมื่อ $\underset{\sim}{\mu}_i = \begin{bmatrix} \mu_{i1} \\ \mu_{i2} \\ \vdots \\ \mu_{iK} \end{bmatrix}$

$$\Sigma = (\sigma_{kk'})_{k,k'=1,2,\ldots,K}$$

และ $P = (\rho_{kk'})_{k,k'=1,2,\ldots,K}$

ขอบเขตการวิจัย

1. การวิจัยนี้จะเปรียบเทียบประสิทธิภาพของวิธีการเปรียบเทียบพหุคูณแบบขั้นลดลงวิธีต่าง ๆ โดยพิจารณาจากความสามารถในการควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 * และอำนาจการทดสอบเฉลี่ยของแต่ละขั้น**

2. ให้สมมุติฐานของตัวแปรที่ k (k = 1,2,...,K) คือ

$$H_0^k : \mu_{1k} - \mu_{2k} = 0$$
$$H_1^k : \mu_{1k} - \mu_{2k} > 0$$

สมมุติฐานที่มีมิติเท่ากับ d (d = 1,2,...,K) คือ

$$H_0^J = \bigcap_{k\in J} H_0^k \ , J \subseteq \{1,2,\ldots,K\}$$

โดยที่ จำนวนสมาชิกของเซต J เท่ากับ d

เช่น d = K จะได้สมมุติฐานที่มีมิติสูงสุด คือ

$$H_0 : \underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2 = \underset{\sim}{0}$$
$$H_1 : \underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2 > \underset{\sim}{0}$$

---

* วัดจากการควบคุมสัดส่วนของการปฏิเสธสมมุติฐานว่างเมื่อสมมุติฐานว่างเป็นจริงสำหรับสมมุติฐานว่างที่มีมิติสูงสุด (Global Hypothesis)

** วัดจากค่าเฉลี่ยของสัดส่วนการปฏิเสธสมมุติฐานเมื่อสมมุติฐานไม่เป็นจริงที่มิติของสมมุติฐานเท่ากัน

หรือ $H_0^k : \mu_{1k} - \mu_{2k} = 0$

$H_1^k : \mu_{1k} - \mu_{2k} > 0$

ทุก k , $k = 1,2,...,K$

3. กำหนดระดับนัยสำคัญเป็น 0.01 และ 0.05

4. กำหนดให้ขนาดตัวอย่างเท่ากันทั้ง 2 กลุ่มประชากร คือ 10 30 และ 50

5. กำหนดจำนวนตัวแปรตาม (K) เท่ากับ 3 5 และ 7

6. กำหนดให้เมทริกซ์ความแปรปรวนร่วมของทั้ง 2 กลุ่มประชากรเท่ากัน และเท่ากับเมทริกซ์สหสัมพันธ์ โดยแบ่งโครงสร้างของสหสัมพันธ์ระหว่างตัวแปรต่าง ๆ ออกเป็น 2 กรณีคือ

6.1 กรณีสหสัมพันธ์เท่ากัน $\rho_{kk'} = \rho$ สำหรับ $k \neq k'$

6.2 กรณีสหสัมพันธ์ไม่เท่ากัน $\rho_{kk'} = \rho^{|k-k'|}$ ($k,k'=1,2,...,K$)

โดยให้ทั้ง 2 กรณี $\rho = 0 \rightarrow 0.9$

7. กำหนดให้ $\Delta_k = \mu_{1k} - \mu_{2k}$ โดยที่ $\mu_{2k} = 0$

7.1 กรณีตรวจสอบความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 :

$\Delta_k = 0$ ทุกค่าของ k

7.2 กรณีพิจารณาอำนาจของการทดสอบ :

$\Delta_k = 0.5$ เมื่อโครงสร้างของสหสัมพันธ์เป็นแบบ 6.1

$\Delta_k = 0.1*(K - k + 1)$ เมื่อโครงสร้างของสหสัมพันธ์เป็นแบบ 6.2

8. การวิจัยนี้จะจำลองสถานการณ์ต่าง ๆ โดยใช้เทคนิคการจำลองแบบมอนติคาร์โล (Monte Carlo Simulation Technique)ทำการจำลองข้อมูลซ้ำ 500 รอบในแต่ละสถานการณ์

เกณฑ์การตัดสินใจ

การเปรียบเทียบประสิทธิภาพของวิธีการต่าง ๆ ในการวิจัยนี้จะพิจารณาจาก

1. ความสามารถในการควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ซึ่งจะวัดจากสัดส่วนของเหตุการณ์ที่ปฏิเสธสมมุติฐานว่างเมื่อสมมุติฐานว่างเป็นจริงในขั้นแรกของการทดสอบ * เปรียบเทียบกับเกณฑ์ที่กำหนด

2. อำนาจการทดสอบ (Power of the test) ซึ่งจะวัดจากสัดส่วนของเหตุการณ์ที่ปฏิเสธสมมุติฐานว่างเมื่อสมมุติฐานว่างไม่เป็นจริงในขั้นแรกของการทดสอบ และโดยเฉลี่ยในแต่ละขั้นของการทดสอบ**

ประโยชน์ที่คาดว่าจะได้รับ

เป็นแนวทางในการเลือกใช้วิธีการเปรียบเทียบพหุคูณแบบขั้นลดลงซึ่งใช้ในการทดสอบความแตกต่างของค่าเฉลี่ยระหว่างสองประชากรบนพื้นฐานของตัวแปรตามพหุคูณ ในแต่ละสถานการณ์ได้อย่างเหมาะสม

---

* การทดสอบสมมุติฐานว่างที่มีมิติสูงสุด (เท่ากับจำนวนตัวแปรตามที่สนใจศึกษา)

** สำหรับขั้นที่ 2 ถึงขั้นที่ K (เมื่อ K คือ จำนวนตัวแปรตามที่สนใจศึกษา)