บทที่ 2

วรรณคดีที่เกี่ยวข้อง

การวิจัยนี้เป็นการศึกษาเพื่อเปรียบเทียบเทคนิคการสุ่มตัวอย่างผู้สอบกับการสุ่มเมตริกพหุคูณ ผู้วิจัยเสนอวรรณคดีที่เกี่ยวข้องกับเทคนิคการสุ่มตัวอย่างดังกล่าว โดยเสนอเป็น 3 ตอนดังนี้

ตอนที่ 1 ลักษณะทั่วไปของการสุ่มตัวอย่าง

ตอนที่ 2 การสุ่มเมตริกพหุคูณ

ตอนที่ 3 งานวิจัยที่เกี่ยวข้อง

<u>ตอนที่ 1</u> ลักษณะทั่วไปของการสุ่มตัวอย่าง

โลกสมัยปัจจุบันมีความเจริญก้าวหน้าทางวิทยาศาสตร์มาก การกระทำการใด ๆ ย่อมต้องวางแผนให้ถูกต้อง เหมาะสมและสิ้นเปลืองน้อยที่สุด โดยให้ผลตอบแทนมากที่สุด โดยเฉพาะอย่างยิ่งโครงการใหญ่ ๆ ย่อมต้องวางแผนให้รัดกุม การตัดสินใจใด ๆ โดยมีเหตุผลย่อมต้องการข้อเท็จจริงมาประกอบการพิจารณา กิจการอุตสาหกรรม ธุรกิจ องค์การรัฐบาล ต้องการข้อมูลสถิติเพื่อประโยชน์ในการวางแผนและการตัดสินใจ ดังนั้นจึงมีการรวบรวมข้อเท็จจริงทางสถิติ และนำเอาข้อมูลเหล่านั้นมาวิเคราะห์และเสนอให้เหมาะสมเพื่อใช้เป็นรากฐานในการตัดสินใจวางนโยบายปฏิบัติกิจการงานต่าง ๆ (นิยม ปุราคำ 2517:1) ในการรวบรวมข้อมูลนั้นอาจทำได้โดยการสำรวจหรือวัดหน่วยทุกหน่วยที่อยู่ในประชากรที่สนใจ ซึ่งเรียกว่าการสำมะโน หรืออาจจะใช้วิธีการรวบรวมข้อเท็จจริงจากหน่วยตัวอย่าง (Sampling units) ที่เลือกมาเป็นตัวแทนจากประชากรที่สนใจ แล้วใช้ข้อเท็จจริงที่ได้กะประมาณค่าพารามิเตอร์ (Parameters) ที่เราสนใจ ซึ่งเรียกว่าการสำรวจตัวอย่าง หรือการสุ่มตัวอย่าง

การที่จะทำให้ตัวอย่างที่สุ่มมาเป็นตัวแทนของประชากรได้ดีเพียงไรนั้นย่อมขึ้นอยู่กับขนาดของตัวอย่าง เทคนิคการสุ่ม และเครื่องมือในการเก็บข้อมูล ถ้าหากตัวอย่างที่สุ่มมานั้นไม่เป็นตัวแทนที่ดีของประชากร ก็จะทำให้การสรุปเกี่ยวกับคุณลักษณะของประชากรผิดพลาดไป (อนันต์ ศรีโสภา 2524 ข : 1) แต่อย่างไรก็ตามก็เป็นที่ยอมรับกันแล้วว่า

การรวบรวมข้อมูลโดยวิธีการสุ่มตัวอย่างก็ย่อมจะผิดพลาดจากที่เป็นจริงได้ แต่ก็มีวิธีการที่จะทำให้ผิดพลาดน้อยที่สุด และถึงแม้จะใช้วิธีสำมะโนก็ตาม งานบางชนิดการสำมะโนจะมีข้อผิดพลาดสูงมาก โดยทั่วไปแล้วการสำรวจด้วยตัวอย่างจะมีข้อดีเหนือการสำมะโนอยู่หลายประการคือ (นิยม ปุราคำ 2517:3)

1. ลดค่าใช้จ่าย (Reduced cost) ทั้งนี้เพราะว่าจะสำรวจจากหน่วยตัวอย่างที่เลือกได้ แทนที่จะสำรวจจากหน่วยทุกหน่วยในประชากร งานสำรวจด้วยตัวอย่างเป็นงานเล็กกว่าการสำมะโน ค่าใช้จ่ายย่อมน้อยกว่า

2. ทำได้รวดเร็วกว่า (Greater speed) เนื่องจากไม่ต้องคลุมให้ครบทุกหน่วยในประชากร ซึ่งบางทีเรียกว่าคุ้มรวม (Population or universe) จึงสำรวจได้ครบที่ต้องการเร็วกว่าการสำมะโน

3. คลุมครอบได้มากกว่า (Greater scope) เพราะว่าการสำรวจด้วยตัวอย่างเกี่ยวข้องกับหน่วยมีจำนวนน้อยกว่า จึงอาจจะคลุมถึงสาระสำคัญต่าง ๆ ได้มากกว่าหนึ่งอย่างในเวลาเดียวกัน การฝึกพนักงานเจ้าหน้าที่สนามหรือบุคคลที่จะรวบรวมข้อเท็จจริงทำได้ง่ายกว่าเพราะเป็นงานเล็กกว่าการสำมะโน งานสำรวจโดยอาศัยตัวอย่างอาจจะดัดแปลงแก้ไขได้ง่ายกว่าการสำมะโน ข้อมูลชนิดต่าง ๆ ที่ต้องการรวบรวมอาจจะเปลี่ยนแปลงได้ง่ายกว่าการสำมะโน

4. ถูกต้องแม่นยำกว่า (Greater accuracy) เนื่องจากคุณภาพของผู้ร่วมงานและพนักงานสนามในงานสำรวจด้วยตัวอย่างอาจจะเลือกสรรหรือฝึกฝนได้ดีกว่าในงานสำมะโน การให้คำแนะนำ การตรวจสอบคุณภาพงานสนาม การประมวลผลทำได้ดีกว่าเมื่อปริมาณงานน้อย ดังนั้นตัวอย่างอาจจะให้ผลที่ถูกต้องแม่นยำกว่าผลจากการสำมะโน

5. ความจำเป็น (Necessity) บังคับให้ใช้การสำรวจด้วยตัวอย่าง เช่นในกรณีที่ถ้านำหน่วยมาทดสอบแล้วจะทำให้เสื่อมคุณภาพหรือใช้ไม่ได้เลย นอกจากนั้นในกรณีที่ประชากรใหญ่เกินกว่าที่จะใช้การสำมะโน ก็จะต้องใช้การสำรวจด้วยตัวอย่าง

สำหรับวิธีการสุ่มตัวอย่างแบ่งออกได้เป็น 2 ประเภทใหญ่ ๆ คือ การสุ่มตัวอย่างที่ไม่เป็นไปตามโอกาสทางสถิติ (Non-probability sampling) และการสุ่มตัวอย่างที่เป็นไปตามโอกาสทางสถิติ (Probability sampling) การสุ่มตัวอย่างประเภทแรกยังแบ่งออกได้เป็นหลายวิธีคือ การสุ่มตัวอย่างแบบบังเอิญ (Accidental sampling) การสุ่มตัวอย่างแบบโควต้า (Quota sampling) การสุ่มตัวอย่างแบบ

เจาะจง (Purposive sampling) ส่วนการสุ่มตัวอย่างประเภทที่เป็นไปตามโอกาสทางสถิติมีวิธีการสุ่มพื้นฐานที่สำคัญ ๆ เช่น การสุ่มอย่างง่าย (Simple random sampling) การสุ่มตัวอย่างแบบแบ่งชั้น (Stratified random sampling) และการสุ่มตัวอย่างแบบกลุ่ม (Cluster sampling) วิธีการสุ่มตัวอย่างแบบต่าง ๆ ของแต่ละประเภทยังนำมาใช้ร่วมกันได้ (สุชาติ ประสิทธิ์รัฐสินธุ์ ธวัชชัย อาทรธุระสุข และพิสิฐ ศุกรียพงศ์ 2523:85)

ความแตกต่างระหว่างการสุ่มตัวอย่างประเภทที่เป็นไปตามโอกาสทางสถิติและที่ไม่เป็นไปตามโอกาสทางสถิติ คือ การสุ่มตัวอย่างประเภทแรก เราสามารถกำหนดได้ว่า หน่วยแต่ละหน่วยของประชากรมีโอกาสเท่าใดที่จะได้รับการเลือกเข้าไปเป็นตัวอย่าง โอกาสที่ได้รับการเลือกนี้อาจจะเท่ากันหรือไม่เท่ากันทุกหน่วยก็ได้ แต่จะต้องระบุให้ชัดเจนว่าแต่ละหน่วยมีโอกาสได้รับเลือกเท่าใด สำหรับการสุ่มตัวอย่างประเภทที่ไม่เป็นไปตามโอกาสทางสถิติ เราไม่สามารถกำหนดหรือประมาณโอกาสที่ใช้เลือกแต่ละหน่วยของประชากรได้ และไม่มีหลักประกันว่าทุกหน่วยมีโอกาสที่จะได้รับเลือกเป็นตัวอย่าง (สุชาติ ประสิทธิ์รัฐสินธุ์ และคณะ 2523:85)

การสุ่มตัวอย่างที่เป็นไปตามโอกาสเป็นวิธีเดียวที่จะวางแผนตัวอย่างให้เป็นตัวแทนประชากรได้ สามารถกำหนดขนาดตัวอย่างตามอัตราของความเชื่อมั่นว่าตัวอย่างที่ได้จะไม่แตกต่างจากประชากรที่ศึกษามากไปกว่าอัตราที่ต้องการ ส่วนการสุ่มตัวอย่างที่ไม่เป็นไปตามโอกาสมีข้อได้เปรียบในแง่ของความสะดวกและประหยัดกว่าการสุ่มที่เป็นไปตามโอกาส (สุชาติ ประสิทธิ์รัฐสินธุ์ และคณะ 2523:86)

สำหรับงานวิจัยทางสังคมศาสตร์ที่ใช้แบบสอบถามหรือแบบทดสอบต่าง ๆ เป็นเครื่องมือในการเก็บรวบรวมข้อมูลนั้น หากเป็นงานวิจัยที่ใช้การสุ่มตัวอย่าง ผู้วิจัยก็จะต้องมีการกำหนดประชากรที่ต้องการสำรวจ การสร้างกรอบสุ่มตัวอย่าง (Sampling frame) และวิธีการคัดเลือก ในการคัดเลือกก็จะต้องเลือกใช้วิธีการสุ่มตัวอย่าง ซึ่งมีหลายวิธี การจะเลือกใช้วิธีใดย่อมขึ้นอยู่กับวัตถุประสงค์ของการวิจัย ผู้วิจัยจะต้องเลือกเอาวิธีที่จะได้มาซึ่งข้อมูลที่ตอบปัญหาของการวิจัยได้ดีที่สุด (สุชาติ ประสิทธิ์รัฐสินธุ์ และคณะ 2523:84) แต่อย่างไรก็ตามจะเห็นได้ว่าในกรณีที่ใช้แบบสอบถามหรือแบบทดสอบกับกลุ่มตัวอย่างนั้นประชากรหรือกลุ่มตัวอย่างที่กล่าวถึงจะหมายถึงผู้ตอบเท่านั้น

ฉะนั้นผู้วิจัยจึงให้ผู้ตอบที่ได้รับการสุ่มมาเป็นตัวอย่างตอบแบบสอบถาม หรือแบบทดสอบที่สร้างขึ้นมาทั้งฉบับ

ในการวิจัยหรือในการประเมินผลใด ๆ ก็ตามที่ใช้แบบสอบเป็นเครื่องมือรวบรวมข้อมูล บางครั้งแบบสอบเหล่านั้นก็จะมีขนาดยาวมาก เพราะการเพิ่มความยาวของแบบสอบจะช่วยเพิ่มความสม่ำเสมอในการสุ่มเนื้อเรื่อง (Content sampling) และแบบสอบยิ่งยาว (จำนวนข้อมาก) ความเที่ยงก็จะมากขึ้น ค่าสัมประสิทธิ์แห่งความเที่ยงของแบบสอบที่ยาวขึ้นหรือสั้นลงจะหาได้จากสูตรเสปียร์แมน-บราวน์ (Spearman-Brown) ดังนี้ (อนาสตาซี 2519:83)

$$r_{tt} = \frac{nr'_{tt}}{1+(n-1)r'_{tt}}$$

เมื่อ r_{tt} เป็นค่าสัมประสิทธิ์ที่ประมาณได้

r'_{tt} เป็นค่าสัมประสิทธิ์ที่ได้จากคะแนนทดสอบ

n เป็นจำนวนเท่าของแบบสอบที่ยาวขึ้นหรือสั้นลง

จากการใช้สูตรเสปียร์แมน-บราวน์ จะพบความสัมพันธ์ระหว่างความเที่ยงและจำนวนข้อในแบบสอบดังแสดงในตารางที่ 1 (อนันต์ ศรีโสภา 2524 ก : 61)

ตารางที่ 1 ความสัมพันธ์ระหว่างความเที่ยงและจำนวนข้อในแบบสอบ

จำนวนข้อสอบ	ความเที่ยง
5	0.20
10	0.33
20	0.50
40	0.67
80	0.80
160	0.89
320	0.94
640	0.97
	1.00

จากตารางจะเห็นว่า เมื่อแบบสอบมีจำนวนข้อมากขึ้นจะมีค่าความเที่ยงสูงขึ้น ทำนองเดียวกันถ้าแบบสอบมีจำนวนข้อลดลงก็จะมีค่าความเที่ยงลดลงด้วย แต่การเพิ่มหรือลดมิได้เป็นสัดส่วนโดยตรงกับจำนวนข้อที่เพิ่มขึ้นหรือลดลง (เอนก เพียรอนุกูลบุตร 2524:449) สำหรับสูตรของเสปียร์แมน-บราวน์นี้ตั้งอยู่บนข้อสมมติ (assumption) สองประการด้วยกัน คือ ข้อสมมติที่เกี่ยวกับสถิติ จำนวนข้อที่นำไปเพิ่มเข้านั้นจะต้องมีคุณสมบัติทางสถิติเหมือนกับข้อสอบเดิม นั่นคือ ข้อสอบที่เพิ่มเข้าไปนั้นจะต้องมีค่าความยากเฉลี่ยเท่ากับค่าความยากเฉลี่ยของข้อสอบเดิม และการเพิ่มจำนวนข้อสอบเข้าไปนี้จะต้องไม่ทำให้ค่าเฉลี่ยของปฏิสัมพันธ์ (Interaction) ระหว่างข้อเปลี่ยนแปลงไป ข้อสมมติประการที่สอง การเพิ่มจำนวนข้อสอบเข้าไปนั้น จะต้องไม่ทำให้วิธีการตอบของผู้ตอบเปลี่ยนแปลงไป โดยเฉพาะเกี่ยวกับคำสั่งชี้แจง และชนิดของข้อสอบจะต้องเหมือนเดิมทุกประการ แต่ในทางปฏิบัติเรามักจะพบว่า เมื่อเราเพิ่มจำนวนข้อสอบให้มากขึ้นนักเรียนจะต้องใช้เวลาทำนานขึ้น นักเรียนมักจะเหนื่อยและเกิดความเบื่อหน่ายในการทำข้อสอบ (อนันต์ ศรีโสภา 2524 ก : 62)

ในทำนองเดียวกัน แบบสอบถามในงานวิจัยบางเรื่องก็มีขนาดยาวคือ มีคำถามมากข้อเพื่อให้ครอบคลุมเนื้อหา เรื่องราวที่จะถาม ฉะนั้นปัญหาที่ตามมาก็คือผู้ตอบจะต้องใช้เวลาทำนานขึ้นและก็จะเกิดความเบื่อหน่ายในการตอบเช่นเดียวกัน แต่อย่างไรก็ตามในบางครั้งก็เป็นการยากที่จะหลีกเลี่ยงการใช้เครื่องมือที่มีขนาดยาวไม่ว่าจะเป็นแบบทดสอบหรือแบบสอบถาม แต่ในกรณีที่เป็นการศึกษาเกี่ยวกับกลุ่ม สิ่งที่ต้องการใช้สำหรับการสรุปผลหรือการตัดสินใจก็คือค่าพารามิเตอร์ ซึ่งเป็นค่าที่บอกให้ทราบถึงลักษณะประชากรแต่ไม่อธิบายเป็นรายบุคคล ฉะนั้นจึงได้มีผู้เสนอวิธีการสุ่มข้อสอบ (Item sampling) โดยมีจุดมุ่งหมายเพื่อที่จะลดเวลาในการทำแบบสอบ ลอร์ดเป็นผู้ที่ได้ชื่อว่าริเริ่มศึกษาเรื่องการสุ่มข้อสอบอย่างจริงจัง เริ่มตั้งแต่ในปี 1955 โดยได้เขียนบทความเพื่อเสนอแนะเทคนิคการสุ่มข้อสอบนี้ลงในวารสารไซโคเมตริกา (Psychometrika) (Petersen 1968:135)

หลังจากที่ลอร์ดได้เสนอเทคนิคการสุ่มข้อสอบขึ้นมาแล้ว ต่อมาก็ได้มีการรวบรวมการสุ่มข้อสอบและการสุ่มผู้สอบเข้าด้วยกัน โดยการสุ่มทั้งข้อสอบและผู้สอบพร้อม ๆ กัน ซึ่งวิธีการนี้เรียกว่าการสุ่มข้อสอบ-ผู้สอบ (Item-examinee sampling) หรือการสุ่มเมตริก (Matrix sampling) นั่นเอง และในปี 1962 ลอร์ดก็ได้เสนอวิธีการขั้นพื้นฐาน รวมทั้งสมการต่าง ๆ และการหาค่าสถิติเพื่อประมาณค่ามัชฌิมเลขคณิตและความแปรปรวนของประชากรด้วย (Kleinke 1972:75)

ตอนที่ 2 การสุ่มเมตริกพหุคูณ

ผู้วิจัยได้ศึกษาค้นคว้าเรื่องการสุ่มเมตริกพหุคูณจากหนังสือและเอกสารต่าง ๆ ซึ่งจะนำความรู้เกี่ยวกับเทคนิคนี้เสนอโดยแยกเป็นหัวข้อย่อยดังต่อไปนี้

ก. ความรู้พื้นฐานเกี่ยวกับการสุ่มเมตริก

ข. ประโยชน์ของการสุ่มเมตริก

ค. ข้อจำกัดของการสุ่มเมตริก

ง. การนำเทคนิคการสุ่มเมตริกพหุคูณไปใช้

จ. ลำดับขั้นการใช้การสุ่มเมตริกพหุคูณ

ฉ. การคำนวณในเรื่องการสุ่มเมตริกพหุคูณ ซึ่งครอบคลุมเนื้อหาดังนี้คือ

การคำนวณค่าสถิติจากตัวอย่างเมตริก

การประมาณค่าพารามิเตอร์

ความคลาดเคลื่อนมาตรฐาน

การคำนวณค่าสัมประสิทธิ์แอลฟาของครอนบาค

ความรู้พื้นฐานเกี่ยวกับการสุ่มเมตริก (Sirotnik 1974:454-469)

ถ้านำคะแนนที่ได้จากประชากรผู้สอบ N คน แต่ละคนทำแบบสอบทั้งฉบับ ซึ่งประกอบด้วยประชากรข้อสอบจำนวน M ข้อมาศึกษา ข้อมูลต่าง ๆ ที่ได้ก็จะเป็นข้อมูลจากประชากร ซึ่งในเรื่องการสุ่มเมตริกเรียกว่า เมตริกประชากร (Matrix population) ดังแสดงในภาพที่ 1

ภาพที่ 1 คะแนนที่ปรากฏในเมตริกประชากร
เมื่อประชากรผู้สอบมีจำนวน N คน ประชากรข้อสอบมีจำนวน M ข้อ และ x_{ij} คือคะแนนของผู้ตอบคนที่ i ข้อกระทงที่ j

ข้อสอบ

ผู้สอบ

$$
\begin{array}{c|ccccccc|}
 & 1 & 2 & 3 & \cdots & j & \cdots & M \\
\hline
1 & x_{11} & x_{12} & x_{13} & \cdots & x_{1j} & \cdots & x_{1M} \\
2 & x_{21} & x_{22} & x_{23} & \cdots & x_{2j} & \cdots & x_{2M} \\
3 & x_{31} & x_{32} & x_{33} & \cdots & x_{3j} & \cdots & x_{3M} \\
\vdots & \vdots & \vdots & \vdots & & \vdots & & \vdots \\
i & x_{i1} & x_{i2} & x_{i3} & \cdots & x_{ij} & \cdots & x_{iM} \\
\vdots & \vdots & \vdots & \vdots & & \vdots & & \vdots \\
N & x_{N1} & x_{N2} & x_{N3} & \cdots & x_{Nj} & \cdots & x_{NM} \\
\hline
\end{array}
$$

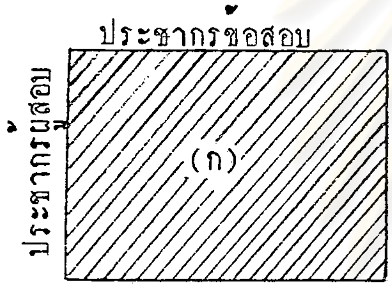
ในกรณีที่การศึกษาจากประชากรเป็นสิ่งที่ทำได้ยาก ก็จะต้องนำการสุ่มตัวอย่างมาใช้ ลักษณะการสุ่มที่เป็นไปได้มี 3 แบบ (แสดงในรูป ข ค และ ง ในภาพที่ 2) ได้แก่

1. การสุ่มผู้สอบ หรือการสุ่มผู้ตอบ (Examinee sampling) วิธีนี้จะสุ่มเฉพาะผู้ตอบ และให้ผู้ตอบที่ได้รับการสุ่มมาทำแบบสอบทั้งฉบับ ซึ่งวิธีการนี้ก็คือวิธีการสุ่มโดยทั่วไปที่นิยมปฏิบัติกันมานั่นเอง

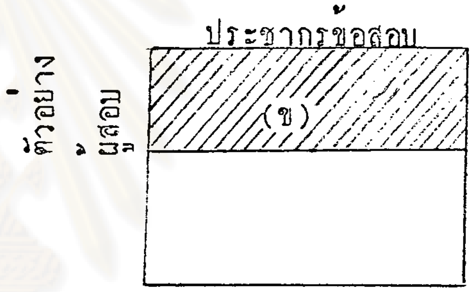
2. การสุ่มข้อสอบหรือการสุ่มข้อกระทง (Item sampling) วิธีนี้จะสุ่มข้อสอบ และให้ประชากรที่ต้องการศึกษาทั้งหมดทำแบบสอบที่ประกอบด้วยตัวอย่างข้อสอบที่ได้รับการสุ่มมา (เรื่องการสุ่มเมตริกในระยะเริ่มแรกนั้น เรียกการสุ่มเมตริกว่า การสุ่มข้อสอบ)

3. การสุ่มเมตริก (Matrix sampling) เป็นการสุ่มทั้งผู้สอบ และข้อสอบ บางทีเรียกการสุ่มเมตริกว่า การสุ่มข้อสอบ-ผู้สอบ (Item-examinee sampling) สำหรับกลุ่มคะแนนที่ได้จากตัวอย่างผู้สอบ 1 กลุ่ม และตัวอย่างข้อสอบ 1 ชุด เรียกว่า ตัวอย่างเมตริก (Matrix sample)

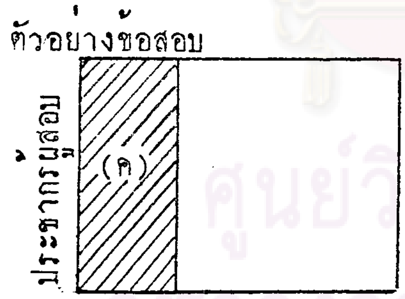
ภาพที่ 2 ลักษณะการสุ่มแบบต่าง ๆ



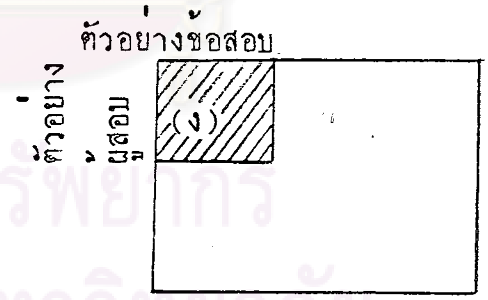
รูป (ก) เมตริกประชากร (Matrix population)



รูป (ข) การสุ่มผู้สอบ(Examinee sampling)



รูป (ค) การสุ่มข้อสอบ(Item sampling)

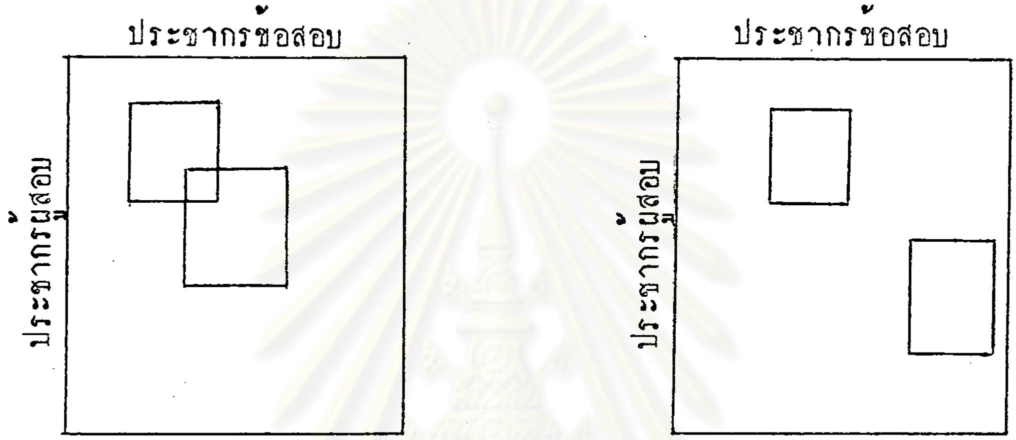


รูป (ง) การสุ่มเมตริก(Matrix sampling)

ถ้ามีการสุ่มตัวอย่างเมตริกหลาย ๆ ตัวอย่างในเมตริกประชากรเดียวกัน เรียกว่า การสุ่มเมตริกพหุคูณ (Multiple matrix sampling)

สำหรับการสุ่มในเรื่องการสุ่มเมตริกนี้อาจทำได้ทั้งการสุ่มแบบคืนที่ (with replacement) และการสุ่มแบบไม่คืนที่ (without replacement) การสุ่มแบบคืนที่จะทำให้เกิดการสุ่มเมตริกพหุคูณแบบคาบเกี่ยวกัน (Overlapping multiple matrix sampling) ส่วนการสุ่มแบบไม่คืนที่ทั้งการสุ่มผู้สอบและการสุ่มข้อสอบจะทำให้ได้การสุ่มเมตริกพหุคูณแบบไม่คาบเกี่ยวกัน (Nonoverlapping multiple matrix sampling) ซึ่งแสดงในภาพที่ 3

ภาพที่ 3 ภาพการสุ่มเมตริกพหุคูณแบบคาบเกี่ยวกัน และการสุ่มเมตริกพหุคูณแบบไม่คาบเกี่ยวกัน

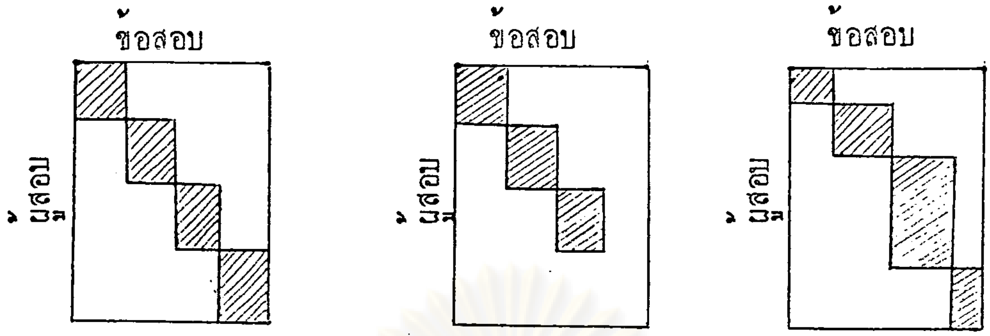


3.1 การสุ่มเมตริกพหุคูณแบบคาบเกี่ยวกัน 3.2 การสุ่มเมตริกพหุคูณแบบไม่คาบเกี่ยวกัน

สำหรับการวิจัยนี้จะกล่าวถึงแต่การสุ่มเมตริกพหุคูณแบบไม่คาบเกี่ยวกันเท่านั้น เพราะการสุ่มแบบนี้เป็นที่นิยมใช้กัน เนื่องจากในทางปฏิบัตินั้น วิธีการสุ่มแบบนี้จะทำได้ง่ายและสะดวกกว่า

ในการสุ่มเมตริกพหุคูณนั้น ตัวอย่างเมตริกในแต่ละเมตริกประชากรเดียวกันที่สุ่มมา อาจจะมีขนาดเท่ากัน (Equal-sized) หรือไม่เท่ากัน (Unequal-sized) ก็ได้ และการสุ่มอาจทำแบบครบหมด (exhausive) คือผู้สอบทุกคนในประชากรผู้สอบได้รับการสุ่มครบทุกคน หรืออาจกล่าวว่า ผู้สอบแต่ละคนในประชากรผู้สอบจะต้องทำตัวอย่างข้อสอบ (แบบสอบชุดย่อย) ชุดใดชุดหนึ่งทุกคน และข้อสอบก็อาจได้รับการสุ่มแบบครบหมดได้ในทำนองเดียวกัน หรือจะสุ่มแบบไม่ครบ (Non-exhausive) ก็ได้ ในภาพที่ 4 แสดงถึงลักษณะการสุ่มเมตริกพหุคูณแบบต่าง ๆ ซึ่งแสดงไว้พอเข้าใจเพียง 5 แบบ ลักษณะของการสุ่มเมตริกพหุคูณยังมีแบบต่าง ๆ อีกมาก ผู้ที่จะนำไปใช้สามารถวางแผนการสุ่มได้ตามความเหมาะสม

ภาพที่ 4 ลักษณะการสุ่มเมตริกพหุคูณแบบต่าง ๆ



ก. สุ่มแบบครบหมด ทั้งประชากรข้อสอบ และประชากรผู้สอบ โดยแต่ละตัวอย่างเมตริก มีขนาดเท่ากัน

ข. สุ่มแบบไม่ครบ ทั้งประชากรข้อสอบ และประชากรผู้สอบ แต่ละตัวอย่างเมตริก มีขนาดเท่ากัน

ค. สุ่มแบบครบหมด ทั้งประชากรข้อสอบ และประชากรผู้สอบ แต่ละตัวอย่างเมตริก มีขนาดไม่เท่ากัน

ง. สุ่มประชากรข้อสอบครบหมด สุ่มประชากรผู้สอบไม่ครบ แต่ละตัวอย่างเมตริกมีขนาดเท่ากัน

จ. สุ่มประชากรข้อสอบไม่ครบ สุ่มประชากรผู้สอบครบหมด แต่ละตัวอย่างเมตริกมีขนาดไม่เท่ากัน

<u>ประโยชน์ของการสุ่มเมตริก</u>

โพแฟม (Popham 1975:229) ได้เสนอประโยชน์ของการสุ่มเมตริกไว้ 4 ประการคือ

1. ลดเวลาในการสอบของนักเรียน (Reduced Testing Time per Student)

2. ทำให้ความคลาดเคลื่อนมาตรฐานของค่าประมาณลดน้อยลง

(Smaller Standard Error of Estimate)

3. ใช้ในการวัดผลที่ต้องใช้ข้อสอบที่มีปริมาณมาก

(Measuring Large Item Domains)

4. ลดความเกรงกลัวการสอบของผู้สอบแต่ละคน

(Less Threatening to Individual Examinees)

ชูเมคเกอร์ (Shoemaker 1973:5-8) ได้เสนอประโยชน์ของการสุ่มเมตริกพหุคูณไว้ดังต่อไปนี้

1. ลดค่าความคลาดเคลื่อนมาตรฐานของค่าประมาณ

(Reduced Standard Error of Estimate)

2. ใช้กับการสอบที่มีประชากรข้อสอบขนาดใหญ่

(Testing Large Item Universes)

3. ลดเวลาในการสอบ

(Reduction in Testing Time)

4. ประโยชน์อื่น ๆ

(Additional Advantages)

จากหัวข้อที่ระบุไว้ในด้านประโยชน์ของการสุ่มเมตริกที่เสนอโดย โพแฟ่ม และชูเมคเกอร์ ดังกล่าวแล้วนั้น จะเสนอเป็นรายละเอียดดังนี้

<u>ลดเวลาในการสอบ</u>

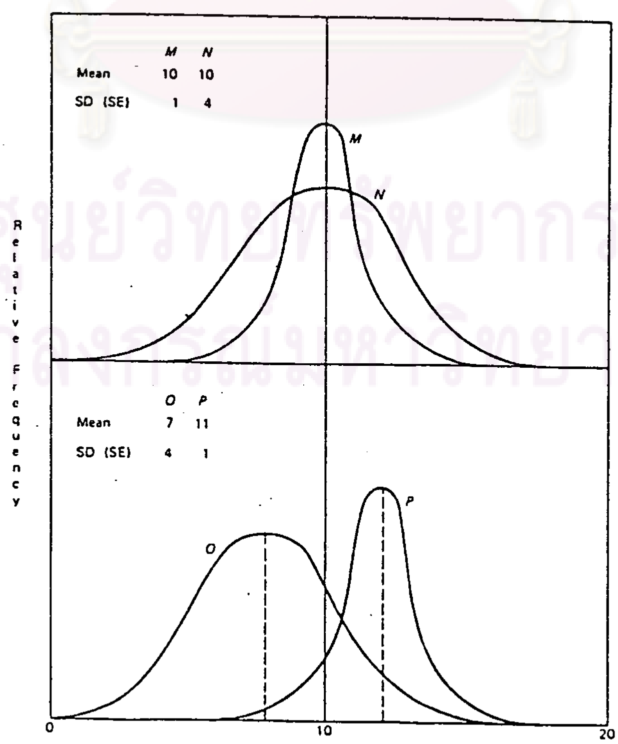
การสุ่มเมตริกที่มีแผนการสุ่มแบบใด ๆ ก็ตามจะช่วยลดเวลาในการสอบของผู้สอบแต่ละคนลงไปได้ ทั้งนี้เพราะผู้สอบไม่ต้องทำข้อสอบทุกข้อ ซึ่งนอกจากจะช่วยลดเวลาในการสอบของผู้สอบแต่ละคนแล้วยังจะช่วยลดเวลาในการดำเนินการสอบ การตรวจข้อสอบ และยังลดค่าใช้จ่ายต่อผู้สอบแต่ละคนอีกด้วย โดยที่เทคนิคนี้เชื่อว่าจะทำให้ได้ค่าประมาณที่มีความเที่ยงตรง (Precision) ในระดับเดียวกับการใช้เทคนิคการสุ่มแบบอื่น ๆ

ลดค่าความคลาดเคลื่อนมาตรฐานของค่าประมาณ

การใช้เทคนิคการสุ่มเมตริกพหุคูณจะทำให้ได้ค่าความคลาดเคลื่อนมาตรฐานของค่าประมาณลดน้อยลงกว่าการใช้การสุ่มผู้สอบ การสุ่มข้อสอบ หรือการสุ่มข้อสอบ-ผู้สอบ (การสุ่มเมตริกที่สุ่มเพียง 1 ตัวอย่างเมตริก) ซึ่งได้มีการแสดงให้เห็นโดยใช้ทั้งขบวนการทางพีชคณิต และการวิจัยเชิงประจักษ์

สำหรับความแตกต่างระหว่างความคลาดเคลื่อนมาตรฐานของค่าประมาณที่แสดงไว้ในภาพที่ 4 (Shoemaker 1973:6) นั้น จะพบว่าการใช้การสุ่มเมตริกพหุคูณ (M) มีความคลาดเคลื่อนมาตรฐานน้อยกว่าการใช้การสุ่มผู้สอบ (N) การสุ่มข้อสอบ (O) หรือการสุ่มข้อสอบ-ผู้สอบ (P) ทั้งนี้ในการเปรียบเทียบนั้นจะต้องให้จำนวนค่าสังเกต (Obsevations) ของแต่ละวิธีมีจำนวนคงที่เท่ากัน (ค่าจำกัดความของ "จำนวนค่าสังเกต" ในที่นี้คือ จำนวนผลรวมของคำตอบของผู้สอบทุกคน หรือจำนวนผลคูณของจำนวนผู้สอบกับจำนวนข้อสอบ เมื่อผู้สอบแต่ละคนทำข้อสอบจำนวนเท่ากัน)

ภาพที่ 5 **การแจกแจงเชิงทฤษฎี** (Hypothetical Distribution) ของค่าสถิติที่ได้จากการทดลองซ้ำ 1,000 ครั้ง ของวิธีการสุ่ม 4 แบบ



ลอร์ดและโนวิค (Lord and Novick 1968:225-226) ได้ใช้วิธีการทางพีชคณิตแสดงให้เห็นว่า เมื่อแบบสอบชุดย่อยประกอบด้วยข้อสอบที่สุ่มมาแบบไม่คืนที่จากประชากรข้อสอบ ความคลาดเคลื่อนมาตรฐานในการประมาณค่ามัชฌิมเลขคณิต โดยใช้การสุ่มเมตริกพหุคูณจะน้อยกว่าความคลาดเคลื่อนมาตรฐานที่ได้จากการสุ่มผู้สอบ และได้พบว่าความคลาดเคลื่อนมาตรฐานของค่าประมาณมัชฌิมเลขคณิตจะมีค่าน้อยที่สุดเมื่อใช้การสุ่มเมตริกพหุคูณที่แบบสอบชุดย่อย 1 ชุด ประกอบด้วยข้อสอบเพียง 1 ข้อ และสุ่มให้ผู้สอบจำนวนมาก ($N \longrightarrow \infty$) แต่ละคนทำข้อสอบคนละ 1 ชุด (1 ข้อ) ซึ่งลอร์ดและโนวิคสรุปว่าวิธีนี้เป็นวิธีที่ดีที่สุดสำหรับการประมาณค่ามัชฌิมเลขคณิต แต่ได้แนะนำว่าวิธีนี้ไม่ใช่วิธีที่ดีที่สุดสำหรับการประมาณค่าความแปรปรวนหรือค่าโมเมนต์เข้าสู่ศูนย์กลาง (Central Moments) อื่น ๆ ได้แก่ ค่าสัมประสิทธิ์ของความเบ้ (Skewness) และค่าสัมประสิทธิ์ของความโด่งแบน (Kurtosis)

สำหรับการวิจัยเกี่ยวกับเรื่องการสุ่มเมตริกนี้นิยมใช้วิธีโพสมอร์เทม (Postmortem) ซึ่งเป็นวิธีที่ผู้วิจัยจะให้ผู้สอบทุกคน (ประชากรผู้สอบ) ทำข้อสอบทุกข้อ (ประชากรข้อสอบ) จากนั้นผู้วิจัยจะนำข้อมูลที่รวบรวมได้มาศึกษาโดยสุ่มตัวอย่างผู้สอบและตัวอย่างข้อสอบตามแผนการสุ่มที่ต้องการศึกษา ซึ่งวิธีการนี้จะทำให้ผู้วิจัยทราบค่าพารามิเตอร์ทั้งหมด และจากการวิจัยโดยใช้วิธีดังกล่าว เช่น การวิจัยของจอห์นสัน และลอร์ด, 1958 ลอร์ด, 1962 พลัมลี, 1964 คุก และสตัฟเฟิลบีม 1967 และชูเมคเกอร์, 1970 (Shoemaker 1973:7) ต่างก็ได้ข้อค้นพบที่สนับสนุนว่าค่าประมาณพารามิเตอร์ที่ได้จากวิธีการสุ่มเมตริกพหุคูณไม่อคติ (Unbias) แต่อย่างไรก็ตาม ชูเมคเกอร์ (1973:7) ได้กล่าวถึงค่าประมาณพารามิเตอร์อื่น ได้แก่ ค่าโมเมนต์ที่ 3 และที่ 4 ว่าเป็นค่าประมาณที่มีอคติสูง แต่ตัวประมาณค่าดังกล่าวนี้จะมีประสิทธิภาพมากขึ้น ถ้าสามารถทราบปริมาณและทิศทางของค่าอคติ ซึ่งจะมีประโยชน์ในการนำไปใช้

<u>ใช้กับการสอบที่ต้องใช้ข้อสอบจำนวนมาก</u>

เมื่อแบบสอบประกอบด้วยข้อสอบจำนวนมาก การใช้เทคนิคการสุ่มเมตริกพหุคูณจะเป็นวิธีการที่ทำให้ได้ค่าประมาณพารามิเตอร์ที่สมเหตุสมผลมากกว่าการใช้การสุ่มข้อสอบ หรือการสุ่มข้อสอบ-ผู้สอบ

ประโยชน์อื่น ๆ

การนำเทคนิคการสุ่มเมตริกไปใช้ในการสุ่มตัวอย่าง อาจมีประโยชน์ในเรื่องต่าง ๆ ต่อไปนี้

1. ช่วยลดความเกรงกลัวการสอบของผู้สอบ การใช้เทคนิคการสุ่มเฉพาะผู้สอบนั้นจำนวนข้อคำถามจะมีจำนวนมากจนน่าเกรงกลัว แต่ถ้าใช้เทคนิคการสุ่มเมตริกแล้วผู้สอบหรือผู้ตอบจะได้รับข้อสอบหรือข้อคำถามจำนวนไม่มาก ซึ่งจะไม่ทำให้เกิดความเกรงกลัวความยาวของแบบทดสอบหรือแบบสอบถาม นอกจากนั้นการใช้เทคนิคนี้ยังทำให้ผู้สอบทราบว่าไม่มีการนำคะแนนของตนไปเปรียบเทียบกับคะแนนของผู้อื่น จึงทำให้ผู้สอบลดความเกรงกลัวการสอบลงไปอีกด้วย

2. การจำกัดเวลาที่ใช้ในการสอบของผู้สอบแต่ละคน โดยใช้เทคนิคการสุ่มเมตริก เมื่อเป็นการสอบที่ประชากรข้อสอบมีขนาดใหญ่จะได้ผลดีกว่าการใช้วิธีการที่เรียกว่าแมทช์-ไอเทม ดีไซน์ (Match-Item Design)

3. นักวัดผลอาจจะประมาณค่าพารามิเตอร์ของแบบสอบ (ประชากรข้อสอบ) หลาย ๆ กลุ่มได้พร้อมกัน โดยที่แบบสอบชุดย่อยแต่ละชุดที่แจกให้แก่ผู้สอบแต่ละคนนั้นประกอบด้วยข้อสอบที่มาจากประชากรข้อสอบต่างกลุ่มกัน

4. การสุ่มเมตริกไม่จำเป็นจะต้องใช้กับงานที่ประกอบด้วยผู้สอบ (แนวขวาง) และข้อสอบ (แนวตั้ง) เท่านั้น แต่สามารถใช้กับงานด้านอื่น ๆ ได้อีกที่เรียกว่าหลักการวัตถุ-ตัวแปร (Object-Variable Rationale) ซึ่งข้อมูลประกอบด้วย วัตถุ (แนวขวาง) และตัวแปร (แนวตั้ง) เช่น เครื่องจักรและการตรวจสอบควัน ปลาและอาการของโรค หินและลักษณะความแข็ง เป็นต้น (Shoemaker 1973:11)

5. ผู้ดำเนินการสอบอาจจะไม่รู้สึกเบื่อเมื่อใช้เทคนิคการสุ่มเมตริกพหุคูณ เพราะผู้สอบทำแบบสอบชุดย่อยที่แตกต่างกัน

ข้อจำกัดของการสุ่มเมตริก

โพแฟม (Popham 1975:230) ได้รวบรวมข้อจำกัดของการสุ่มเมตริก ซึ่งมีดังต่อไปนี้

1. ไม่สามารถทราบคะแนนของผู้สอบแต่ละคนได้ (Performances of Individuals Unknown)

2. ความยุ่งยากในการดำเนินการสอบเนื่องจากการใช้แบบสอบหลายชุด (Logistics of Administering Multiple Tests)

3. เกิดผลกระทบจากบริบทเมื่อใช้แบบสอบอาศัยความเร็ว (Context Effects for Speeded Tests)

4. ความจำเป็นที่จะต้องใช้เครื่องคอมพิวเตอร์ในการประมวลผลข้อมูล (Data-Processing Requirements)

5. การไม่เน้นเรื่องการวัดผลแต่ละบุคคลมีผลต่อการทำแบบสอบอย่างเต็มความสามารถ (Impersonal Measurement and Optimal Performance)

ชูเมคเกอร์ (Shoemaker 1973:8-11) ได้กล่าวถึงข้อจำกัดของการสุ่มเมตริกพหุคูณ ซึ่งจะพบในเรื่องต่าง ๆ ดังต่อไปนี้

1. ผลกระทบจากบริบท (Context Effect)

2. การสุ่มผู้สอบ (Sampling of Examinees)

√3. การประมาณคะแนนประชากรข้อสอบของผู้สอบแต่ละคน โดยใช้คะแนนจากตัวอย่างข้อสอบ (Estimating the Universe Score for Each Examinee)

4. ความยุ่งยากในการใช้แบบสอบชุดย่อยหลายชุด (Logistics of Multiple Subtests)

5. กระบวนการทางสถิติอันยุ่งยากก่อให้เกิดความจำเป็นที่จะต้องใช้เครื่องคอมพิวเตอร์ (Necessary Statistical Procedures)

ไซรอทนิค (Sirotnik 1974:473-476) ได้บรรยายเกี่ยวกับข้อจำกัดของการสุ่มเมตริก ซึ่งสามารถสรุปได้ดังต่อไปนี้

1. เทคนิคนี้ใช้สำหรับการประมาณค่าพารามิเตอร์ของกลุ่ม และจะยุ่งยากมากถ้าจะใช้ข้อมูลหรือคะแนนที่รวบรวมโดยเทคนิคนี้ในการศึกษารายบุคคล

2. การนำเทคนิคนี้ไปใช้จะต้องตั้งอยู่บนข้อตกลงที่ว่า การตอบข้อสอบข้อหนึ่งข้อใดที่ปรากฏอยู่ในกลุ่มตัวอย่างข้อสอบ ต้องเป็นไปในทำนองเดียวกับการที่ข้อสอบนั้นอยู่ในประชากรข้อสอบ ฉะนั้นจึงนำเทคนิคนี้ไปใช้กับแบบสอบอาศัยความเร็วไม่ได้

3. ทฤษฎีที่จะใช้เป็นแนวทางยังไม่ชัดเจน (Nonexistence of Theoretical Guidelines)

4. ปัญหาเกี่ยวกับการคำนวณที่ยุ่งยาก จึงทำให้มีความจำเป็นที่จะต้องใช้เครื่องคอมพิวเตอร์

5. ความยุ่งยากในการสร้างแบบสอบและดำเนินการสอบ

รายละเอียดของข้อจำกัดของการสุ่มเมตริกในด้านต่าง ๆ มีดังต่อไปนี้

1. ข้อมูลที่รวบรวมโดยใช้เทคนิคนี้ไม่สามารถนำไปใช้ในกรณีที่ต้องการศึกษาเป็นรายบุคคล

เนื่องจากการใช้เทคนิคนี้ ผู้สอบจะทำข้อสอบเพียงบางข้อที่สุ่มมาจากแบบสอบฉบับเต็ม ฉะนั้นจึงไม่ทราบว่าถ้าผู้สอบคนนี้ทำแบบสอบทั้งฉบับแล้วจะได้คะแนนเท่าใด ด้วยเหตุนี้ข้อมูลที่รวบรวมโดยเทคนิคนี้จึงไม่สามารถนำไปใช้ในกรณีที่ต้องการศึกษาผู้สอบเป็นรายบุคคลได้ อย่างไรก็ตามได้มีการศึกษาเพื่อหาวิธีการที่จะประมาณคะแนนประชากรข้อสอบ (Universe Score) ของผู้สอบแต่ละคนโดยใช้คะแนนที่ได้จากการทำตัวอย่างข้อสอบ เช่น การศึกษาของคลีน (Kleinke 1972:75-84) บันดา (Bunda 1973:117-130) ซาชาร์ และซัพเพส (Sachar and Suppes 1980:687-699) เป็นต้น

2. การใช้เทคนิคนี้ต้องคำนึงถึงผลกระทบจากบริบท (Context Effect)

ข้อตกลงประการหนึ่งในการนำเทคนิคนี้ไปใช้ก็คือ การตอบข้อสอบข้อหนึ่งข้อใดที่ปรากฏอยู่ในกลุ่มตัวอย่างข้อสอบ ต้องเป็นไปในทำนองเดียวกับการที่ข้อสอบนั้นอยู่ในประชากรข้อสอบ (Sirotnik 1974:474) แต่ถ้าการตอบข้อสอบเมื่อข้อสอบข้อนั้นอยู่ในประชากรข้อสอบ แตกต่างจากผลการตอบเมื่อข้อสอบนั้นอยู่ในตัวอย่างข้อสอบแล้ว เรียกว่าเกิดผลกระทบจากบริบท (Popham 1975:230)

การเกิดผลกระทบจากบริบทเป็นเรื่องสำคัญมากเรื่องหนึ่งในเรื่องการใช้เทคนิคการสุ่มเมตริก เพราะถ้าเกิดผลกระทบจากบริบทก็จะมีผลกระทบไปถึงค่าประมาณพารามิเตอร์ต่าง ๆ อีกด้วย จึงได้มีผู้ศึกษาเกี่ยวกับเรื่องนี้เช่น การศึกษาของคาเฮน รอมเบิร์กและซไวร์เนอร์ (Cahen, Romberg and Zwirner 1970:41-60) เฟลดท์และฟอร์สิท (Feldt and Forsyth 1974:73-82) ซึ่งผลจากการวิจัยสรุปได้ว่าค่าประมาณพารามิเตอร์ที่คำนวณได้มีค่าใกล้เคียงกันไม่ว่าคะแนนจากข้อสอบที่ใช้ในการวิจัยจะเป็นคะแนนที่ได้มาจากการตอบแบบสอบเพียงบางส่วน หรือทำแบบสอบทั้งฉบับ

สำหรับในกรณีแบบสอบอาศัยความเร็วแล้วไม่ควรใช้เทคนิคการสุ่มเมตริก เพราะจะมีผลกระทบจากบริบทมากเนื่องจากแบบสอบดังกล่าวประกอบด้วยข้อสอบที่มีความยากต่ำโดยตลอด ข้อสอบทั้งหมดออกแบบไว้ภายในระดับความสามารถของผู้ที่ถูกสอบ เวลาที่กำหนดให้ค่อนข้างสั้นจนไม่มีใครทำเสร็จภายในเวลาที่กำหนดให้ (อนาสตาซี 2519:89) และการที่ผู้สอบทำข้อสอบไม่ครบนี้เอง จะมีผลต่อการเกิดผลกระทบจากบริบท

3. ทฤษฎีที่จะใช้เป็นแนวทางยังไม่ชัดเจน

ปัญหาอีกประการหนึ่งสำหรับการนำเทคนิคการสุ่มเมตริกพหุคูณไปใช้ก็คือ การกำหนดแผนการสุ่มให้มีประสิทธิภาพที่สุดสำหรับงานนั้น ๆ ทั้งนี้เพราะยังขาดการค้นคว้าวิจัยในด้านนี้ คำถามที่จะพบเมื่อนำเทคนิคนี้ไปใช้ก็คือ จำนวนตัวอย่างเมตริกที่เหมาะสมควรจะเป็นเท่าใด แต่ละตัวอย่างเมตริกควรจะประกอบด้วยจำนวนผู้ตอบและจำนวนข้อกระทงเท่าใด การสุ่มควรจะเป็นการสุ่มแบบครบหมดหรือสุ่มไม่ครบ และควรจะเป็นการสุ่มแบบคืนที่หรือไม่คืนที่ (Sirotnik 1974:475)

4. การสุ่มผู้สอบ

ในการนำเทคนิคการสุ่มเมตริกไปใช้ได้มีข้อตกลงว่า แบบสอบชุดย่อยแต่ละชุดจะนำไปใช้กับกลุ่มตัวอย่างผู้สอบซึ่งผู้สอบแต่ละคนจะสุ่มมาจากประชากรผู้สอบ ฉะนั้นรายชื่อประชากรที่จะใช้ในการศึกษาจะต้องครบถ้วนซึ่งบางครั้งเป็นเรื่องที่กระทำได้ยาก ดังนั้นจึงมีการเลือกใช้วิธีการสุ่มตัวอย่างแบบอื่นบ้าง เช่น การสุ่มตัวอย่างแบบแบ่งกลุ่มสองชั้นและการสุ่มตัวอย่างแบบกลุ่ม (Two-stage Sampling and Cluster Sampling) ซึ่งทำให้ตัวอย่างของกลุ่มย่อยคือ ชั้นเรียน โรงเรียน หรือเขตการศึกษา เป็นต้น ซึ่งเป็นการฝ่าฝืนข้อตกลงในเรื่องการสุ่มผู้สอบ อันเป็นผลให้ค่าพารามิเตอร์ที่ได้

มีอคติ แม้ว่าจะสามารถลดค่าอคติลงได้โดยการเพิ่มจำนวนผู้สอบของแต่ละแบบสอบชุดย่อย แต่กระนั้นก็ยังส่อให้เห็นข้อจำกัดของวิธีการสุ่มตัวอย่างแบบเมตริกพหุคูณ (Shoemaker 1973:9-10)

5. การคำนวณยุ่งยากต้องใช้เครื่องคอมพิวเตอร์ช่วยในการคำนวณ

เนื่องจากการคำนวณหาค่าสถิติของแบบสอบชุดย่อยเพื่อจะนำไปใช้หาค่าประมาณพารามิเตอร์นั้นมีวิธีการคำนวณที่ยุ่งยาก ดังนั้นจึงต้องใช้เครื่องคอมพิวเตอร์ช่วยในการคำนวณ ซึ่งความจำเป็นในการที่จะต้องใช้เครื่องคอมพิวเตอร์นี้เองเป็นสาเหตุหนึ่งที่ทำให้เทคนิคนี้ไม่สามารถแพร่หลายเท่าที่ควร

6. ความยุ่งยากในการจัดการสอบ

เนื่องจากเทคนิคนี้ต้องใช้แบบสอบที่แตกต่างกันกับผู้สอบหลายกลุ่ม ฉะนั้นจึงมีความยุ่งยากในการจัดเตรียมแบบสอบ จัดกลุ่มผู้สอบ ตลอดจนการตรวจให้คะแนน ซึ่งแบบสอบชุดย่อยแต่ละชุดก็มีค่าเฉลี่ยแตกต่างกันไป

7. การไม่เน้นเรื่องการวัดผลแต่ละบุคคล มีผลต่อการทำแบบสอบอย่างเต็มความสามารถ

การที่ผู้สอบแต่ละคนไม่ต้องทำแบบสอบทั้งฉบับอาจมีผลดีในกรณีที่ทำให้ผู้สอบไม่รู้สึกเกรงกลัวการสอบ ทั้งนี้เพราะการที่แต่ละคนทำข้อสอบแต่เพียงบางข้อนั้น การนำผลสอบของแต่ละคนมาเปรียบเทียบกันย่อมเป็นไปไม่ได้ แต่ผลเสียที่เกิดขึ้นก็คือผู้สอบบางคนจะไม่พยายามทำข้อสอบจนสุดความสามารถของตน เพียงแต่ทำเท่าที่จะทำได้เท่านั้น ฉะนั้นคะแนนที่ได้ก็จะไม่ใช่คะแนนที่มาจากการทำแบบสอบอย่างเต็มความสามารถของผู้สอบ หรือกล่าวได้ว่าเป็นคะแนนที่แตกต่างจากเมื่อผู้สอบทำแบบสอบครบทั้งฉบับ ซึ่งผลของคะแนนในลักษณะนี้จะทำให้เกิดผลกระทบจากบริบทเช่นกัน (Popham 1973:11-13)

<u>การนำเทคนิคการสุ่มเมตริกพหุคูณไปใช้</u> (Shoemaker 1973:11-13)

เทคนิคการสุ่มเมตริกพหุคูณเป็นวิธีการที่สามารถนำไปใช้ได้อย่างกว้างขวางทั้งในงานด้านการศึกษาและงานด้านอื่น ๆ แต่เนื่องจากเทคนิคนี้ใช้กันมากในด้านการศึกษาโดยเฉพาะเรื่องการวัดผลสัมฤทธิ์จึงทำให้เทคนิคนี้ถูกจำกัดอยู่ในวงแคบ

หลักการวัตถุ-ตัวแปร (Object-Variable Rationale)

นักวิจัยมักใช้เมตริกในรูปผู้สอบ (แนวขวาง) - ข้อสอบ (แนวตั้ง) ซึ่งถ้าพิจารณาพื้นฐานของข้อมูลต่าง ๆ จะพบว่าเมตริกสามารถอยู่ในรูปวัตถุ (แนวขวาง) - ตัวแปร (แนวตั้ง) ได้ ซึ่งเมตริกผู้สอบ - ข้อสอบก็เป็นรูปแบบพิเศษแบบหนึ่งของเมตริกวัตถุ - ตัวแปร นั่นเอง ฉะนั้นในกรณีผู้อยู่อาศัย (แนวขวาง) - แบบสำรวจ (แนวตั้ง) เครื่องจักร (แนวขวาง) - การตรวจสอบการควบคุมควัน (แนวตั้ง) ปลา (แนวขวาง) - อาการของโรค (แนวตั้ง) และหิน (แนวขวาง) - ลักษณะความแข็ง (แนวตั้ง) ก็เป็นเมตริกที่สามารถนำกระบวนการเทคนิคการสุ่มเมตริกพหุคูณไปใช้ได้เช่นกัน

ตัวอย่างการนำไปใช้เฉพาะด้าน

แม้ว่าเทคนิคการสุ่มเมตริกพหุคูณจะสามารถนำไปใช้ได้กว้างขวางในหลายสาขา แต่เป็นที่รู้จักกันดีในวงการศึกษาและการวิจัยการศึกษามากกว่าในสาขาวิชาอื่น ๆ ซึ่งในด้านการศึกษาสามารถนำเทคนิคการสุ่มเมตริกพหุคูณไปใช้ได้ในเรื่องต่าง ๆ ต่อไปนี้

1. การประเมินผลหลักสูตร (Curriculum Evaluation) ในระดับชาติ ระดับเขตการศึกษา หรือระดับท้องถิ่นสามารถนำการสุ่มเมตริกพหุคูณไปใช้ได้เมื่อการประเมินผลโปรแกรมการสอนต่าง ๆ เหล่านั้นเกี่ยวข้องกับประชากรข้อสอบ การใช้เทคนิคนี้จะสามารถเพิ่มเนื้อหาในแบบสอบมาตรฐานเพื่อให้ครอบคลุมสิ่งที่ต้องการวัดอย่างเพียงพอ กระบวนการนี้จะลดทั้งเวลาสอบของผู้สอบแต่ละคนและจำนวนผู้สอบที่ต้องใช้ในการทดสอบด้วย

2. การประเมินโปรแกรมการสอน (Evaluation of Competing Instructional Program) ในการใช้ประชากรข้อสอบเพื่อประเมินผลการสอนในสาขาวิชาต่าง ๆ ซึ่งแต่ละสาขาวิชาประกอบด้วยเนื้อหาวิชาที่สำคัญหลายประเภทรวมกันสามารถใช้การสุ่มเมตริกพหุคูณเพื่อประเมินผลสัมฤทธิ์ของนักเรียน โดยใช้ประชากรข้อสอบซึ่งประกอบด้วยเนื้อหาของแต่ละโปรแกรมย่อย ๆ ผลของค่าประมาณพารามิเตอร์สามารถใช้เป็นพื้นฐานสำหรับกำหนดความสัมพันธ์ระหว่างจุดแข็งและจุดอ่อนของแต่ละโปรแกรม

3. การทดสอบสมมุติฐาน (Hypothesis Testing) ค่าประมาณพารามิเตอร์ที่คำนวณจากข้อมูลที่รวบรวมโดยใช้เทคนิคการสุ่มเมตริกพหุคูณสามารถนำมาใช้กับกระบวนการทดสอบสมมุติฐานต่าง ๆ ที่ใช้กันโดยทั่วไปได้อย่างง่ายดาย ตัวอย่างเช่น คะแนนจากข้อกระทงที่ได้มาจากแบบสอบชุดย่อยแต่ละชุดสามารถใช้แผนการวิเคราะห์ความแปรปรวนร่วม

ในแบบสุ่มสมบูรณ์และแบบแฟคทอเรียล

4. แบบสอบถามและการสำรวจ (Questionnaires and Surveys) เทคนิคการสุ่มเมตริกพหุคูณสามารถนำไปใช้ในการรวบรวมข้อมูลและประมาณค่าพารามิเตอร์ของคะแนนจากแบบสอบถามและการสำรวจ ซึ่งเทคนิคนี้ได้นำมาใช้ในเรื่องนี้กันมาก

5. การวัดเจตคติ (Attitude Measurement) ในการวัดเจตคติสามารถนำเทคนิคการสุ่มเมตริกพหุคูณไปใช้ และเทคนิคการสุ่มแบบนี้สามารถใช้ได้ดีกับข้อกระทงแบบแพร์-คอมแพริสัน (Paired-comparison) และแบบลิคเคิท

6. แบบสอบมาตรฐาน (Standardized Tests) การสุ่มเมตริกพหุคูณสามารถนำไปใช้ในการประมาณค่าการแจกแจงคะแนนทั้งหมด (Normative Distribution) ของแบบสอบมาตรฐาน ปัญหาประการหนึ่งของการสร้างปกติวิสัย คือการรวบรวมข้อมูลจากกลุ่มตัวอย่าง เพราะกลุ่มตัวอย่างที่ดีจะต้องเป็นตัวแทน (Representative) ของประชากร แต่เนื่องจากความยาวของแบบสอบ ซึ่งทำให้ผู้สอบต้องใช้เวลามากในการทำแบบสอบอาจมีผลต่อการให้ความร่วมมือ ดังนั้นการใช้เทคนิคการสุ่มเมตริกพหุคูณจะทำให้เวลาในการทำแบบสอบของผู้สอบแต่ละคนลดน้อยลง

<u>ลำดับขั้นการใช้การสุ่มเมตริกพหุคูณ</u>

เมื่อใช้เทคนิคการสุ่มเมตริกพหุคูณในการรวบรวมข้อมูลควรกระทำตามลำดับต่อไปนี้

ขั้นที่ 1 สร้างหรือเลือกประชากรข้อสอบ (จำนวน M ข้อกระทง) ในกรณีที่เป็นแบบสอบวัดผลสัมฤทธิ์ควรรวบรวมข้อกระทงเหล่านี้โดยแยกเป็นชั้น (Strata) ตามระดับความยาก (Difficulty Level)

ขั้นที่ 2 กำหนดขอบเขตและข้อจำกัดของกระบวนการดำเนินการสอบ

ขั้นที่ 3 เลือกแผนการสุ่มที่เหมาะสม

ขั้นที่ 4 ดำเนินการสอบโดยใช้แบบสอบชุดย่อยกับผู้สอบด้วยความระมัดระวังอย่าให้มีความสับสนเกี่ยวกับแบบสอบชุดย่อยและผู้สอบกลุ่มย่อยแต่ละกลุ่ม นอกจากนี้ต้องพยายามทำให้ภายในกลุ่มย่อยของผู้สอบมีความเป็นเอกพันธ์ (Homogeneous)

ขั้นที่ 5 คำนวณค่าประมาณพารามิเตอร์ต่าง ๆ โดยใช้เครื่องคอมพิวเตอร์

การคำนวณในเรื่องการสุ่มเมตริกพหุคูณ

เมื่อใช้เทคนิคการสุ่มเมตริกพหุคูณในการรวบรวมข้อมูล การคำนวณค่าประมาณพารามิเตอร์ และความคลาดเคลื่อนมาตรฐานของค่าประมาณแต่ละตัว มี 2 ขั้นตอนต่อไปนี้

ขั้นที่ 1 การคำนวณค่าสถิติจากตัวอย่างเมตริก คำนวณหาค่าสถิติต่าง ๆ จากแบบสอบชุดย่อย (ตัวอย่างเมตริก) แต่ละชุด

ขั้นที่ 2 การประมาณค่าพารามิเตอร์ นำค่าสถิติจากแบบสอบชุดย่อยแต่ละชุดซึ่งเป็นอิสระต่อกันมาใช้ในการคำนวณหาค่าประมาณพารามิเตอร์ที่ดีที่สุดของพารามิเตอร์แต่ละตัว

ในการประมาณค่าพารามิเตอร์นั้น ค่าสถิติจากตัวอย่างเมตริกเพียง 1 ตัวอย่างก็สามารถใช้ในการประมาณค่าพารามิเตอร์ได้ แต่การใช้ค่าสถิติจากตัวอย่างเมตริกหลายตัวอย่างมาหาค่าเฉลี่ย แล้วนำค่าเฉลี่ยนั้นมาใช้ประมาณค่าพารามิเตอร์จะทำให้ได้ค่าประมาณ "รวม" (The "pooled" estimate of parameter) ใกล้เคียงกับค่าพารามิเตอร์ที่แท้จริงมากกว่า นอกจากนี้ความคลาดเคลื่อนมาตรฐานของค่าประมาณก็จะลดน้อยลงด้วย (Shoemaker 1973:26)

การคำนวณค่าสถิติจากตัวอย่างเมตริก

ก่อนที่จะคำนวณหาค่าสถิติต่าง ๆ ของแต่ละตัวอย่างเมตริก จะต้องคำนวณหาค่า E, E^2, I, I^2, X^2, A, B และ C ของแต่ละตัวอย่างเมตริก โดยกำหนดให้

n คือ จำนวนตัวอย่างผู้สอบที่สุ่มมาจากประชากรผู้สอบ N คน

m คือ จำนวนตัวอย่างข้อสอบที่สุ่มมาจากประชากรข้อสอบ M ข้อ

X_{ij} คือ คะแนนของคนที่ i ข้อที่ j

E_i คือ คะแนนรวมของคนที่ i

I_j คือ คะแนนรวมของข้อสอบข้อที่ j

ภาพที่ 6 การคำนวณในแต่ละตัวอย่างเมตริกซึ่งประกอบด้วยตัวอย่างผู้สอบ n คน และตัวอย่างข้อสอบ m ข้อ โดยสุ่มมาจากประชากรผู้สอบ N คน และประชากรข้อสอบ M ข้อ

ข้อสอบ

ผู้สอบ

$$
\begin{array}{c|cccccc|cc}
 & 1 & 2 & 3 & \cdots j \cdots & & m & & \\
\hline
1 & X_{11} & X_{12} & X_{13} & \cdots\cdots & & X_{1m} & E_1 & E_1^2 \\
2 & X_{21} & X_{22} & X_{23} & \cdots\cdots & & X_{2m} & E_2 & E_2^2 \\
3 & X_{31} & X_{32} & X_{33} & \cdots\cdots & & X_{3m} & E_3 & E_3^2 \\
\cdot & \cdot & \cdot & \cdot & \cdots\cdots & & \cdot & \cdot & \cdot \\
i & \cdot & \cdot & \cdot & \cdot\, \cdot X_{ij} \cdots & & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdots\cdots & & \cdot & \cdot & \cdot \\
n & X_{n1} & X_{n2} & X_{n3} & \cdots\cdots & & X_{nm} & E_n & E_n^2 \\
\hline
 & I_1 & I_2 & I_3 & \cdot & \cdot & I_m & & \\
 & I_1^2 & I_2^2 & I_3^2 & \cdot & \cdot & I_m^2 & & \\
\end{array}
$$

$$\Sigma E = E_1 + E_2 + E_3 + \cdots\cdots + E_n$$

$$\Sigma E^2 = E_1^2 + E_2^2 + E_3^2 + \cdots\cdots + E_n^2$$

$$\Sigma I = I_1 + I_2 + I_3 + \cdots\cdots + I_m$$

$$\Sigma I^2 = I_1^2 + I_2^2 + I_3^2 + \cdots\cdots + I_m^2$$

$$\Sigma x^2 = x_{11}^2 + x_{12}^2 + x_{13}^2 + \cdots\cdots + x_{nm}^2$$

หมายเหตุ : ΣE ต้องเท่ากับ ΣI

หาค่า A, B และ C โดยใช้สูตร (Sirotnik 1974:480)

$$A = \left[\sum E^2/m - (\sum E)^2/nm\right] /(n-1)$$

$$B = \left[\sum I^2/n - (\sum I)^2/nm\right] /(m-1)$$

$$C = \left[\sum X^2 - \sum E^2/m - \sum I^2/n + (\sum E)^2/nm\right] /(n-1)(m-1)$$

นำค่าที่คำนวณได้ดังกล่าวข้างต้นมาคำนวณหาค่าสถิติต่าง ๆ ดังต่อไปนี้

1. มัชฌิมเลขคณิต ($\hat{\mu}$) ของผู้สอบ (หรือข้อสอบ) ในแต่ละตัวอย่างเมตริก (Sirotnik 1974:481)

$$\hat{\mu} = \frac{\sum E}{mn} \text{ หรือ } = \frac{\sum I}{mn} \quad \ldots\ldots\ldots\ldots\ldots\ldots (1)$$

2. ความแปรปรวนของคะแนนของผู้สอบ ($\hat{\sigma}_E^2$) ในแต่ละตัวอย่างเมตริก (Sirotnik 1974:481)

$$\hat{\sigma}_E^2 = \frac{N-1}{N}\left[\frac{A - (1-m/M)C}{m}\right] \ldots\ldots\ldots\ldots\ldots\ldots (2)$$

3. ความแปรปรวนของคะแนนของข้อสอบ ($\hat{\sigma}_I^2$) ในแต่ละตัวอย่างเมตริก (Sirotnik 1974:482)

$$\hat{\sigma}_I^2 = \frac{M-1}{M}\left[\frac{B - (1-n/N)C}{n}\right] \ldots\ldots\ldots\ldots\ldots\ldots (3)$$

4. ความแปรปรวนของปฏิสัมพันธ์ (Interaction Variance = $\hat{\sigma}_{EI}^2$) ในแต่ละตัวอย่างเมตริก (Sirotnik 1974:487)

$$\hat{\sigma}_{EI}^2 = \frac{(N-1)(M-1)}{NM} C \ldots\ldots\ldots\ldots\ldots\ldots (4)$$

สำหรับตัวอย่างเมตริก n x m เมื่อคะแนนของคนที่ i ข้อที่ j คือ X_{ij} ถ้าพิจารณาในเรื่องการวิเคราะห์ความแปรปรวน ตัวอย่างผู้สอบ n คน เป็นตัวอย่างที่สุ่มมาจากประชากรผู้สอบ N คน และตัวอย่างข้อสอบ m ข้อ เป็นตัวอย่างที่สุ่มมาจาก

ประชากรข้อสอบ m ข้อ จะเขียนสมการเชิงเส้นตรงได้ดังนี้

$$X_{ij} = \mu + \epsilon_i + \iota_j + \epsilon\iota_{ij}$$

เมื่อ μ คือ มัชฌิมเลขคณิตของประชากร (เมื่อผู้สอบทุกคนทำแบบสอบทั้งฉบับ)

ϵ_i คือ อิทธิพลสุ่ม (random effect) สำหรับผู้สอบคนที่ i (เป็นคะแนนที่เกี่ยวข้องกับประชากรข้อสอบ) $\epsilon_i \sim IIN\ (0, \sigma_\epsilon^2)$

ι_j คือ อิทธิพลสุ่มสำหรับข้อสอบที่ j (เป็นคะแนนที่เกี่ยวข้องกับประชากรผู้สอบ) $\iota_j \sim IIN\ (0, \sigma_\iota^2)$

$\epsilon\iota_{ij}$ คือ อิทธิพลสุ่มสำหรับปฏิสัมพันธ์ระหว่างผู้สอบคนที่ i และข้อสอบที่ j ภายในประชากร $\epsilon\iota_{ij} \sim IIN\ (0, \sigma_{\epsilon\iota}^2\ )$

อิทธิพลสุ่มแต่ละตัวตรงกับองค์ประกอบของความแปรปรวนที่ใช้สัญลักษณ์ σ_ϵ^2 σ_ι^2 และ $\sigma_{\epsilon\iota}^2$ ตามลำดับ (สำหรับสูตรที่ใช้ในการวิจัยนี้จะใช้สัญลักษณ์ σ_E^2 σ_I^2 และ σ_{EI}^2 แทนสัญลักษณ์ดังกล่าวตามลำดับ)

รูปแบบดังที่ระบุนี้จะมีข้อตกลงเบื้องต้นที่สำคัญก็คือ ϵ_i และ ι_j จะมีลักษณะที่เป็นตัวแปรสุ่ม โดยสุ่มอย่างเป็นอิสระต่อกัน มีค่าความแปรปรวนเท่ากัน และมีลักษณะเป็นการแจกแจงปกติ (Normal Distribution)

ตารางที่ 2 แสดงค่าที่คาดหวังของกำลังสองเฉลี่ย (The Expected Mean Squares - E[MS]) เมื่อใช้การสุ่มเมตริกพหุคูณที่สุ่มเพียง 1 ตัวอย่างเมตริก (หรือที่เรียกกันโดยทั่วไปว่า การสุ่มเมตริก) สำหรับสูตรที่ (1) ถึง (4) เป็นการประมาณค่าที่สอดคล้องกับองค์ประกอบของความแปรปรวนของแผนการนี้ ดังนั้น ค่ามัชฌิมเลขคณิตของแต่ละตัวอย่างเมตริกจึงเป็นค่าที่สามารถนำไปใช้ในการประมาณค่าพารามิเตอร์ได้ดีเมื่อคำนวณโดยใช้สูตร

(1) $\hat{\mu} = \Sigma E/nm$

ตารางที่ 2 แหล่งวิเคราะห์ความแปรปรวน ชั้นของความเป็นอิสระ (df) กำลังสองเฉลี่ย (MS) และค่าที่คาดหวังของกำลังสองเฉลี่ย (E $[MS]$)

แหล่ง	df	MS	E $[MS]$
ผู้สอบ (E)	n-1	MS_E	$(1-m/M)\sigma_{\epsilon\iota}^{2} + m\sigma_{\epsilon}^{2}$
ข้อกระทง (I)	m-1	MS_I	$(1-n/N)\sigma_{\epsilon\iota}^{2} + n\sigma_{\iota}^{2}$
EI	(n-1)(m-1)	MS_{EI}	$\sigma_{\epsilon\iota}^{2}$

หมายเหตุ : ในที่นี้ MS_E, MS_I และ MS_{EI} คือ A, B และ C ตามลำดับ

เพื่อที่จะแปลงค่าประมาณความแปรปรวนส่วนต่าง ๆ ให้อยู่ในรูปของ MS ฉะนั้นจึงกำหนดให้ σ_{ϵ}^{2}, σ_{ι}^{2} และ $\sigma_{\epsilon\iota}^{2}$ เท่ากับ $\sum\epsilon_{i}^{2}/N-1$, $\sum\iota_{j}^{2}/M-1$ และ $\sum\sum\epsilon\iota_{ij}^{2}/(N-1)(M-1)$ โดยเหตุที่ต้องการประมาณค่าความแปรปรวนที่แท้จริง ซึ่งความแปรปรวนส่วนต่าง ๆ โดยปกติคือ $\sum\epsilon_{i}^{2}/N$, $\sum\iota_{j}^{2}/M$ และ$\sum\sum\epsilon_{ij}^{2}/NM$ จึงต้องคูณด้วย "ค่าแก้" (Correction Factors) ต่อไปนี้ (N-1)/N, (M-1)/M และ (N-1)(M-1)/NM ตามลำดับ และให้ใช้สัญลักษณ์ของของความแปรปรวนส่วนต่าง ๆ เหมือนเดิมตามลำดับ สำหรับค่าประมาณความแปรปรวนส่วนต่าง ๆ ซึ่งแปลงค่าให้อยู่ในรูปของ MS ซึ่งสอดคล้องกับ E $[MS]$ ก็จะได้สมการต่าง ๆ ดังต่อไปนี้

$$(2)\quad \gamma_{\epsilon}^{2} = \frac{N-1}{N}\left[\frac{MS_E - (1-m/M)MS_{EI}}{m}\right]$$

$$(3)\quad \gamma_{\iota}^{2} = \frac{M-1}{M}\left[\frac{MS_I - (1-n/N)MS_{EI}}{n}\right]$$

$$(4)\quad \gamma_{\epsilon\iota}^{2} = \frac{(N-1)(M-1)}{NM}(MS_{EI})$$

<u>การประมาณค่าพารามิเตอร์</u>

1. การประมาณค่าพารามิเตอร์โดยใช้ค่าสถิติจากตัวอย่างเมตริกเพียง 1 ตัวอย่าง มีสูตรที่ใช้ในการคำนวณดังต่อไปนี้ (Sirotnik 1974:482)

ก. ค่าประมาณมัชฌิมเลขคณิตของคะแนนของผู้สอบทั้งหมด

Estimated mean $= M\hat{\mu}$ .... (5)

ข. ค่าประมาณความแปรปรวนของคะแนนของผู้สอบทั้งหมด

Estimated variance for examinee total scores $= M^2\hat{\sigma}^2_E$ ...(6)

ค. ค่าประมาณความแปรปรวนของคะแนนของข้อสอบทั้งหมด

Estimated variance for item total scores $= N^2\hat{\sigma}^2_I$ ..(7)

สำหรับค่าประมาณส่วนเบี่ยงเบนมาตรฐานใช้สูตรต่อไปนี้

Estimated standard deviation for examinee total scores $= M\hat{\sigma}_E$ ... (8)

Estimated standard deviation for item total scores $= N\hat{\sigma}_I$ ....(9)

2. การประมาณค่าพารามิเตอร์โดยใช้ค่าสถิติจากตัวอย่างเมตริกหลาย ๆ ตัวอย่างที่อยู่ในแผนการสุ่มเดียวกัน (การสุ่มเมตริกพหุคูณ) มีขั้นตอนดังต่อไปนี้

2.1 นำค่าสถิติจากตัวอย่างเมตริกในแผนการสุ่มเดียวกันมาหาค่าเฉลี่ย ในกรณีที่ขนาดตัวอย่างเมตริกเท่ากันสามารถหาค่าเฉลี่ยได้เลย แต่ถ้าขนาดตัวอย่างเมตริกไม่เท่ากันต้องมีการถ่วงน้ำหนัก (Sirotnik 1974:485)

ก. ค่าเฉลี่ยของมัชฌิมเลขคณิต ($\bar{\hat{\mu}}$)

ในกรณีที่ขนาดตัวอย่างเมตริกเท่ากัน (equal-sized)

$$\bar{\hat{\mu}} = [\hat{\mu}_1 + \hat{\mu}_2 + \cdots\cdots + \hat{\mu}_k]/k \quad \ldots\ldots\ldots\ldots\ldots(10)$$

ในกรณีที่ขนาดตัวอย่างเมตริกไม่เท่ากัน

$$\bar{\hat{\mu}} = [m_1n_1\hat{\mu}_1+m_2n_2\hat{\mu}_2+ \ldots + m_kn_k\hat{\mu}_k]/(m_1n_1+m_2n_2+ \ldots + m_kn_k) \quad (11)$$

ข. ค่าเฉลี่ยของความแปรปรวน ($\bar{\hat{\sigma}}^2_E$)

ในกรณีที่ขนาดตัวอย่างเมตริกเท่ากัน

$$\bar{\hat{\sigma}}^2_E = [\hat{\sigma}^2_1 + \hat{\sigma}^2_2 + \ldots + \hat{\sigma}^2_k]/k \ldots\ldots (12)$$

ในกรณีที่ขนาดตัวอย่างเมตริกไม่เท่ากัน

$$\bar{\hat{\sigma}}^2_E = [m_1n_1\hat{\sigma}^2_1+m_2n_2\hat{\sigma}^2_2+\ldots+m_kn_k\hat{\sigma}^2_k]/(m_1n_1+m_2n_2+\ldots+m_kn_k) \ldots\ldots (13)$$

2.2 นำค่าเฉลี่ยมาคำนวณหาค่าประมาณพารามิเตอร์ด้วยสูตรต่อไปนี้

ก. ค่าประมาณมัชฌิมเลขคณิตของคะแนนของผู้สอบทั้งหมด

$$\text{Estimated mean} = M\bar{\hat{\mu}} \ldots\ldots (14)$$

ข. ค่าประมาณความแปรปรวนของคะแนนของผู้สอบทั้งหมด

$$\text{Estimated variance for examinee total scores} = M^2\bar{\hat{\sigma}}^2_E \ldots\ldots (15)$$

ค. ค่าประมาณความแปรปรวนของคะแนนของข้อสอบทั้งหมด

$$\text{Estimated variance for item total scores} = N^2\bar{\hat{\sigma}}^2_I \ldots\ldots (16)$$

<u>ความคลาดเคลื่อนมาตรฐาน</u>

ในกรณีที่การสุ่มเมตริกพหุคูณที่ใช้เป็นการสุ่มแบบไม่คืนที่และสุ่มแบบครบหมดทั้งผู้สอบและข้อสอบ การคำนวณหาค่าความคลาดเคลื่อนมาตรฐานของค่ามัชฌิมเลขคณิตใช้สูตร (Sirotnik 1974:488)

$$\hat{\sigma}_{\hat{\mu}} = \sqrt{\frac{k-1}{(N-1)(M-1)}\bar{\hat{\sigma}}^2_{EI}} \ldots\ldots (17)$$

หมายเหตุ : ในกรณีที่เป็นการสุ่มเมตริกพหุคูณค่า $\hat{\sigma}^2_{EI}$ ต้องเป็นค่าเฉลี่ยเสมอ

ในกรณีที่เป็นการสุ่มแบบไม่ครบจะต้องใช้สูตรซึ่งประกอบด้วยค่าประมาณความแปรปรวนทั้ง 3 ส่วน โดยมีข้อแม้ว่าขนาดของตัวอย่างเมตริกจะต้องเท่ากัน (Sirotnik 1974:488)

$$\hat{\sigma}_{\hat{\mu}} = \sqrt{\frac{1}{knm(N-1)(M-1)}\left[m(M-1)(N-nk)\hat{\sigma}_E^2 + n(N-1)(M-mk)\hat{\sigma}_I^2 + \left[(N-n)(M-m)+nm(k-1)\right]\hat{\sigma}_{E\,I}^2\right]} \quad \ldots.(18)$$

สำหรับในกรณีที่นอกเหนือไปจากนี้สามารถนำวิธี "แจคไนฟ์" (Jackknife Procedure) มาใช้เพื่อคำนวณหาค่าความคลาดเคลื่อนมาตรฐาน ซึ่งวิธีการนี้สามารถใช้หาค่าความคลาดเคลื่อนมาตรฐานของค่าประมาณ "รวม" ของพารามิเตอร์ ซึ่งได้แก่ มัชฌิมเลขคณิต ความแปรปรวน รวมทั้งค่าโมเมนต์ที่ 3 และที่ 4 ด้วย นอกจากนี้วิธีแจคไนฟ์ยังใช้ได้กับแผนการสุ่มทุกแบบ (ตัวอย่างเช่น สุ่มแบบคืนที่หรือแบบไม่คืนที่ สุ่มแบบครบหมดหรือไม่ครบขนาดของตัวอย่างเมตริกเท่ากันหรือไม่เท่ากัน) จึงนับว่าวิธีการนี้สามารถใช้หาค่าความคลาดเคลื่อนมาตรฐานของค่าประมาณในเรื่องการสุ่มเมตริกพหุคูณได้อย่างกว้างขวาง (Shoemaker 1973:42-43) วิธีแจคไนฟ์มีวิธีการคำนวณดังต่อไปนี้

เมื่อ k คือ จำนวนตัวอย่างเมตริก

Y_{all} คือ ค่าเฉลี่ยของค่าสถิติที่คำนวณมาจากตัวอย่างเมตริกทั้งหมด

Y_s คือ ค่าสถิติของตัวอย่างเมตริกที่ s

$$Y_{(s)} = \frac{\sum_{s=1}^{k} O_s Y_s - O_s Y_s}{\left(\sum_{s=1}^{k} O_s\right) - O_s} \quad \text{เมื่อ} \quad O_s = m_s n_s$$

Y_{*s} คือ ค่าเทียม (Pseudovalue) ซึ่งคำนวณโดยใช้สูตร

$$Y_{*s} = kY_{all} - (k-1)Y_{(s)} \quad \text{เมื่อ } s = 1, 2, \ldots k$$

หาค่าเฉลี่ยของค่าเทียมโดยใช้สูตร

$$Y_{*} = (Y_{*1} + Y_{*2} + \ldots. + Y_{*k})/k$$

หาค่าความคลาดเคลื่อนมาตรฐานของค่าประมาณโดยใช้สูตร

$$\sigma^2_{Y_{*}} = \frac{\sum_{s=1}^{k}(Y_{*s} - Y_{*})^2}{k(k-1)} \quad \ldots\ldots\ldots\ldots\ldots\ldots(19)$$

<u>การคำนวณค่าสัมประสิทธิ์แอลฟาของครอนบาค</u>

การหาค่าความเที่ยงของแบบสอบเมื่อใช้ข้อมูลที่รวบรวมโดยใช้เทคนิคการสุ่มเมตริกพหุคูณ สามารถคำนวณโดยวิธีหาค่าสัมประสิทธิ์แอลฟาของครอนบาค (α) จากสูตรต่อไปนี้ (Sirotnik 1974:493)

$$\alpha = \frac{(M-1)\sigma^2_{E} - \sigma^2_{EI}}{(M-1)\sigma^2_{E}} \quad \ldots\ldots\ldots\ldots\ldots(20)$$

หมายเหตุ : ในกรณีที่เป็นการสุ่มเมตริกพหุคูณ $\hat{\sigma}^2_{E}$ และ $\hat{\sigma}^2_{EI}$ ต้องเป็นค่าเฉลี่ย

ถ้านำวิธีวิเคราะห์ความแปรปรวนที่ฮอยท์ (Hoyt) ได้นำมาใช้ในการหาค่าความเที่ยงมาใช้ ค่าสัมประสิทธิ์แอลฟาของแบบสอบซึ่งประกอบด้วยประชากรข้อสอบ M ข้อ จะมีสูตรดังนี้ (Sirotnik 1974:523)

$$\alpha = (MS_{E} - MS_{EI})/MS_{E}$$

สำหรับประชากรเมตริกขนาด N x M ซึ่งกำหนดให้

$$\sigma^2_{\epsilon} = \frac{N-1}{NM} MS_{E} \quad \text{และ} \quad \sigma^2_{\epsilon\iota} = \frac{(N-1)(M-1)}{NM} MS_{EI}$$

ดังนั้น

$$\alpha = \frac{NM\sigma^2_{\epsilon}/(N-1) - NM\sigma^2_{\epsilon\iota}/(N-1)(M-1)}{NM\sigma^2_{\epsilon}/(N-1)}$$

$$= \frac{(M-1)\sigma^2_{\epsilon} - \sigma^2_{\epsilon\iota}}{(M-1)\sigma^2_{\epsilon}}$$

ตอนที่ 3 งานวิจัยที่เกี่ยวข้อง

ลอร์ด (1962:259-267, quoted in Cook and Stufflebeam 1967:601-602; Plumlee 1964:623-624) ได้เสนอทั้งทฤษฎีและผลการวิจัยเพื่อสนับสนุนหลักการที่ว่า "ปกติวิสัยที่สร้างขึ้นโดยใช้คนจำนวนมากตอบคำถามเพียง 2-3 ข้อ จะใกล้เคียงกับปกติวิสัยที่สร้างโดยใช้กลุ่มตัวอย่างผู้สอบขนาดใหญ่ทำแบบสอบทั้งฉบับ มากกว่าปกติวิสัยที่สร้างโดยใช้คนจำนวนน้อยทำแบบสอบทั้งฉบับ" ลอร์ดได้เสนอผลงานวิจัยเพื่อสนับสนุนหลักการนี้โดยนำผลที่ได้จากการสุ่มเมตริก (ในยุคนั้นเรียกว่าการสุ่มข้อสอบ) และการสุ่มโดยทั่วไปซึ่งสุ่มเฉพาะผู้สอบ หรือเรียกว่า การสุ่มผู้สอบ เปรียบเทียบกับค่าสถิติของปกติวิสัย (The norm statistics) ที่ได้มาจากการใช้แบบสอบคัดเลือกบุคลากรอุตสาหกรรม (Industrial Personnel Testing Situation) ทั้งฉบับ ซึ่งประกอบด้วยประชากรข้อสอบจำนวน 70 ข้อ กับประชากรผู้สอบจำนวน 1,000 คน ซึ่งเป็นวิธีการที่เรียกว่าโพสมอร์เทม (Post-mortem) วิธีนี้จะทำให้ทราบค่าพารามิเตอร์ทั้งหมด

ในการสุ่มเมตริกนั้น แผนการสุ่มที่ลอร์ดใช้คือสุ่มตัวอย่างเมตริก 10 กลุ่ม แต่ละกลุ่มประกอบด้วยตัวอย่างข้อสอบ 7 ข้อ และตัวอย่างผู้สอบ 100 คน เป็นการสุ่มเมตริกพหุคูณแบบไม่คาบเกี่ยวกัน เนื่องจากสุ่มแบบไม่คืนที่ทั้งการสุ่มข้อสอบและผู้สอบ สำหรับการสุ่มผู้สอบ ซึ่งให้ตัวอย่างผู้สอบที่ได้รับการสุ่มมาทุกคนทำแบบสอบทั้งฉบับ คือทำทั้ง 70 ข้อ ลอร์ดได้สุ่มตัวอย่างเป็นจำนวน 10 กลุ่ม แต่ละกลุ่มมีตัวอย่างผู้สอบ 100 คน จากนั้นนำค่าสถิติที่คำนวณได้จากเมื่อใช้การสุ่มผู้สอบ 10 กลุ่ม และค่าประมาณที่คำนวณได้จากการใช้การสุ่มเมตริกมาเปรียบเทียบกับค่าพารามิเตอร์ พบว่าค่ามัชฌิมเลขคณิตจากการใช้การสุ่มผู้สอบนั้น มีเพียง 3 ใน 10 กลุ่มที่มีค่าใกล้เคียงกับมัชฌิมเลขคณิตของประชากรมากกว่าค่ามัชฌิมเลขคณิตที่คำนวณโดยวิธีประมาณค่าเมื่อใช้การสุ่มเมตริก ส่วนความแปรปรวนนั้น จากกลุ่มตัวอย่างที่ใช้การสุ่มผู้สอบจำนวน 10 กลุ่ม มี 5 กลุ่มที่ความแปรปรวนมีค่าใกล้เคียงกับความแปรปรวนที่คำนวณจากประชากรมากกว่าการใช้การสุ่มเมตริก และการแจกแจงคะแนนของแบบสอบที่สร้างขึ้นโดยใช้ค่าประมาณที่ได้มาจากการสุ่มเมตริกใกล้เคียงกับการแจกแจงคะแนนของปกติวิสัยที่สร้างขึ้นโดยใช้ข้อมูลจากประชากร

จากผลการวิจัยลอร์ดได้สรุปว่า การสร้างปกติวิสัยของแบบสอบที่ใช้ผู้สอบ 1,000 คน ทำแบบสอบที่มีข้อสอบ 70 ข้อนั้น การประมาณค่าปกติวิสัยโดยให้ผู้สอบ 1,000 คน ทำข้อสอบคนละ 7 ข้อ จะดีกว่าการใช้ผู้สอบ 100 คน ทำข้อสอบคนละ 70 ข้อ และในการอภิปรายผลลอร์ดได้แสดงความต้องการที่จะให้มีการตรวจสอบผลสรุปดังกล่าวนี้โดยการใช้ข้อมูลกลุ่มอื่น ๆ อีก

พลัมลี (Plumlee 1964:623-630) ได้ตรวจสอบการวิจัยของลอร์ดโดยทดลองใช้การสุ่มเมตริกกับประชากรที่มีขนาดค่อนข้างเล็ก กลุ่มประชากรที่ใช้ในการวิจัยนี้เป็นผู้สมัครเป็นเสมียนจำนวน 200 คน แบบสอบที่ใช้เป็นแบบสอบคัดเลือกบุคลากรอุตสาหกรรม (Industrial Personnel Situation) เกี่ยวกับการเว้นวรรคตอนจำนวน 30 ข้อ วิธีวิจัยใช้วิธีโพสมอร์เทมเช่นเดียวกับการวิจัยของลอร์ดโดยให้ผู้สอบทั้ง 200 คน ทำแบบสอบทั้งฉบับ และใช้การสุ่มเมตริก ซึ่งการวิจัยนี้เป็นการสุ่มเมตริกพหุคูณแบบไม่ควบเกี่ยวกัน เนื่องจากสุ่มแบบไม่คืนที่ทั้งข้อสอบและผู้สอบ โดยจะสุ่มตัวอย่างเมตริกจำนวน 10 กลุ่ม แต่ละตัวอย่างเมตริกประกอบด้วยตัวอย่างข้อสอบ 3 ข้อ และตัวอย่างผู้สอบจำนวน 20 คน ส่วนการสุ่มผู้สอบนั้น พลัมลีได้สุ่มตัวอย่างผู้สอบจำนวน 10 กลุ่ม จากนั้นนำค่ามัชฌิมเลขคณิต และส่วนเบี่ยงเบนมาตรฐานที่ได้มาจากการใช้เทคนิคการสุ่มเมตริก และการสุ่มผู้สอบมาเปรียบเทียบกับมัชฌิมเลขคณิต และส่วนเบี่ยงเบนมาตรฐานที่คำนวณจากประชากร พบว่า มัชฌิมเลขคณิตจากการสุ่มผู้สอบทั้ง 10 กลุ่มนั้นมีเพียง 2 กลุ่มที่ใกล้เคียงกับมัชฌิมเลขคณิตของประชากรมากกว่ามัชฌิมเลขคณิตจากการสุ่มเมตริก แต่ส่วนเบี่ยงเบนมาตรฐานนั้น จากการสุ่มผู้สอบจำนวน 10 กลุ่ม มีถึง 9 กลุ่มที่ส่วนเบี่ยงเบนมาตรฐานมีค่าใกล้เคียงกับส่วนเบี่ยงเบนมาตรฐานที่คำนวณโดยใช้ประชากรมากกว่าส่วนเบี่ยงเบนมาตรฐานที่คำนวณโดยวิธีประมาณค่าเมื่อใช้การสุ่มเมตริก แสดงว่าในกรณีประชากรขนาดเล็ก การสร้างปกติวิสัยโดยใช้เทคนิคการสุ่มเมตริกเป็นไปตามปกติไม่ได้ ทั้งนี้เนื่องจากมีปัญหาเกี่ยวกับการประมาณค่าส่วนเบี่ยงเบนมาตรฐาน

พลัมลีได้ทดลองรวมกลุ่มตัวอย่างผู้สอบจาก 10 กลุ่มให้เป็น 5 กลุ่ม แต่ละกลุ่มมีผู้สอบ 40 คน และรวมให้เป็น 2 กลุ่ม แต่ละกลุ่มมีผู้สอบ 100 คน พบว่าในกรณีที่มีผู้สอบกลุ่มละ 40 คน จะมีค่ามัชฌิมเลขคณิตจากการใช้การสุ่มผู้สอบเพียง 1 กลุ่มที่ใกล้เคียงกับมัชฌิมเลขคณิตของประชากรมากกว่าที่ใช้การสุ่มเมตริก และในกรณีที่มีผู้สอบกลุ่มละ

100 คน ค่ามัชฌิมเลขคณิตที่ใช้การสุ่มเมตริกใกล้เคียงกับมัชฌิมเลขคณิตของประชากรมากกว่าที่ใช้การสุ่มผู้สอบทั้ง 2 กลุ่ม แต่สำหรับส่วนเบี่ยงเบนมาตรฐานจากกลุ่มตัวอย่างทั้งหมด ทั้งในกรณีที่ผู้สอบกลุ่มละ 40 คน และ 100 คน ส่วนเบี่ยงเบนมาตรฐานที่ใช้การสุ่มผู้สอบทุกกลุ่มมีค่าใกล้เคียงกับส่วนเบี่ยงเบนมาตรฐานของประชากรมากกว่าที่ใช้การสุ่มเมตริก

จากผลการทดลอง พลัมลีสรุปว่า การสุ่มเมตริกเป็นวิธีการที่ใช้ในการสร้างมาตราคะแนนมาตรฐาน (Standard score scale) สำหรับแบบสอบอาศัยความสามารถ (power test) ที่ยังไม่มีข้อสรุปที่แน่นอน

คุก และสตัฟเฟิลบีม (Cook and Stufflebeam 1967:601-610) ได้ขยายการวิจัยในเรื่องการสุ่มเมตริกสำหรับการประมาณค่าปกติวิสัยของแบบสอบโดยใช้ตัวอย่างข้อสอบและตัวอย่างผู้สอบขนาดต่าง ๆ ได้แก่กลุ่มตัวอย่างขนาด 10 %, 25 %, 33 % และ 50 % ของประชากร ซึ่งกลุ่มตัวอย่างขนาด 10 % นั้น ลอร์ดและพลัมลีได้ใช้ในการวิจัยมาแล้ว

การวิจัยนี้ใช้แบบสอบวัดผลสัมฤทธิ์เรื่องอนามัยในวิทยาลัย (College Hygiene) ซึ่งได้ปรับปรุงเพื่อใช้ในสถาบันเหล่าทหารบกแห่งสหรัฐอเมริกา (United States Armed Forces Institute) โดยหน่วยพัฒนาแบบสอบของมหาวิทยาลัยโอไฮโอ (The Test Development Center of the Ohio State University) ซึ่งแบบสอบนี้ประกอบด้วยข้อสอบแบบเลือกตอบ จำนวน 115 ข้อ ใช้ทดสอบกับนักเรียนจำนวน 1,239 คน และวิธีวิจัยใช้วิธีโพสมอร์เหมเช่นเดียวกัน ในการสุ่มเมตริกนั้น การสุ่มผู้สอบใช้การสุ่มตัวอย่างแบบมีระบบทั้งในกรณีที่ใช้ตัวอย่างผู้สอบ 10 %, 25 %, 33 % และ 50 % ส่วนการสุ่มข้อสอบสุ่มโดยใช้ตารางเลขสุ่ม สำหรับการสุ่มผู้สอบ ซึ่งสุ่มเฉพาะผู้สอบก็ใช้การสุ่มตัวอย่างแบบมีระบบเช่นเดียวกัน คำนวณหามัชฌิมเลขคณิต และความแปรปรวนที่ได้จากกลุ่มตัวอย่างที่ใช้การสุ่มผู้สอบจำนวน 19 กลุ่ม และจากการสุ่มเมตริกที่ใช้แผนการสุ่มต่างกัน 4 แบบ (จำนวนตัวอย่างเมตริกเท่ากับ 10, 4, 3 และ 2 กลุ่ม) นำมาเปรียบเทียบกับพารามิเตอร์ ซึ่งพบว่าค่ามัชฌิมเลขคณิตของประชากรเท่ากับ 66.94 ค่ามัชฌิมเลขคณิตจากกลุ่มตัวอย่างที่ใช้การสุ่มผู้สอบมีค่าระหว่าง 65.35-68.05 ส่วนจากการสุ่มเมตริกมีค่าระหว่าง 65.86-66.73 ค่าความแปรปรวนของประชากรเท่ากับ 313.64

ความแปรปรวนจากกลุ่มตัวอย่างที่ใช้การสุ่มผู้สอบมีค่าระหว่าง 290.02-357.97 และจากการสุ่มเมตริกมีค่าระหว่าง 274.96-327.13 ซึ่ง คุก และสตัฟเฟิลบีม สรุปว่า ค่าประมาณของมัชฌิมเลขคณิตและความแปรปรวนจากการใช้การสุ่มเมตริกมีค่าต่ำกว่า จากการใช้การสุ่มผู้สอบ

นอกจากนี้คุกและสตัฟเฟิลบีมได้ทดสอบภาวะทับกันสนิท (Test for Goodness of Fit) ระหว่างการแจกแจงคะแนนปกติวิสัย (The fitted norm score distribution) และการแจกแจงคะแนนของกลุ่มตัวอย่าง (The fitted sample score distribution) โดยใช้การทดสอบด้วยไคสแควร์ (Chi-square) พบว่าค่าประมาณจากการสุ่มเมตริก นั้นใช้ประมาณค่าการแจกแจงคะแนนปกติวิสัยได้เหนือกว่า (superior) การใช้ค่า ประมาณจากการสุ่มผู้สอบ ยกเว้นในกรณีที่กลุ่มตัวอย่างมีขนาด 33 % ซึ่งแสดงให้เห็นว่า ความสัมพันธ์ระหว่างขนาดของตัวอย่างข้อสอบกับค่าไคสแควร์นั้นมีไม่มากนัก และจากผล การวิจัย คุกและสตัฟเฟิลบีมได้สรุปว่า การวิจัยนี้สนับสนุนข้อสรุปของลอร์ดที่ว่า "การสุ่ม เมตริกนั้น ถ้าประสิทธิภาพไม่สูงกว่าการสุ่มผู้สอบก็มีประสิทธิภาพเท่าเทียมกัน"

โนแวค (Novak 1974:4000-A) ได้ทดลองใช้การสุ่มเมตริกพหุคูณ ในโรงเรียนระดับประถมศึกษา โดยใช้แบบสอบวัดผลสัมฤทธิ์ที่มีชื่อว่าแบบสอบทักษะ พื้นฐานไอโอวา (Iowa Tests of Basic Skills-ITBS) การวิจัยนี้มีจุดมุ่งหมาย ศึกษาเรื่องการสุ่มเมตริกพหุคูณในด้านต่อไปนี้

1. ผลการเปลี่ยนแปลงเนื้อหาอันเนื่องมาจากการสุ่มข้อสอบ ที่มีต่อค่าประมาณ ของตัวอย่างเมตริก

2. ผลของการเคยเห็นข้อสอบมาก่อนที่มีต่อค่าประมาณของตัวอย่างเมตริก

3. เปรียบเทียบประสิทธิภาพของวิธีการสุ่มเมตริกพหุคูณกับการสุ่มผู้สอบ

กลุ่มตัวอย่างประชากรที่ใช้ในการวิจัยนี้เป็นนักเรียนโรงเรียนระดับประถมศึกษา 2 โรงเรียน เป็นนักเรียนระดับ 4 จำนวน 124 คน นักเรียนระดับ 5 จำนวน 119 คน และนักเรียนระดับ 6 จำนวน 109 คน โดยใช้แบบสอบ ITBS แบบฟอร์มที่ 5 ซึ่ง ประกอบด้วยแบบสอบชุดย่อย 3 ชุด ได้แก่ คำศัพท์ การสะกดคำ และมโนภาพเกี่ยวกับ คณิตศาสตร์

โนแวคได้แบ่งแบบสอบออกเป็น 2 พวก โดยพวกที่หนึ่งได้จัดให้มีการสอบก่อน 1 สัปดาห์ จากนั้นจึงนำมาใช้สอบซ้ำอีกครั้งหนึ่ง พร้อมกับแบบสอบพวกที่ 2 ซึ่งวิธีดำเนินการสอบมีทั้งการใช้การสุ่มเมตริก และการสุ่มผู้สอบ ทำให้ได้ข้อมูลภายใต้ภาวะการณ์ 4 แบบ ดังต่อไปนี้

1. แบบสอบที่ประกอบด้วยตัวอย่างข้อสอบที่สุ่มมา และผู้สอบไม่เคยเห็นข้อสอบมาก่อน

2. แบบสอบที่มีข้อสอบครบทั้งฉบับ และผู้สอบเคยเห็นข้อสอบมาก่อน

3. แบบสอบที่มีข้อสอบครบทั้งฉบับ และผู้สอบไม่เคยเห็นข้อสอบมาก่อน

4. แบบสอบที่ประกอบด้วยตัวอย่างข้อสอบที่สุ่มมา และผู้สอบเคยเห็นข้อสอบมาก่อน

นำค่าประมาณจากแต่ละตัวอย่างเมตริกแยกตามแบบสอบชุดย่อยและระดับชั้นเรียน จำนวน 9 ค่า ซึ่งรวบรวมภายใต้ภาวะการณ์ 4 แบบดังกล่าว และค่าประมาณที่ได้จากข้อมูลที่รวบรวมโดยการสุ่มผู้สอบ จำนวน 10 ค่า ที่มีจำนวนคำตอบของผู้สอบทั้งหมดในแต่ละกลุ่มเท่ากับเมื่อใช้การสุ่มเมตริกมาวิเคราะห์ความแปรปรวนแบบพหุคูณ (Multivariate Analysis of Variance) โดยที่คะแนนจากแบบสอบชุดย่อยเกี่ยวกับคำศัพท์ การสะกดคำ และมโนภาพ เกี่ยวกับคณิตศาสตร์เป็นตัวแปรตาม และใช้การทดสอบค่าทีในการเปรียบเทียบข้อมูลทีละคู่ (Paired data <u>t-tests</u>) โดยการเปรียบเทียบค่าประมาณจากการสุ่มเมตริกพหุคูณกับค่าประมาณจากการสุ่มผู้สอบจากกลุ่มตัวอย่างที่มีความแม่นยำที่สุด และในการวิเคราะห์ใช้เฉพาะค่าประมาณจากการสุ่มเมตริกที่ต้องพบในการนำเทคนิคนี้ไปประยุกต์ใช้นั่นคือ แบบสอบที่ประกอบด้วยตัวอย่างข้อสอบที่สุ่มมาโดยการสุ่มเมตริก และผู้สอบไม่เคยเห็นข้อสอบมาก่อน จากผลการวิจัยพบว่า

1. ค่าประมาณมัชฌิมเลขคณิตจากการใช้การสุ่มเมตริกพหุคูณมีความแม่นยำมากกว่าค่าประมาณจากการใช้การสุ่มผู้สอบที่นำมาเปรียบเทียบ และค่าประมาณความแปรปรวนมีความแม่นยำเท่ากับค่าประมาณจากการใช้การสุ่มผู้สอบที่นำมาเปรียบเทียบ

2. การเปลี่ยนแปลงของเนื้อหาอันเนื่องมาจากการสุ่มข้อสอบเมื่อใช้การสุ่มเมตริกมีผลกระทบต่อค่าประมาณความแปรปรวนของตัวอย่างเมตริก แต่ไม่มีผลกระทบต่อค่าประมาณมัชฌิมเลขคณิต

3. การเคยเห็นข้อสอบมาก่อน มีผลกระทบต่อค่าประมาณมัชฌิมเลขคณิตของตัวอย่างเมตริก แต่ไม่มีผลกระทบต่อค่าประมาณความแปรปรวนของตัวอย่างเมตริก ค่าประมาณที่คำนวณจากคะแนนที่ได้จากการทำแบบสอบในครั้งที่สองมีความเที่ยงตรงน้อยกว่าค่าประมาณที่ได้จากการทำแบบสอบในครั้งแรก

จากผลการวิจัยนี้ โนแวคได้สรุปว่าวิธีการสุ่มเมตริกพหุคูณเป็นวิธีที่สามารถนำมาใช้รวบรวมข้อมูลในการทดสอบนักเรียนระดับประถมศึกษาโดยใช้แบบสอบวัดผลสัมฤทธิ์

ปีเตอร์เสน (Petersen 1968:135-149) เป็นคนแรกที่นำการสุ่มเมตริกมาใช้กับแบบมาตราเจตคติแบบลิคเคิต ซึ่งปีเตอร์เสนได้กล่าวถึงการสุ่มเมตริกซึ่งใช้กันในขอบเขตจำกัด คือใช้เฉพาะแบบสอบวัดผลสัมฤทธิ์ และข้อกระทงแบบไบนารี (Binary items-เมื่อตอบถูกได้ 1 คะแนน และเมื่อตอบผิดได้ 0 คะแนน) ซึ่งลอร์ดและโนวิคแนะนำให้ใช้ในกรณีที่

1. การใช้แบบสอบทั้งฉบับกับผู้สอบทุกคนเป็นสิ่งที่เป็นไปไม่ได้
2. ถ้าให้ผู้สอบทุกคนทำแบบสอบทั้งฉบับแล้ว จะมีอคติเกิดขึ้นเนื่องจากความไม่เต็มใจในการให้ความร่วมมือของผู้สอบ หรือผู้สอบทำแบบสอบไม่ครบ
3. เมื่อไม่มีแบบสอบที่ใช้สำหรับทดสอบก่อนการเรียนการสอนหรือการปฏิบัติ

เมื่อพิจารณาจากข้อแนะนำดังกล่าว ปีเตอร์เสนเห็นว่ามีเหตุผลเพียงพอที่จะนำเทคนิคการสุ่มเมตริกมาใช้กับการประเมินเจตคติต่อสภาพแวดล้อมในสถาบัน โดยการวิจัยของปีเตอร์เสนต้องการเปรียบเทียบผลจากการใช้การสุ่มเมตริก และการสุ่มผู้สอบเมื่อใช้แบบมาตราเจตคติแบบลิคเคิตเป็นเครื่องมือในการรวบรวมข้อมูล เพื่อวิเคราะห์ว่าการใช้เทคนิคการสุ่มที่แตกต่างกันจะเป็นสาเหตุที่ทำให้มีความแตกต่างระหว่างค่ามัชฌิมเลขคณิต และส่วนเบี่ยงเบนมาตรฐานของข้อกระทงหรือไม่

ปีเตอร์เสนใช้แบบมาตราเจตคติที่มีชื่อว่า แคมปัส เอนไวรอนเมนต์ สตัดี (Campus Environment Study-CES) ซึ่งมีข้อความจำนวน 150 ข้อ ประกอบด้วยองค์ประกอบเกี่ยวกับสภาพแวดล้อมในวิทยาลัย 6 ด้าน คือ (1) วิชาการ (2) อนามัย (3) วัฒนธรรม (4) การสื่อสาร (5) ความสัมพันธ์ในชุมชน และ (6) ศีลธรรม-จริยธรรม แต่ละด้านมีข้อความจำนวน 25 ข้อ มีคำตอบ 5 รายการ คือ เห็นด้วยอย่างยิ่ง เห็นด้วย ไม่แน่ใจ ไม่เห็นด้วย และไม่เห็นด้วยอย่างยิ่ง และจะให้คะแนน 5, 4, 3, 2

และ 1 ตามลำดับ เมื่อเป็นข้อความเชิงนิมานให้คะแนน 1, 2, 3, 4 และ 5 ตามลำดับ เมื่อเป็นข้อความเชิงนิเสธ

การรวบรวมข้อมูลเริ่มกระทำในเดือนพฤษภาคม ค.ศ.1968 โดยใช้แบบมาตราเจตคติ CES กับนักศึกษา 25 สถาบัน จำนวน 13,500 คน การสุ่มผู้สอบใช้ตัวอย่างผู้ตอบซึ่งเป็นนักศึกษาวิทยาลัยแห่งรัฐแมนคาโต (Mankato State College) จำนวน 1,300 คน และได้รับแบบมาตราเจตคติกลับคืน 791 ฉบับ สำหรับการสุ่มเมตริกนั้น ปีเตอร์เสนได้สร้างแบบมาตราเจตคติชุดย่อยโดยใช้เครื่องคอมพิวเตอร์ในการสุ่มข้อกระทง การสุ่มเป็นแบบไม่คืนที่ และแบบมาตราเจตคติชุดย่อย แต่ละฉบับประกอบด้วยข้อความ 6 ข้อ จากนั้นนำแบบมาตราเจตคติชุดย่อยเหล่านี้ไปใช้กับตัวอย่างผู้ตอบจำนวน 5,958 คน ในเดือนพฤศจิกายน ค.ศ.1968 และรวบรวมข้อมูลโดยใช้แบบมาตราเจตคติชุดย่อยเหล่านี้อีกในเดือนกุมภาพันธ์ ค.ศ.1969 โดยใช้ตัวอย่างผู้ตอบจำนวน 5,421 คน จากการใช้การสุ่มเมตริกรวบรวมข้อมูล 2 ครั้งนั้น ค่าเฉลี่ยของจำนวนผู้ตอบแต่ละข้อกระทงเท่ากับ 455 คน

จากข้อมูลที่รวบรวมได้ทั้งหมด ปีเตอร์เสนนำมาหาความแตกต่างของค่ามัชฌิมเลขคณิต และส่วนเบี่ยงเบนมาตรฐาน ที่คำนวณมาจากข้อมูลที่รวบรวมโดยใช้เทคนิคการสุ่มที่แตกต่างกัน ใช้วิธีคำนวณค่าสัมประสิทธิ์สหสัมพันธ์แบบโปรดัก-โมเมนต์ของเพียร์สัน (Pearson product-moment coefficients of correlation) แยกตามองค์ประกอบ 6 ด้าน ซึ่งแต่ละด้านประกอบด้วยข้อความจำนวน 25 ข้อ และทดสอบความมีนัยสำคัญของค่าสัมประสิทธิ์สหสัมพันธ์โดยใช้อัตราส่วนที (t-ratios) จากผลการวิจัยพบว่า ค่าสหสัมพันธ์ระหว่างมัชฌิมเลขคณิตจากการสุ่มผู้สอบ และการสุ่มเมตริกมีความคงเส้นคงวาสูง แม้ว่าการวิจัยนี้จะมีสาเหตุที่เป็นไปได้ที่ทำให้เกิดความผันแปรในค่ามัชฌิมเลขคณิตของข้อกระทง 4 ประการ ดังต่อไปนี้

1. การเปลี่ยนแปลงเจตคติของนักศึกษา เพราะระยะเวลาในการเก็บรวบรวมข้อมูลนี้ยาวมาก

2. ในกรณีที่ใช้การสุ่มผู้สอบนั้น ได้รับแบบมาตราเจตคติคืนมาเพียง 60%

3. ความคลาดเคลื่อนจากการสุ่ม (Sampling Error) ทั้งการสุ่มผู้สอบและการสุ่มเมตริก

4. ใช้เทคนิคการสุ่มเมตริกกับนักศึกษาปีแรก และใช้การสุ่มผู้สอบกับนักศึกษาปีสุดท้าย

สำหรับสหสัมพันธ์ระหว่างส่วนเบี่ยงเบนมาตรฐานจากการสุ่มผู้สอบและการสุ่มเมตริกมีความคงเส้นคงวาสูงเช่นกัน และพบว่าค่าเฉลี่ยของส่วนเบี่ยงเบนมาตรฐานของแต่ละข้อกระทงที่ใช้การสุ่มเมตริกมีค่าต่ำกว่าที่ใช้การสุ่มผู้สอบ ทั้งนี้อาจเป็นเพราะว่านักศึกษามีความระมัดระวังในการตอบข้อกระทงเพียง 6 ข้อมากกว่าพวกที่ต้องตอบข้อกระทงถึง 150 ข้อ หรือแสดงให้เห็นว่าการใช้การสุ่มเมตริกทำให้การตอบของผู้ตอบมีความแปรปรวน (Variability) น้อยกว่าเมื่อใช้การสุ่มผู้สอบ

จากผลการวิจัย ปีเตอร์เสนได้สรุปว่า อย่างน้อยผลจากการใช้การสุ่มเมตริกก็เชื่อมั่นได้เท่ากับการสุ่มผู้สอบ และได้เสนอแนะว่าควรมีการทำงานวิจัยในทำนองเดียวกันนี้อีก โดยใช้เทคนิคการสุ่มตัวอย่างที่ต่างกัน 2 วิธีในระยะเวลาเดียวกันเพื่อลดสาเหตุความผันแปรของค่ามัชฌิมเลขคณิตของข้อกระทงลงไป ตลอดจนศึกษาหาเทคนิคการสร้างแบบมาตราเจตคติชุดย่อยที่มีประสิทธิภาพอีกด้วย

พัช (Pugh 1971:54-56, quoted in Shoemaker 1973:81) ได้ศึกษาความสัมพันธ์ของการสุ่มเมตริกและการสุ่มผู้สอบเมื่อใช้มาตราเจตคติแบบลิคเคิต งานวิจัยนี้ใช้วิธีโพสมอร์เทม โดยรวบรวมข้อมูลทั้งหมดจากการให้นักเรียนระดับ 6 จำนวน 600 คน ตอบแบบมาตราเจตคติแบบลิคเคิตซึ่งมีความยาว 60 ข้อกระทง วิธีนี้ทำให้ทราบค่าพารามิเตอร์ จากนั้นพัชได้ทดลองใช้การสุ่มแบบเมตริกพหุคูณและการสุ่มผู้สอบหลายแบบ และนำข้อมูลที่ได้จากการสุ่มแบบต่าง ๆ เหล่านี้มาคำนวณหาค่าประมาณมัชฌิมเลขคณิตและความแปรปรวน จากนั้นนำค่าที่ได้มาเปรียบเทียบกับพารามิเตอร์ ซึ่งผลจากการวิจัยพัชได้สรุปว่าค่าประมาณที่คำนวณมาจากข้อมูลที่รวบรวมโดยใช้การสุ่มเมตริกพหุคูณมีแผนการสุ่มแบบต่าง ๆ ใกล้เคียงกับพารามิเตอร์ และพบว่าการสุ่มเมตริกให้ค่าประมาณพารามิเตอร์ที่มีความแม่นยำมากกว่าการสุ่มผู้สอบ

จากงานวิจัยที่เกี่ยวข้องดังกล่าวเป็นงานวิจัยในต่างประเทศทั้งสิ้น สำหรับในประเทศไทยยังไม่ปรากฏว่ามีงานวิจัยเกี่ยวกับเรื่องนี้ แต่จากงานวิจัยที่เกี่ยวข้องดังกล่าวมาทั้งหมดนั้นจะเห็นว่าส่วนใหญ่เป็นงานวิจัยที่ใช้วิธีโพสมอร์เทม ซึ่งแม้ว่าจะสามารถเปรียบเทียบค่าประมาณกับพารามิเตอร์ที่แท้จริงได้ แต่ก็ทำให้มองเห็นปัญหาในด้านการนำเทคนิค

นี้ไปใช้ในสภาพการณ์ที่แท้จริงว่าผลที่ได้จากการใช้เทคนิคการสุ่มเมตริกจะแตกต่างไปจากเมื่อใช้การสุ่มผู้สอบซึ่งเป็นเทคนิคที่ยอมรับกันโดยทั่วไปหรือไม่ แม้ว่าการวิจัยของปีเตอร์เสนจะได้ทดลองใช้การสุ่มเมตริกกับสภาพการณ์จริง แต่วิธีดำเนินการวิจัยก็ยังมีสาเหตุที่ทำให้เกิดความแปรปรวนได้หลายประการดังได้กล่าวมาแล้ว ดังนั้นผู้วิจัยจึงสนใจที่จะทำการศึกษาเกี่ยวกับการนำเทคนิคการสุ่มเมตริกมาใช้กับแบบมาตราเจตคติแบบลิคเคิตในสภาพการณ์จริงเพื่อจะได้ข้อค้นพบใหม่ ๆ ที่จะทำให้ความรู้เกี่ยวกับเรื่องการสุ่มเมตริกได้ขยายขอบเขตออกไป และเป็นแนวทางการนำเทคนิคนี้มาใช้ในประเทศไทยในโอกาสต่อไปอีกด้วย