หอสมุดกลาง สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 1

บทนำ

ความเป็นมาและความสำคัญของปัญหา

ในการออกแบบการวิจัยโดยเฉพาะอย่างยิ่งในการวิจัยเชิงทดลอง ซึ่งนอกจากผู้วิจัยจะต้องมีความสามารถในการออกแบบการทดลองที่ดี และมีความเข้าใจในเรื่องที่จะศึกษาตลอดจนลักษณะข้อมูลที่เกี่ยวข้อง เพื่อการสรุปผลของการตอบปัญหาการวิจัยได้ถูกต้องและสมเหตุสมผลแล้ว ยังต้องขึ้นอยู่กับเทคนิคในการวางแผนและการวิเคราะห์ทางสถิติเป็นอันมาก การเลือกใช้วิธีการทดสอบทางสถิติที่เหมาะสมเป็นความสามารถของผู้วิจัยที่จะทำให้ได้ผลสรุปอันมีความหมายและเหมาะสมกับสถานการณ์

การวิจัยเชิงทดลองโดยเฉพาะอย่างยิ่งในทางการศึกษาและจิตวิทยานั้น ส่วนมากมักจะเผชิญกับปัญหาในการใช้สถิติทดสอบ ปัญหาที่เกิดขึ้นก็คือ สถิติทดสอบที่ใช้ไม่สอดคล้องกับข้อตกลงเบื้องต้น (assumptions) ของสถิติตัวนั้น ซึ่งในกรณีเช่นนี้จะเกิดทางเลือกขึ้นมาว่า จะยังคงใช้สถิติตัวนั้นทดสอบโดยยอมฝ่าฝืนข้อตกลงเบื้องต้น หรือจะเลือกใช้วิธีการทดสอบแบบอื่นทดแทน เช่น ใช้การทดสอบทางนันพาราเมตริก (Nonparametic) หรือการทดสอบแบบดิสตริบิชั่นฟรี (Distribution Free Test) ซึ่งสามารถจะนำมาใช้ทดสอบได้โดยไม่ต้องฝ่าฝืนข้อตกลงเบื้องต้น

การใช้สถิติทดสอบแบบดั้งเดิม (classical test statistic) เช่น การทดสอบที (t-test) และการทดสอบเอฟ (F-test) เป็นต้นมาวิเคราะห์ข้อมูลโดยไม่สนใจในความผิดพลาดของข้อตกลงเบื้องต้น มาจากพื้นฐานความเชื่อที่ว่า การที่มีการกระทบหรือฝ่าฝืน (violate) ข้อตกลงเบื้องต้นจะมีผลกระทบเพียงเล็กน้อยต่อการแจกแจงการสุ่มของสถิติทดสอบ (sampling distribution of the test statistic) ซึ่งถ้าหากสภาพความเชื่อนี้เป็นจริง เราก็จะเรียกสภาวะของสถิติตัวนั้นว่า มีความแกร่ง (robust) ซึ่ง Box (1953) เป็นคนแรกที่ใช้คำนี้ บางทีอาจจะเป็นการพอเพียงที่จะกล่าวว่าสถิติตัวนั้นมีความแกร่งถ้าหากว่าผลกระทบหรือการเปลี่ยนข้อตกลงเบื้องต้นไม่ได้ทำให้โอกาส

ที่จะเกิดความคลาดเคลื่อนประเภทที่ 1 (type I error) หรือความคลาดเคลื่อนประเภทที่ 2 (type II error) มากขึ้น (Sawat Pratoomraj 1970 : 1)

ในการวิจัยเชิงทดลองเกี่ยวกับการเปรียบเทียบกลุ่มตัวอย่างสุ่มสองกลุ่มซึ่งเป็นอิสระต่อกัน (indepedent groups) เพื่อศึกษาเปรียบเทียบความแตกต่างอย่างมีนัยสำคัญทางสถิติของประชากรซึ่งเป็นที่มาของกลุ่มตัวอย่างทั้งสอง เรามักจะเลือกใช้การทดสอบที (t-test) เพื่อทำการทดสอบข้อมูล แต่จากข้อตกลงเบื้องต้นของการทดสอบทีที่ว่า ลักษณะการแจกแจงของประชากรเป็นแบบปกติ (normal distribution) และความแปรปรวนของประชากร (variance) ของประชากรทั้งสองกลุ่มต้องเท่ากัน ในสภาพการณ์โดยทั่ว ๆ ไป ข้อเท็จจริงอาจจะมิได้เป็นไปตามข้อตกลงเบื้องต้นข้อใดข้อหนึ่งซึ่งมักจะเป็นเรื่องของการแจกแจงของประชากรหรือทั้งสองข้อก็ได้ ในกรณีนี้จะเกิดทางเลือกขึ้นมาสองกรณี คือ

กรณีที่ 1 ยังคงใช้การทดสอบทีทดสอบข้อมูล

กรณีที่ 2 เลือกใช้สถิติทดสอบนันพาราเมตริก หรือสถิติทดสอบดิสตริบิ้วชั่นฟรี ซึ่งมีข้อตกลงเบื้องต้นที่อ่อนกว่ามาก ทั้งลักษณะการแจกแจงและค่าความแปรปรวนของประชากร

<u>กรณีที่ 1</u> ยังใช้การทดสอบทีทดสอบข้อมูล

จากการศึกษาและวิจัยของ Hatch และ Posten (1966) ที่ได้ทำการวิจัยเพื่อตรวจสอบความแกร่งของการทดสอบทีภายใต้สภาพการณ์ของลักษณะการแจกแจงของประชากรไม่เป็นแบบปกติ (non-normality) และความไม่เท่ากันของความแปรปรวนของประชากร (heterogeneity of variance) สรุปผลได้ว่า การทดสอบทีมีความแกร่งแม้ว่าการกระจายและความแปรปรวนของประชากรไม่เป็นไปตามข้อตกลงเบื้องต้นของการทดสอบที และภายใต้สภาพการณ์ของลักษณะการแจกแจงของประชากรไม่เป็นแบบปกติการทดสอบทีจะมีความแกร่งมากกว่าภายใต้สภาพการณ์ความไม่เท่ากันของความแปรปรวนของประชากร (Sawat Pratoomraj 1970: 16)

Ramsey (1980: 337) กล่าวไว้ว่า ภายใต้สภาพการณ์ความไม่เท่ากันของความแปรปรวนของประชากรการทดสอบที่ไม่ได้มีความแกร่ง เสมอไปถึงแม้จะมีขนาดของกลุ่มตัวอย่าง เท่ากันก็ตาม และ Ramsey (1980 : 346) ยังได้พิสูจน์ว่า เมื่อความแปรปรวนของประชากรไม่เท่ากันถึงแม้ว่าจะมีลักษณะการแจกแจงของประชากร เป็นแบบปกติ และขนาดของกลุ่มตัวอย่าง เท่ากันก็ตามไม่สามารถที่จะพิจารณาถึงความสมบูรณ์ของความแกร่งของการทดสอบที่ได้

กรณีที่ 2 เลือกใช้สถิติทดสอบนันพารา เมตริก หรือสถิติทดสอบดิสตริบิชั่นฟรีซึ่งมีข้อตกลงเบื้องต้นเพียงแต่ลักษณะสุ่มมาตร เท่านั้น

สถิติทดสอบนันพารา เมตริกหรือดิสตริบิชั่นฟรีที่จะใช้แทนได้ในกรณีที่กลุ่มตัวอย่างทั้งสอง เป็นอิสระต่อกันมีดังนี้ คือ วิลค็อกซอน เทส [Wilcoxon Test or Wilcoxon's Rank-Sum Test (Wilcoxon, 1949)], เทอรี่-โฮฟฟ์ดิง นอร์มอล-สกอร์ เทส [Terry-Hoeffding Normal-Scores Test (Terry and Hoeffding, 1952)] และแวน เดอ แวร์เดน นอร์มอล-สกอร์ เทส [Van der Waerden Normal-Scores Test (Van der Waerden, 1953)]

Marascuilo and McSweeney (1977: 280) กล่าวว่า นักพฤติกรรมศาสตร์ได้ให้ความสนใจกับนอร์มอลสกอร์ (normal scores) มาก เพราะว่าค่า A.R.E. (Asymtotic Relative Efficiency) ของนอร์มอลสกอร์เมื่อเทียบกับการทดสอบที่มีค่า เท่ากับ 1 เมื่อข้อมูลเป็นไปตามข้อตกลงเบื้องต้นของการทดสอบที่ใน เรื่องของลักษณะการแจกแจงของประชากร เป็นแบบปกติ และความแปรปรวนของประชากรทั้งสองกลุ่ม เท่ากัน เชาคาดว่า A.R.E. จะมากกว่า 1 ถ้าข้อมูลไม่เป็นไปตามข้อตกลงเบื้องต้นของการทดสอบที นอกจากนั้นเมื่อข้อมูลเป็นไปตามข้อตกลงเบื้องต้นของการทดสอบที Relative Efficiency ของการทดสอบของวิลค็อกซอน (Wilcoxon Test) เมื่อเทียบกับการทดสอบที่มีค่าประมาณ $\frac{3}{\pi}$ หรือ 0.955 และภายใต้สภาพการณ์เดียวกันนี้สถิติพวกนอร์มอลสกอร์จะมีประสิทธิภาพมากกว่าการทดสอบของวิลค็อกซอน

การทดสอบของวิลค็อกซอนมีอำนาจของการทดสอบสูงที่สุด เมื่อการแจกแจงของประชากร เป็นโลจิสติค (logistic) ในกรณีที่กลุ่มตัวอย่างมีขนาดเล็กการทดสอบของวิลค็อกซอน

จะมีอำนาจของการทดสอบสูงกว่าการทดสอบที ถ้าข้อมูลที่ได้เป็นข้อมูลประเภทต่อเนื่อง มีรูปร่างเป็นแบบเดียวกันและมีความแปรปรวนเท่ากันไม่ว่าจะมีตำแหน่งของมัชฌิมเลขคณิตเท่ากันหรือไม่ก็ตามค่า A.R.E. ของการทดสอบของวิลค็อกซอนเมื่อเทียบกับการทดสอบทีจะมีค่าตั้งแต่ 0.864 ถึงอินฟินิตี้ (∞) แต่ถ้าลักษณะการแจกแจงของประชากรทั้งสองกลุ่มเป็นแบบปกติซึ่งเป็นไปตามข้อตกลงเบื้องต้นของการทดสอบที A.R.E. ของการทดสอบของวิลค็อกซอนเมื่อเทียบกับการทดสอบทีมีค่าประมาณ $\frac{3}{\pi}$ หรือ 0.955 ซึ่งเท่ากับ A.R.E. ของการทดสอบของวิลค็อกซอนเมื่อเทียบกับการทดสอบของเทอรี่-โฮฟฟ์ดิง (Terry-Hoeffding Normal-Scores Test) และภายใต้สภาพการณ์เดียวกันนี้สถิติทดสอบประเภทนอร์มอลสกอร์จะมีอำนาจของการทดสอบสูงกว่าการทดสอบของวิลค็อกซอน ในกรณีที่การแจกแจงของประชากรทั้งสองกลุ่มเป็นยูนิฟอร์ม A.R.E. ของการทดสอบของวิลค็อกซอนเมื่อเทียบกับการทดสอบทีมีค่าเป็น 1 แต่ถ้าการแจกแจงของประชากรมีรูปร่างต่างกันและความแปรปรวนไม่เท่ากันค่า A.R.E. ของการทดสอบของวิลค็อกซอนเมื่อเทียบกับการทดสอบทีอาจจะต่ำถึงศูนย์ได้ สิ่งที่สำคัญคือ ถ้าข้อมูลต่างกันมาก ๆ การทดสอบของวิลค็อกซอนจะใช้ได้ดีกว่าสถิติทดสอบดิสตริบิชั่นฟรีและสถิติทดสอบพารามิเตอร์ต่าง ๆ รวมทั้งมีค่า A.R.E. สูงกว่าด้วย ถ้าเปรียบเทียบการทดสอบของวิลค็อกซอนกับสถิติทดสอบประเภทนอร์มอลสกอร์ทั้งหลายแล้วจะเห็นว่าการทดสอบของวิลค็อกซอนง่ายกว่าในการคิดคำนวณ ง่ายกว่าในการตีความหมาย และตารางที่ใช้ก็เป็นที่นิยมกันอย่างแพร่หลาย (Bradley 1978: 108-109)

ในการใช้การทดสอบที่ทดสอบข้อมูลจะต้องอาศัยข้อตกลงเบื้องต้นในเรื่องของลักษณะการแจกแจงของประชากรต้องเป็นแบบปกติและความแปรปรวนของประชากรต้องเท่ากัน การทดสอบทีจะมีอำนาจของการทดสอบสูงมากถ้าการแจกแจงของประชากรทั้งสองกลุ่มเป็นแบบปกติ ในทางตรงกันข้ามถ้าลักษณะการแจกแจงของประชากรไม่เป็นแบบปกติ อาจจะยังคงใช้การทดสอบที่ทดสอบข้อมูลซึ่งจะทำให้อำนาจของการทดสอบน้อยลงไป หรือบางครั้งอาจจะใช้การทดสอบที่ทดสอบข้อมูลไม่ได้ การทดสอบของเทอรี่-โฮฟฟ์ดิงเป็นสถิติทดสอบที่จะช่วยปรับข้อมูลที่ได้มาให้มีลักษณะการแจกแจงใกล้เคียงแบบปกติที่มีมัชฌิมเลขคณิตเป็น 0 และความแปรปรวนเท่ากับ 1 ให้มากที่สุด ซึ่งต้องอาศัยตารางการแปลงคะแนนหลายครั้งและต้องอาศัยตารางถึงสองตาราง แม้ว่าวิธีการของเทอรี่-โฮฟฟ์ดิงยุ่งยากมากกว่าวิธีการของวิลค็อกซอนและการทดสอบทีแต่การคำนวณก็ไม่ยากนัก ถ้าลักษณะการแจกแจงของ

ประชากร เป็นแบบปกติการทดสอบของ เทอรี่-โฮฟฟ์ดิงจะมีอำนาจของการทดสอบใกล้เคียงกับอำนาจของการทดสอบทีแต่น้อยกว่า แต่ถ้าลักษณะการแจกแจงของประชากรไม่เป็นแบบปกติ การทดสอบของ เทอรี่-โฮฟฟ์ดิงจะมีอำนาจของการทดสอบดีกว่า A.R.E. ของการทดสอบของ เทอรี่-โฮฟฟ์ดิง เมื่อเทียบกับการทดสอบของวิลค็อกซอนมีค่าอยู่ระหว่าง 0.524 ถึง อินฟินิตี้ ส่วนจะมีค่า เท่าไรนั้นขึ้นอยู่กับลักษณะการแจกแจงของประชากร ถ้าลักษณะการแจกแจงของประชากร เป็นแบบปกติ A.R.E. จะมีค่าประมาณ $\frac{\pi}{3}$ หรือ 1.047 อย่างไรก็ตาม การทดสอบของ เทอรี่-โฮฟฟ์ดิงก็ยังคง เป็นการทดสอบที่มีอำนาจของการทดสอบสูงที่สุดในจำพวกสถิติทดสอบที่ใช้อันดับ (Bradley 1978: 149-151)

การทดสอบของแวน เดอ แวร์เดน (Van der Waerden Normal-Scores Test) เป็นการทดสอบที่มีวิธีการคล้ายคลึงกับการทดสอบของ เทอรี่-โฮฟฟ์ดิง และมีอำนาจของการทดสอบ เท่า ๆ กัน (Bradley 1978: 159)

Bradley (1978: 108) ยังได้กล่าวไว้ว่า ในการทดสอบความ เป็นอันหนึ่งอันเดียวกันของประชากรซึ่ง เป็นที่มาของกลุ่มตัวอย่างสุ่มสองกลุ่มซึ่ง เป็นอิสระต่อกัน มีวิธีการทดสอบในรูปแบบของดิสตริบิชั่นฟรีหลายวิธีด้วยกัน แต่วิธีการทดสอบใดจะมีอำนาจของการทดสอบสูงกว่ากันย่อมขึ้นอยู่กับลักษณะการแจกแจงของประชากร

ในทางการศึกษารูปแบบของนอร์มอลสกอร์ เป็นที่สนใจมากของนักพฤติกรรมศาสตร์ นักวิจัยและนักจิตวิทยา เพราะว่าผู้วิจัย เหล่านี้ต้องการใช้คะแนนที่เป็นมาตรฐานในการวัด (Marascuilo & McSweeney 1977: 284) และการที่เราจะพิจารณาว่าการทดสอบใดดีมากน้อย เพียงใดก็ให้ดูที่ค่า efficiency หรือค่าอำนาจของการทดสอบ จะ เห็นได้ว่างานวิจัยโดยทั่ว ๆ ไปที่เกี่ยวกับการวิเคราะห์ข้อมูลที่ได้มาจากกลุ่มตัวอย่างสุ่มสองกลุ่มซึ่ง เป็นอิสระต่อกันแล้ว มักจะนิยมใช้การทดสอบที่ทดสอบข้อมูล โดยสมมุติว่าลักษณะการแจกแจงของประชากร เป็นแบบปกติและความแปรปรวนของประชากร เท่ากัน ทั้ง ๆ ที่การแจกแจงและความแปรปรวนของประชากรอาจจะไม่ได้เป็นไปตามที่สมมุติไว้ก็ได้ แต่ก็พยายามที่จะใช้การทดสอบทีโดยไม่สนใจที่จะใช้สถิติทดสอบอื่นซึ่ง เทียบ เท่ากับการทดสอบที ตามความ เป็นจริงแล้วอาจจะใช้การทดสอบของวิลค็อกซอน, การทดสอบของ เทอรี่-โฮฟฟ์ดิง และการทดสอบของแวน เดอ แวร์เดนทำการทดสอบแทนได้

เมื่อเกิดการฝ่าฝืนข้อตกลงเบื้องต้นในเรื่องของความเท่ากันของความแปรปรวนของประชากร ผู้วิจัยอาจจะทำการทดสอบได้ด้วยสถิติทดสอบ F_{max} แต่การฝ่าฝืนข้อตกลงเบื้องต้นของลักษณะการแจกแจงของประชากรเป็นแบบปกติ ผู้วิจัยไม่อาจจะทราบได้แน่ชัดถึงลักษณะการแจกแจงของประชากรที่ศึกษาว่าเป็นการแจกแจงในรูปแบบใด แม้ว่าผู้วิจัยพยายามจะแก้ปัญหานี้โดยอาศัยทฤษฎี Central Limit Theorem โดยได้กำหนดให้กลุ่มตัวอย่างมีขนาดใหญ่เพื่อให้การแจกแจงของค่าสถิติมีลักษณะเป็นแบบปกติ โดยทฤษฎีนี้กล่าวว่าถ้าสุ่มตัวอย่างจากประชากรที่มีค่าเฉลี่ย μ และส่วนเบี่ยงเบนมาตรฐาน σ มีขนาดของกลุ่มตัวอย่าง (n) เพิ่มขึ้นแล้ว การแจกแจงของค่าเฉลี่ย $\bar{X}$ จะยิ่งเข้าสู่ลักษณะการแจกแจงเป็นแบบปกติที่มีค่าเฉลี่ย μ และส่วนเบี่ยงเบนมาตรฐาน $\sigma/\sqrt{n}$ แต่ในสภาพของการวิจัยเชิงทดลอง ผู้วิจัยไม่อาจที่จะหลีกเลี่ยงต่อการทดลองที่ต้องใช้กลุ่มตัวอย่างขนาดกลาง เช่น 10, 15 และ 20 หรือกลุ่มตัวอย่างขนาดเล็ก เช่น 5 ได้ ดังนั้น ผลการวิจัยจะมีคุณภาพมากน้อยเพียงใดจึงขึ้นอยู่กับการตัดสินใจของผู้ทดลองในการที่จะเลือกเทคนิค หรือวิธีการในการทดสอบข้อมูล ปัญหาจึงอยู่ที่ว่า การทดสอบใดจะให้ผลสรุปที่ถูกต้องมากที่สุดในทุกสภาพการณ์ของข้อมูล นั่นก็คือ การทดสอบใดมีอำนาจของการทดสอบ (Power) สูงสุด ซึ่งยังไม่มีหลักฐานและผลสรุปเพียงพอเกี่ยวกับอำนาจของการทดสอบทั้งสี่นี้ภายใต้สภาพการณ์ของลักษณะการแจกแจงของประชากรแตกต่างกัน ถึงแม้ว่าการเปรียบเทียบอำนาจของการทดสอบทีและการทดสอบของวิลค็อกซอนภายใต้ระดับอัตราความคลาดเคลื่อนที่ระบุ (α) เดียวกัน เมื่อข้อมูลไม่เป็นไปตามข้อตกลงเบื้องต้นของการทดสอบทีในเรื่องของลักษณะการแจกแจงของประชากรเป็นสิ่งที่ไม่ควรกระทำอย่างยิ่ง เนื่องมาจากการที่ข้อมูลไม่เป็นไปตามข้อตกลงเบื้องต้นของการทดสอบทีจะมีผลกระทบต่อความคลาดเคลื่อนประเภทที่ 1 ของการทดสอบที แต่ Bradley (1978: 109) ได้กล่าวไว้ว่า การที่จะตัดสินอำนาจของการทดสอบของสถิติทดสอบต่าง ๆ ได้อย่างแท้จริงนั้น จะต้องละเลยในเรื่องเหล่านี้คือ การรู้ระดับที่แท้จริงของ α ของแต่ละวิธีการทดสอบ และความสัมพันธ์ระหว่างอำนาจของการทดสอบกับระดับที่แท้จริงของ α และประการสุดท้าย การศึกษาอำนาจของการทดสอบในรูปของซิมูเลชั่น (Simulation) ซึ่งทำให้ผลสรุปเด่นชัดในสภาวะของสถานการณ์ที่คล้ายกับการทดลอง คือ สามารถที่จะระบุหรือจำกัดขนาดของกลุ่มตัวอย่าง ความแปรปรวน และ

ลักษณะของการแจกแจงของประชากรได้ด้วย เครื่องจักรกลคอมพิวเตอร์โดย เทคนิคที่เรียกว่า มอนติคาร์โล (Monte Carlo) ซึ่งเป็นวิธีที่สามารถดำเนินการได้ในปัจจุบัน

วัตถุประสงค์ของการวิจัย

การวิจัยครั้งนี้ต้องการที่จะตรวจสอบผลของการฝ่าฝืนข้อตกลงเบื้องต้นด้านการแจกแจงของประชากร เมื่อความแปรปรวนของประชากรเท่ากัน ในการหาผลสรุปเกี่ยวกับอำนาจของการทดสอบ 4 วิธี คือ

1. การทดสอบที (t-test)

2. การทดสอบของวิลค็อกซอน (Wilcoxon Test)

3. การทดสอบของเทอรี่-โฮฟฟ์ดิง (Terry-Hoeffding Normal-Scores Test)

4. การทดสอบของแวน เดอ แวร์เดน (Van der Waerden Normal - Scores Test)

ภายใต้เงื่อนไขต่อไปนี้ คือ

1. ลักษณะการแจกแจงของประชากรเป็นแบบปกติ (normal distribution) เป็นแบบยูนิฟอร์ม (uniform distribution) และเป็นแบบโลจิสติค (logistic distribution)

2. จำนวนกลุ่มตัวอย่าง 2 กลุ่ม และมีขนาดเท่ากัน

3. ความแปรปรวนของประชากรเท่ากัน

สมมุติฐานของการวิจัย

ผลของการฝ่าฝืนข้อตกลงเบื้องต้นด้านการแจกแจงของประชากรจะทำให้ค่าอำนาจของการทดสอบทั้ง 4 วิธีแตกต่างกัน

ข้อตกลงเบื้องต้น

1. การวิจัยครั้งนี้ถือว่า ค่าอำนาจของการทดสอบเป็นดัชนีสำคัญที่นักวิจัยจะใช้เป็นเกณฑ์วินิจฉัยเลือกสถิติทดสอบ

2. การวิจัยครั้งนี้เลือกใช้โปรแกรมสัปรูทีนที่มีหลักฐาน และการศึกษาในรูปการแปลงข้อมูลไปยังประชากรที่ต้องการวิจัย

ขอบเขตของการวิจัย

1. การวิจัยครั้งนี้ศึกษาตัวแปรดังต่อไปนี้

1.1 ตัวแปรตาม (dependent variable) คือ ค่าอำนาจของการทดสอบทั้ง 4 วิธีดังกล่าวแล้วข้างต้น

1.2 ตัวแปรอิสระ (independent variable) คือ

1.2.1 ลักษณะการแจกแจงของประชากร

1.2.2 แบบของการทดสอบ

1.2.3 ขนาดของกลุ่มตัวอย่าง

2. ศึกษาความแตกต่างของค่าอำนาจของการทดสอบของสถิติทดสอบต่าง ๆ เฉพาะลักษณะการแจกแจงของประชากรที่กำหนดให้ซึ่งมีลักษณะการแจกแจงเป็นแบบปกติ เป็นแบบยูนิฟอร์ม และแบบโลจิสติค

3. ศึกษาเฉพาะจำนวนกลุ่มตัวอย่าง 2 กลุ่ม ซึ่งมีขนาดเท่ากัน โดยกำหนดขนาดของกลุ่มตัวอย่างในการทดลองเท่ากับ 5, 10 และ 15

4. กำหนดพารามิเตอร์ $\mu = 500$ และ $\sigma^2 = 100$ สำหรับการทดลองเมื่อการแจกแจงเป็นแบบปกติ ยูนิฟอร์ม และโลจิสติค

5. จะคำนวณหาค่าอำนาจของการทดสอบเป็นจุด ๆ เมื่อเดลต้า ($\delta_i = \mu_1 - \mu_2$) มีค่าต่าง ๆ ดังนี้ $0, 0.25\sigma, 0.50\sigma, 0.75\sigma, 1.00\sigma, 1.25\sigma, 1.50\sigma$ และ 1.75σ

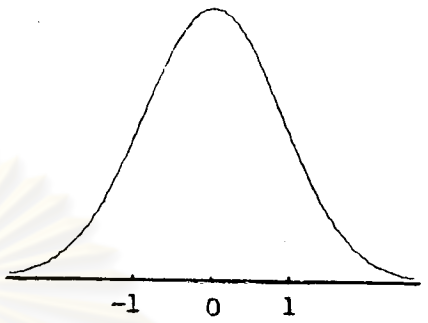
6. การวิจัยนี้จำลองการทดลองขึ้นในคอมพิวเตอร์ โดยอาศัยเทคนิคมอนติคาร์โลซิมูเลชั่น (Monte Carlo Simulation Technique) และโปรแกรมสัปรูทีน โดยจำกัดการศึกษาเฉพาะลักษณะการแจกแจงของประชากร 3 รูปแบบคือ การแจกแจงเป็นแบบปกติ การแจกแจงแบบยูนิฟอร์ม และการแจกแจงแบบโลจิสติค ซึ่งเป็นรูปแบบการแจกแจงที่พบมากในทางการศึกษาและจิตวิทยา

ฟังก์ชันของการแจกแจง (density functions) ซึ่งใช้ในการวิจัยมีดังนี้

7.1 ลักษณะการแจกแจงของประชากรเป็นแบบปกติ

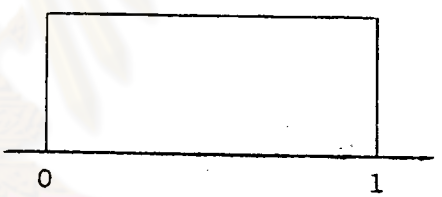
$$f(x) = \left(\frac{1}{2\pi}\right)^{\frac{1}{2}} e^{\frac{-x^2}{2}}$$

$-\infty < x < \infty$



7.2 ลักษณะการแจกแจงของประชากรแบบยูนิฟอร์ม

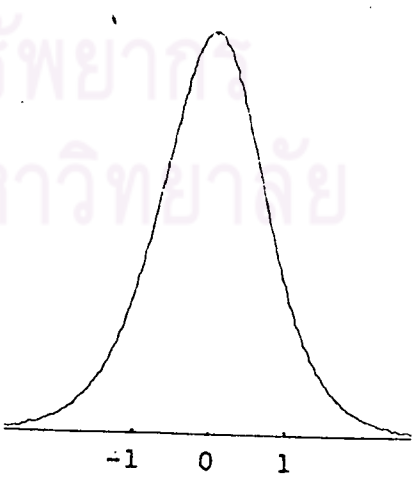
$$f(x) = \begin{cases} 1, & 0 \leq x \leq 1 \\ 0, & \text{otherwise} \end{cases}$$



7.3 ลักษณะการแจกแจงของประชากรแบบโลจิสติค

$$f(x) = \frac{e^{-x}}{(1+e^{-x})^2}$$

$-\infty < x < \infty$



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

7. ศึกษาเฉพาะกลุ่มตัวอย่างขนาดเล็กและขนาดกลางโดยกำหนดระดับอัตราความคลาดเคลื่อนประเภทที่ 1 (α) สองระดับคือ .05 และ .01

8. ศึกษาเฉพาะการทดสอบที่ การทดสอบของวิลค็อกซอน การทดสอบของเทอรี่-โฮฟฟ์ดิง และการทดสอบของแวน เดอ แวร์เดน

9. การซิมูเลทการทดลองนี้จะกระทำซ้ำ 1,000 ครั้งในแต่ละสถานการณ์ของการทดลอง เพื่อประหยัดเวลาในการทำงานของเครื่องคอมพิวเตอร์

คำจำกัดความ

ค่าอำนาจของการทดสอบ (Power) หมายถึง ความน่าจะเป็น (probability) ที่จะปฏิเสธสมมุติฐานสูญเมื่อสมมุติฐานสูญนั้นผิด ซึ่งจะมีค่าเท่ากับ $(1-\beta)$ เมื่อ β คือความคลาดเคลื่อนประเภทที่ 2 (Minium 1978: 364)

ความคลาดเคลื่อนประเภทที่ 2 (Type II error) หมายถึง ความคลาดเคลื่อนที่เกิดขึ้นเนื่องจากยอมรับสมมุติฐานสูญเมื่อสมมุติฐานสูญนั้นผิด

Relative Efficiency ของการทดสอบ P เมื่อเทียบกับการทดสอบ Q หมายถึง ค่า $\frac{m}{n}$ เมื่อการทดสอบ P และ Q ทดสอบสมมุติฐานอย่างเดียวกัน $(\theta=\theta_o)$ และมีอำนาจของการทดสอบเท่ากัน การทดสอบ P จะต้องใช้จำนวนตัวอย่างเท่ากับ n และการทดสอบ Q จะต้องใช้จำนวนตัวอย่างเท่ากับ m (Sawat Pratoomraj 1970: 3)

Asymtotic Relative Efficiency (A.R.E.) ของการทดสอบ P เมื่อเทียบกับการทดสอบ Q หมายถึง ค่าลิมิตของค่า Relative Efficiency ของการทดสอบ P เมื่อเทียบกับการทดสอบ Q เมื่อ n เข้าใกล้อินฟินิตี้ และ θ เข้าใกล้ θ_o (Sawat Pratoomraj 1970: 3)

เดลต้า (δ_i) หมายถึง ความแตกต่างระหว่างค่ามัชฌิมเลขคณิตของประชากร $(\mu_1-\mu_2)$

ความคลาดเคลื่อนประเภทที่ 1 หมายถึง ความคลาดเคลื่อนที่เกิดจากการปฏิเสธสมมุติฐานสูญเมื่อสมมุติฐานสูญนั้นถูก โอกาสที่ทำให้เกิดความคลาดเคลื่อนประเภทที่ 1 กำหนดด้วย α

อัตราความคลาดเคลื่อนประเภทที่ 1 จากผลการทดลอง หมายถึงอัตราความคลาดเคลื่อนประเภทที่ 1 ที่เกิดขึ้นจริงในการทดสอบสมมุติฐานของการทดลองตามแผนการทดลองครั้งนี้

อัตราความคลาดเคลื่อนที่ระบุ หมายถึงอัตราความคลาดเคลื่อนประเภทที่ 1 ที่ควบคุมด้วยระดับนัยสำคัญ α ซึ่งผู้วิจัยกำหนด

ประโยชน์ของการวิจัย

เพื่อช่วยให้ผู้ใช้สถิติมีผลสรุปและหลักฐานในการเลือกใช้สถิติวิเคราะห์ข้อมูลเพื่อเปรียบเทียบความแตกต่างอย่างมีนัยสำคัญทางสถิติของประชากรซึ่งเป็นที่มาของกลุ่มตัวอย่างสุ่มสองกลุ่มซึ่งเป็นอิสระต่อกันได้อย่างมีประสิทธิภาพและเหมาะสมกับสถานการณ์และข้อตกลงเบื้องต้น

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย