การกำหนดขนาดตัวอย่างในการทดสอบสมมุติฐานวิจัย

นงลักษณ์ วิรัชชัย

(เอกสารประกอบการบรรยาย เรื่อง “วิธีการที่ถูกต้องและทันสมัยในการกำหนดขนาดตัวอย่าง” ในโครงการ Research Zone จัดโดย ศูนย์การเรียนรู้ทางการวิจัย ณ อาคารศูนย์การเรียนรู้ทางการวิจัย สำนักงานคณะกรรมการวิจัยแห่งชาติ (วช.) วันที่ 26 มกราคม 2555 เวลา 9.00-12.00 น.)

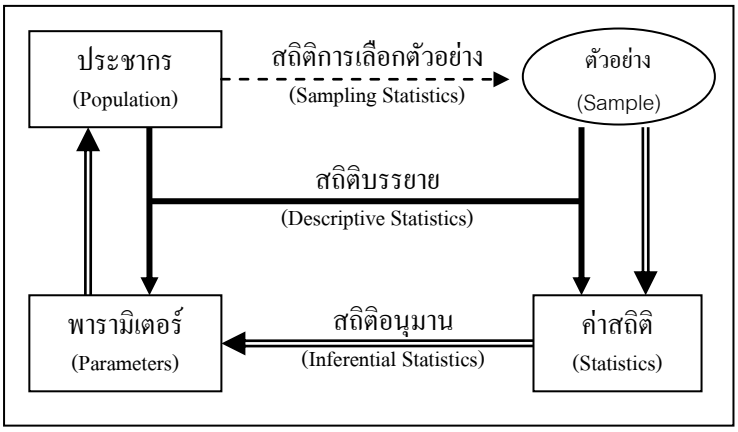
การวิจัยเชิงปริมาณที่นักวิจัยต้องใช้สถิติวิเคราะห์ในการทดสอบสมมุติฐานทางสถิติ และ/หรือการประมาณค่าพารามิเตอร์ นักวิจัยต้องมีความรู้ความสามารถในการกำหนดขนาดตัวอย่างสำหรับการวิจัย วิธีการกำหนดขนาดตัวอย่างที่ใช้กันอยู่ในวงการวิจัยของไทยในปัจจุบันมีหลายวิธี วิธีที่รู้จักและใช้กันอยู่ทั่วไป คือ การใช้ตารางสำเร็จของ Krejcie and Morgan (1970) Yamane (1970) และ Cohen (1977) โดยที่นักวิจัยส่วนใหญ่ใช้ตารางสำเร็จโดยไม่เข้าใจที่มาของวิธีการดังกล่าว ในขณะที่วงการวิจัยต่างประเทศได้เลือกวิธีการที่ดี และพัฒนาโปรแกรมคอมพิวเตอร์ขึ้นเพื่อช่วยให้นักวิจัยสามารถกำหนดขนาดตัวอย่างได้ถูกต้องและรวดเร็ว ผู้เขียนเห็นความสำคัญของการกำหนดตัวอย่างด้วยวิธีการที่ถูกต้องและทันสมัย ซึ่งจะช่วยให้การทำวิจัยมีประสิทธิผลดีมากขึ้น และได้ผลการวิจัยที่มีคุณภาพเป็นที่ยอมรับให้พิมพ์เผยแพร่ได้ในวารสารวิชาการระดับนานาชาติ จึงเขียนบทความนี้ขึ้น โดยมุ่งหวังให้นักวิจัยไทยได้ใช้วิธีการกำหนดขนาดตัวอย่างที่เหมาะสมและทันสมัยมากขึ้น สาระในบทความแยกเป็น 3 ตอน คือ ความรู้พื้นฐานเกี่ยวกับการกำหนดขนาดตัวอย่าง วิธีการกำหนดขนาดตัวอย่างจากตารางสำเร็จ และวิธีการกำหนดขนาดตัวอย่างโดยใช้โปรแกรมคอมพิวเตอร์ ดังรายละเอียดต่อไปนี้

ตอนที่ 1 ความรู้พื้นฐานเกี่ยวกับการกำหนดขนาดตัวอย่าง

การกำหนดขนาดตัวอย่างหรือการประมาณค่าขนาดตัวอย่าง (estimation of sample size) เป็นขั้นตอนหนึ่งของกระบวนการเลือกตัวอย่าง ซึ่งมี 6 ขั้นตอน คือ 1) การกำหนดวัตถุประสงค์ของการใช้สถิติวิเคราะห์และการเลือกใช้สถิติทดสอบ (test statistics) 2) การกำหนดลักษณะประชากรและตัวอย่าง 3) การกำหนดระดับนัยสำคัญ (significance level = α) และอำนาจการทดสอบ (power of the test = 1- β) 4) การกำหนดขนาดตัวอย่าง 5) การกำหนดวิธีการเลือกตัวอย่าง และ 6) การดำเนินการเลือกตัวอย่างตามแนวทางที่กำหนดไว้ (Howell, 2010; Kerlinger and Lee, 2000) จากขั้นตอนของกระบวนการเลือกตัวอย่าง จะเห็นได้ว่าการกำหนดขนาดตัวอย่างเกี่ยวข้องกับความรู้พื้นฐานรวม 3 เรื่อง คือ ก) ประชากร ตัวอย่าง และสถิติวิเคราะห์ ข) การแจกแจงของค่าสถิติจากตัวอย่าง และความคลาดเคลื่อนมาตรฐาน และ ค) การทดสอบสมมุติฐานและการประมาณค่าพารามิเตอร์ ซึ่งมีคำสำคัญ ได้แก่ ระดับนัยสำคัญ อำนาจการทดสอบ จุดวิกฤติ ความคลาดเคลื่อนในการทดสอบสมมุติฐาน อันเป็นความรู้ที่ควรต้องเข้าใจก่อนทำความเข้าใจเรื่องการกำหนดขนาดตัวอย่าง การเสนอสาระในตอนนี้จึงเป็นการนำเสนอความรู้พื้นฐานดังกล่าว (หากผู้อ่านมีความเข้าใจความรู้พื้นฐานแล้ว ท่านอาจข้ามไปอ่านสาระตอนที่ 2 ได้) โดยแยกการนำเสนอสาระเป็น 3 หัวข้อ ดังนี้

1. ประชากร ตัวอย่าง และสถิติวิเคราะห์ (Population, Sample Analytical Statistics)

ในการวิจัยเชิงปริมาณ นักวิจัยส่วนใหญ่ไม่สามารถศึกษากับประชากรที่มีขนาดใหญ่ได้ และใช้การศึกษากับตัวอย่าง ซึ่งเป็นเซ็ทย่อยของประชากร นักวิจัยใช้สถิติวิเคราะห์ประเภทการเลือกตัวอย่าง (sampling statistics) ในการเลือกตัวอย่างที่เป็นตัวแทนที่ดีของประชากรและมีขนาดพอเพียง ใช้สถิติบรรยาย (descriptive statistics) ในการนำเสนอข้อมูลในรูปแผนภูมิ และค่าสถิติ (statistics) สำหรับการบรรยายลักษณะตัวอย่าง หรือ ค่าพารามิเตอร์ (parameters) สำหรับบรรยายลักษณะประชากร และใช้สถิติอนุมาน (inferential statistics) ในการสรุปผลการวิเคราะห์ข้อมูลจากตัวอย่างซึ่งอยู่ในรูปค่าสถิติ และอ้างอิงไปสู่ค่าพารามิเตอร์ เพื่อสรุปให้ได้ข้อค้นพบสำหรับ ประชากร สถิติพาราเมตริกแบ่งตามระเบียบวิธีทางสถิติได้เป็น 2 แบบ คือ การประมาณค่าพารามิเตอร์ (parameter estimation) และการทดสอบสมมุติฐานทางสถิติ (hypothesis testing) เมื่อแบ่งตามลักษณะคำถามวิจัยได้เป็น 2 แบบ คือ สถิติวิเคราะห์สำหรับการศึกษาเปรียบเทียบ และสถิติวิเคราะห์สำหรับการศึกษาความสัมพันธ์ แต่ละแบบยังแบ่งตามระดับการวัดตัวแปรและลักษณะของตัวอย่างอีกหลายแบบ และแบ่งตามข้อกำหนดเกี่ยวกับค่าพารามิเตอร์ได้เป็น 2 แบบ คือ สถิติพาราเมตริก (parametric statistics) ที่มีข้อตกลงเบื้องต้นเกี่ยวกับลักษณะประชากร และสถิตินันพาราเมตริก (non-parametric statistics) ที่มีข้อตกลงเบื้องต้นเกี่ยวกับลักษณะประชากรน้อยหรือไม่มีเลย สำหรับความสัมพันธ์ระหว่างประชากร ตัวอย่าง และสถิติวิเคราะห์แสดงได้ดังภาพต่อไปนี้

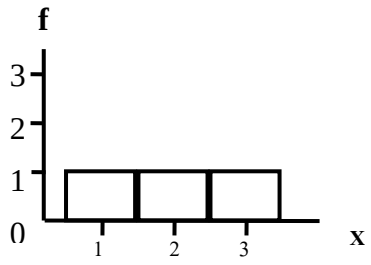


ภาพ 1 ความเกี่ยวข้องสัมพันธ์ระหว่างประชากร ตัวอย่าง และสถิติวิเคราะห์ 3 ประเภท

2. การแจกแจงของค่าสถิติจากตัวอย่าง (Sampling Distribution of Statistics or Distribution of Sample Statistics) และความคลาดเคลื่อนมาตรฐาน (Standard Error)

การแจกแจง (distribution) ที่ใช้ในสถิติวิเคราะห์ มี 3 แบบ คือ ก) การแจกแจงของประชากร (population distribution) ซึ่งเป็นผลการนำค่าของตัวแปรที่ได้จากสมาชิกทุกคนในประชากรมาแจกแจง ข) การแจกแจงของตัวอย่าง (sample distribution) ซึ่งเป็นผลการนำค่าของตัวแปรที่ได้จากสมาชิกทุกคนในตัวอย่างมาแจกแจง และ ค) การแจกแจงของค่าสถิติจากการเลือกตัวอย่าง หรือเรียกสั้นๆว่า **การแจกแจงของค่าสถิติจากตัวอย่าง** เป็นผลจากการนำค่าสถิติ ที่ได้จากตัวอย่างทั้งหมดทุกกลุ่มที่เป็นไปได้ (all possible samples) มาแจกแจง ในกรณีที่มีประชากรขนาด N หน่วย การเลือกตัวอย่าง ขนาด n หน่วย เลือกได้สอง

แบบ **แบบแรก คือ การเลือกตัวอย่างแบบใส่คืน (sampling with replacement)** ได้ตัวอย่างที่เป็นไปได้ทั้งหมด N^n กลุ่ม **แบบที่สอง คือ การเลือกตัวอย่างแบบไม่ใส่คืน (sampling without replacement)** ได้ตัวอย่างที่เป็นไปได้ทั้งหมดจำนวน ${}^{N}C_{n}$ กลุ่ม เมื่อคำนวณค่าสถิติจากตัวอย่างทั้งหมดทุกกลุ่ม จะได้ค่าสถิติจากตัวอย่าง (sample statistics) และนำมาแจกแจงความถี่ จะได้ " **การแจกแจงค่าสถิติจากตัวอย่าง"** ซึ่งมีลักษณะแตกต่างกันตามค่าสถิติที่นำมาแจกแจง เพื่อให้เข้าใจชัดเจนผู้เขียนยกตัวอย่างการแจกแจงของตัวแปร X ที่ได้จากประชากรขนาดเล็กมาก (N = 3) มาแสดงให้เห็นการแจกแจงทั้ง 3 แบบ คือ ก) การแจกแจงของประชากรขนาด 3 คน (N = 3) มีค่าของตัวแปร X = 1, 2, 3 และ ข) การแจกแจงของตัวอย่างหนึ่งกลุ่มที่สุ่มเลือกมาขนาด 2 คน (n = 2) มีค่าของตัวแปร X = 2, 3 และ ค) การแจกแจงของค่าเฉลี่ยตัวแปร X หรือ $\bar{X}$ จากตัวอย่างที่เลือกแบบใส่คืนทั้งหมดที่เป็นไปได้ มีจำนวน 9 กลุ่ม ดังนี้



ก) การแจกแจงของประชากร (x มีการแจกแจงแบบเอกรูป)

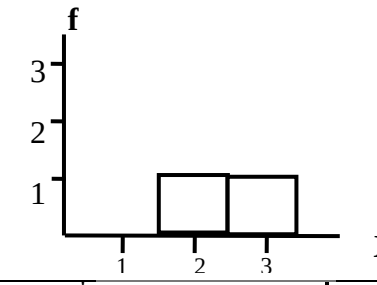
ค่าของตัวแปร X = 1, 2, 3; N = 3

พารามิเตอร์

ค่าเฉลี่ยของ X = μ = (1+2+3)/3 = 6/3 = 2

ความแปรปรวนของ X = σ^2 = $\{(-1)^2+(0)^2+(1)^2\}/3$ = 2/3 = 0.67

ส่วนเบี่ยงเบนมาตรฐานของ X = σ = $\sqrt{2/3}$ = 0.8165



ข) การแจกแจงของตัวอย่าง (x มีการแจกแจงแบบเอกรูป)

ค่าของตัวแปร X = 2, 3; n = 2

ค่าสถิติ

ค่าเฉลี่ยของตัวแปร X = $\bar{X}$ = (2+3)/2 = 2.50

ความแปรปรวนของตัวแปร X = SD^2 = $\{(-0.5)^2+(0.5)^2\}/(2-1)$ = 0.50

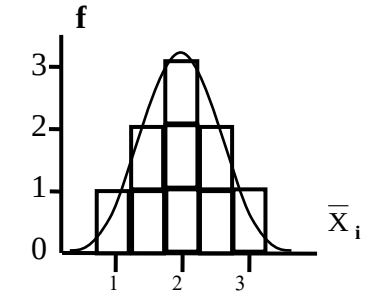
ส่วนเบี่ยงเบนมาตรฐานของตัวแปร X = SD = $\sqrt{0.50}$ = 0.7071

ตัวอย่างที่เป็นไปได้ = 9 กลุ่ม

กลุ่มที่	ค่าตัวแปร X	$\bar{X}_i$
1	1, 1	$\bar{X}_1$ = 1.0
2	1, 2	$\bar{X}_2$ = 1.5
3	1, 3	$\bar{X}_3$ = 2.0
4	2, 1	$\bar{X}_4$ = 1.5
5	2, 2	$\bar{X}_5$ = 2.0
6	2, 3	$\bar{X}_6$ = 2.5
7	3, 1	$\bar{X}_7$ = 2.0
8	3, 2	$\bar{X}_8$ = 2.5
9	3, 3	$\bar{X}_9$ = 3.0

ค) การแจกแจงของค่าเฉลี่ยจากตัวอย่าง ($\bar{X}$ มีการแจกแจงแบบปกติ)

$\bar{X}_i$	f	$f(\bar{X}_i)$	$f(\bar{X}_i-\mu)^2$
1.0	1	1.0	1.00
1.5	2	3.0	0.50
2.0	3	6.0	0.00
2.5	2	5.0	0.50
3.0	1	3.0	1.00
รวม	**9**	**18.0**	**3.00**



ค่าเฉลี่ยของ $\bar{X}_i$ = $\mu_{\bar{X}}$ = (18)/(9) = 2.00 = μ

ความแปรปรวนของ $\bar{X}_i$ = $\sigma^2_{\bar{X}}$ = (3)/(9) = 1/3 = 0.33

ส่วนเบี่ยงเบนมาตรฐานของ $\bar{X}_i$ = $\sigma_{\bar{X}}$ = $\sqrt{0.33}$ = 0.5744

ภาพ 2 การแจกแจงของประชากร ของตัวอย่าง และของค่าเฉลี่ยจากตัวอย่าง

เมื่อปรียบเทียบกัน ค่าพารามิเตอร์ที่ได้จากการแจกแจงประชากร และค่าสถิติของกลุ่มตัวอย่าง พบว่า มีค่าแตกต่างกันทั้งค่าเฉลี่ยและค่าความแปรปรวน แต่พารามิเตอร์จากการแจกแจงของประชากรและ

การแจกแจงของค่าเฉลี่ยจากตัวอย่าง มีพารามิเตอร์ค่าเฉลี่ยเท่ากัน และพารามิเตอร์ความแปรปรวนมีความเกี่ยวข้องกัน ตาม**ทฤษฎีขีดจำกัดกลาง (Central Limit Theorem)** ซึ่งสรุปได้ดังต่อไปนี้

พารามิเตอร์ค่าเฉลี่ยของ $\overline{X}_i$ $= \mu_{\overline{X}} = 2.00 = \mu$ = พารามิเตอร์ค่าเฉลี่ยของประชากร

พารามิเตอร์ความแปรปรวนของ $\overline{X}_i$ $= \sigma^2_{\overline{X}} = \frac{(1)}{(3)} = \frac{(2)}{(3)(2)} = \left[\frac{(2)}{(3)}\right]\left[\frac{(1)}{(2)}\right] = \frac{\sigma^2}{2} = \frac{\sigma^2}{n}$

ทฤษฎีขีดจำกัดกลาง กล่าวว่า เมื่อเลือกตัวอย่างขนาด n หน่วย เป็นจำนวนมาก จากประชากรขนาด N หน่วย ที่มีพารามิเตอร์ค่าเฉลี่ย = μ และความแปรปรวน = σ^2 มาแจกแจง จะได้การแจกแจงของ ค่าเฉลี่ยจากตัวอย่าง (distribution of sample means or $\overline{X}_i$) มีลักษณะดังนี้

1) ถ้าขนาดตัวอย่าง (n) ใหญ่ขึ้น การแจกแจงค่าเฉลี่ย จากตัวอย่าง ($\overline{X}_i$) มีการแจกแจง ใกล้เคียงกับการแจกแจงปกติมากขึ้น ไม่ว่าประชากรจะมีการแจกแจงเป็นหรือไม่เป็นโค้งปกติ

2) พารามิเตอร์ค่าเฉลี่ยของ $\overline{X}_i = \mu_{\overline{X}} = \mu$ = พารามิเตอร์ค่าเฉลี่ยของประชากร

3) พารามิเตอร์ความแปรปรวน และส่วนเบี่ยงเบนมาตรฐาน มีค่าแตกต่างกันเป็นสองแบบ

แบบที่หนึ่ง กรณีเลือกตัวอย่างแบบคืนที่ $\sigma^2_{\overline{X}} = \frac{\sigma^2}{n}$; $\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$

แบบที่สอง กรณีเลือกตัวอย่างแบบไม่คืนที่ $\sigma^2_{\overline{X}} = \left[\frac{N-n}{N-1}\right]\left[\frac{\sigma^2}{n}\right]$; $\sigma_{\overline{X}} = \left[\sqrt{\frac{N-n}{N-1}}\right]\left[\frac{\sigma}{\sqrt{n}}\right]$

ในที่นี้ $\sigma^2_{\overline{X}}$ = พารามิเตอร์**ความแปรปรวนมาตรฐาน (standard variance)** ของ $\overline{X}_i$ เรียกย่อๆ ว่า $SV_{\overline{X}}$
= พารามิเตอร์ความแปรปรวนของค่าเฉลี่ยจากตัวอย่าง ($\overline{X}_i$)

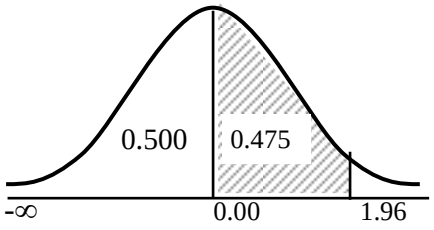
$\sigma_{\overline{X}}$ = พารามิเตอร์**ความคลาดเคลื่อนมาตรฐาน (standard error)** ของ $\overline{X}_i$ เรียกย่อๆ ว่า $SE_{\overline{X}}$
= พารามิเตอร์ส่วนเบี่ยงเบนมาตรฐานของค่าเฉลี่ยจากตัวอย่าง ($\overline{X}_i$)
= พารามิเตอร์บ่งชี้ความแตกต่างระหว่างค่า เฉลี่ยของตัวอย่าง กับค่าพารามิเตอร์ ค่าเฉลี่ยของประชากร จึงเรียกอีกชื่อหนึ่งว่า **ความคลาดเคลื่อนของการเลือกตัวอย่าง (sampling error)**

4) ความคลาดเคลื่อนมาตรฐานของค่าเฉลี่ย ($\sigma_{\overline{X}}$) มีความสัมพันธ์แบบผกผันกับขนาดตัวอย่าง (n) นั่นคือ ตามสูตร $\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$ ถ้าใช้ตัวอย่างขนาดใหญ่มากขึ้น ทำให้มี ความคลาดเคลื่อนมาตรฐาน หรือความคลาดเคลื่อนจากการเลือกตัวอย่างมีค่าลดน้อยลง

5) เนื่องจากตัวแปร ($\overline{X}_i$) มีการแจกแจงแบบปกติ และมีค่าเฉลี่ยเท่ากับ μ ค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ $\sigma_{\overline{X}}$ จึงสามารถเปลี่ยนรูปให้เป็นคะแนนมาตรฐาน (z) ได้ โดยใช้สูตร ซึ่งได้การแจกแจงของตัวแปร $z_{\overline{X}}$ เรียกว่า **การแจกแจงปกติมาตรฐาน (standard normal distribution)**

$$z_{\overline{X}} = \frac{\overline{X} - \mu_{\overline{X}}}{\sigma_{\overline{X}}} = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$$

นักสถิติได้คำนวณและจัดทำตารางพื้นที่ภายใต้โค้งปกติของคะแนนมาตรฐาน เพื่ออำนวยความสะดวกให้นักวิจัยใช้ประโยชน์หาพื้นที่ภายใต้โค้งปกติได้ เมื่อรู้ค่าของตัวแปร z ดังภาพต่อไปนี้

การหาพื้นที่ภายใต้โค้งปกติ เมื่อ $-\infty < z_{\bar{x}} \leq 1.96$ จากตารางพื้นที่ภายใต้โค้งปกติ พื้นที่ใต้โค้งปกติเมื่อ $0.00 < z_{\bar{x}} \leq 1.96 = P(0.00 < z_{\bar{x}} \leq 1.96) = 0.475$ พื้นที่ใต้โค้งปกติเมื่อ $-\infty < z_{\bar{x}} \leq 0.00 = P(-\infty < z_{\bar{x}} \leq 0.00) = 0.500$ ดังนั้น พื้นที่ใต้โค้งปกติเมื่อ $-\infty < z_{\bar{x}} \leq 1.96 = 0.500+0.475 = 0.975$ นั่นคือ มี**ตัวอย่างร้อยละ 97.50 ของตัวอย่างทั้งหมดทุกกลุ่มที่เป็นไปได้** ซึ่งมีคะแนนมาตรฐานของค่าเฉลี่ย $z_{\bar{x}} \leq 1.96$	 **พื้นที่ภายใต้โค้งปกติ เมื่อ $-\infty < z_{\bar{x}} \leq 1.96$**

ภาพ 3 การใช้ประโยชน์ตารางพื้นที่ภายใต้โค้งปกติในการหาพื้นที่ เมื่อทราบค่าคะแนน $z_{\bar{x}}$

3. การทดสอบสมมุติฐาน (Hypothesis Testing) และการประมาณค่าพารามิเตอร์ (Parameter Estimation)

การเปลี่ยนรูปตัวแปรค่าสถิติจากตัวอย่าง ในที่นี้คือ ค่าสถิติ $\bar{X}$ ให้เป็นคะแนนมาตรฐาน z จะได้สูตรพื้นฐาน 2 สูตร คือ สูตรสำหรับการทดสอบสมมุติฐาน และสูตรสำหรับการประมาณค่าพารามิเตอร์ในรูปช่วงเชื่อมั่น (confidence interval) ซึ่งได้มาจากการแทนค่า z ที่ระดับนัยสำคัญ (significance level = α) ที่กำหนด และย้ายข้างสมการ ดังภาพแสดงสูตร 4 แบบ คือ 1) z-test สำหรับการทดสอบสมมุติฐานเกี่ยวกับค่าเฉลี่ย 2) สูตรทั่วไปสำหรับการทดสอบสมมุติฐานเกี่ยวกับค่าสถิติชนิดอื่นๆ 3) สูตรการประมาณค่าพารามิเตอร์ค่าเฉลี่ย และ 4) สูตรทั่วไปสำหรับการประมาณค่าพารามิเตอร์ชนิดอื่นๆ

1. z-test: $z_{\bar{x}} = \dfrac{\bar{X} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \dfrac{\bar{X} - \mu}{\sigma/\sqrt{n}}$ สำหรับการทดสอบ H_0: $\mu = 0$	2. สถิติทดสอบ (test statistics) $= \dfrac{\text{สถิติ} - \text{พารามิเตอร์}}{\text{ส่วนเบี่ยงเบนมาตรฐานของสถิติ}}$ สำหรับการทดสอบ H_0: $\theta = 0$ เมื่อ θ = พารามิเตอร์ชนิดอื่นๆ
3. การประมาณค่าพารามิเตอร์ μ ที่ช่วงเชื่อมั่น 95% (95% confidence interval of parameter μ) $= \bar{X} \pm (1.96)(\sigma_{\bar{x}})$	4. ช่วงเชื่อมั่นการประมาณค่าพารามิเตอร์ θ ที่ช่วงเชื่อมั่น $(1-\alpha)$ % (95% confidence interval of parameter θ) เมื่อ θ = พารามิเตอร์ชนิดอื่นๆ $= \hat{\theta} \pm (z_{\alpha/2})(\sigma_{\hat{\theta}})$

ภาพ 4 สูตรสำหรับการทดสอบสมมุติฐาน และการประมาณค่าพารามิเตอร์

Kirk (1995) อธิบายว่า การทดสอบสมมุติฐาน เป็นกระบวนการประเมินนัยสำคัญทางสถิติ (statistical significance) ของข้อค้นพบจากตัวอย่าง หรือ กระบวนการตัดสินใจปฏิเสธหรือไม่ปฏิเสธ (reject or not reject) สมมุติฐานทางสถิติ โดยใช้ข้อมูลค่าสถิติจาตัวอย่าง เพื่อให้ได้ข้อสรุปเกี่ยวกับพารามิเตอร์ของประชากร การทดสอบสมมุติฐาน หลังจากกำหนดสมมุติฐานวิจัยแล้ว มีขั้นตอนรวม 6 ขั้นตอน คือ 1) การกำหนดสมมุติฐานทางสถิติ 2) การระบุสถิติทดสอบ (test statistics) 3) การกำหนดการแจกแจงค่าสถิติจากตัวอย่าง และขนาดตัวอย่าง 4) การกำหนดระดับนัยสำคัญทางสถิติ 5) การเลือกตัวอย่างสุ่ม การรวบรวมข้อมูล และการวิเคราะห์คำนวณค่าสถิติทดสอบ และ 6) การตัดสินใจปฏิเสธ/ไม่ปฏิเสธสมมุติฐานหลัก และแปลความหมาย ส่วนการประมาณค่าพารามิเตอร์ ทำได้ 2 แบบ คือการประมาณค่าแบบจุดด้วยค่าสถิติ และการประมาณค่าแบบช่วง โดยการกำหนดช่วงคะแนนที่เชื่อมั่นได้ว่ารวมค่าพารามิเตอร์อยู่ในช่วงคะแนนนั้น เรียกว่า ช่วงเชื่อมั่น (confidence interval) ในอดีตนักวิจัยนิยมใช้การทดสอบสมมุติฐานมากกว่าการประมาณ

ค่าพารามิเตอร์ แต่ปัจจุบันนักวิจัยเริ่มใช้การประมาณค่าพารามิเตอร์มากขึ้น เพราะให้สารสนเทศมากกว่าสารสนเทศที่ได้จากการทดสอบสมมุติฐาน

การทดสอบสมมุติฐานและการประมาณค่าพารามิเตอร์ มีสาระสำคัญที่ต้องทำความเข้าใจ ได้แก่ การกำหนดสมมุติฐานทางสถิติ การตัดสินใจในการทดสอบสมมุติฐาน ประเภทของการทดสอบสมมุติฐาน ดังความหมายโดยสรุป (Kirk, 1995; Howell, 2010; Kerlinger & Lee, 2000; นงลักษณ์ วิรัชชัย, 2553) ตามสาระสังเขป และภาพ 5 ต่อไปนี้

3.1 การกำหนดสมมุติฐานทางสถิติ นักวิจัยกำหนดสมมุติฐานทางสถิติให้สอดคล้องกับสมมุติฐานวิจัย รวม 2 ชุด ชุดแรก เรียกว่า**สมมุติฐานหลัก** หรือ **สมมุติฐานศูนย์ (null hypothesis)** และชุดที่สองเรียกว่า **สมมุติฐานเลือก (alternative hypothesis)** สมมุติฐานทั้งสองชุด เป็นข้อความเกี่ยวกับพารามิเตอร์ของประชากรตั้งแต่หนึ่งค่าขึ้นไป นิยมเขียนในรูปสัญลักษณ์ทางสถิติ โดยมีหลักการกำหนดสมมุติฐานทางสถิติรวม 3 ประการ คือ ก) สมมุติฐานหลักเป็นข้อความแสดงความเท่ากันของพารามิเตอร์ ข) สมมุติฐานเลือกเป็นข้อความที่แสดงลักษณะพารามิเตอร์ซึ่งตรงตามสมมุติฐานวิจัย และ ค) ข้อความในสมมุติฐานทางสถิติทั้งสมมุติฐานหลัก และสมมุติฐานเลือก ครอบคลุมสภาพของพารามิเตอร์ที่เป็นได้ทั้งหมด

3.2 การตัดสินใจในการทดสอบสมมุติฐาน เมื่อมีสมมุติฐานหลักเป็นจริงหรือเป็นเท็จ ทำให้ได้ผลการตัดสินใจในการทดสอบสมมุติฐาน รวม 4 แบบ ดังนี้

แบบแรก การตัดสินใจไม่ถูกต้อง เกิด ความคลาดเคลื่อนประเภทที่ I (type I error = α) หมายถึง ความน่าจะเป็น ในการปฏิเสธสมมุติฐานหลัก (H_0) เมื่อสมมุติฐานหลักถูกต้อง คำที่เกี่ยวข้อง คือ **ระดับนัยสำคัญ (significance level = α)** หมายถึง ระดับความน่าจะเป็นที่นักวิจัยยอมให้เกิดความคลาดเคลื่อนประเภทที่ I

แบบที่สอง การตัดสินใจถูกต้อง หมายถึง การตัดสินใจไม่ปฏิเสธสมมุติฐานหลักที่ถูกต้องเป็นจริง มีค่าความน่าจะเป็น = 1-α คำที่เกี่ยวข้อง คือ **ระดับความเชื่อมั่น (confidence level = 1-α)** หมายถึง ระดับความน่าจะเป็น ที่ตัดสินใจถูกต้องเมื่อสมมุติฐานหลักถูกต้อง และ **จุดวิกฤติ (critical point)** หมายถึง จุดที่แบ่งเขตความคลาดเคลื่อนประเภทที่ I หรือ**เขตวิกฤติ (critical region)** กับเขตการตัดสินใจที่ถูกต้อง

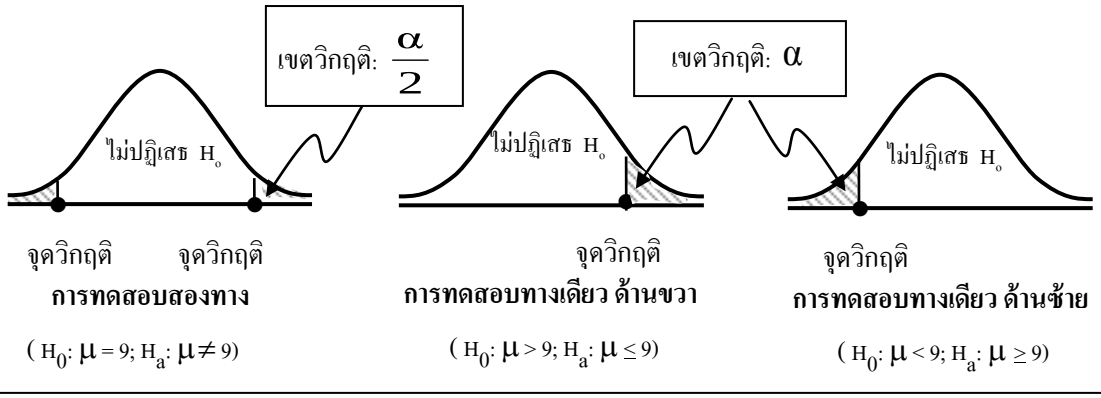
แบบที่สาม การตัดสินใจที่ไม่ถูกต้อง เกิดความ **คลาดเคลื่อนประเภทที่ II (type II error = β)** หมายถึง ความน่าจะเป็นในการไม่ปฏิเสธสมมุติฐานหลักเมื่อสมมุติฐานหลักเป็นเท็จ ตามภาพ 5 จะเห็นได้ชัดว่า ค่า α และ β เกี่ยวข้องสัมพันธ์กัน เมื่อค่าหนึ่งลดลงอีกค่าหนึ่งจะเพิ่มขึ้น

แบบที่สี่ การตัดสินใจถูกต้อง เมื่อตัดสินใจปฏิเสธสมมุติฐานหลักที่ถูกต้องเป็นจริง คำที่เกี่ยวข้อง คือ **อำนาจการทดสอบ (power of the test = 1 - β)**

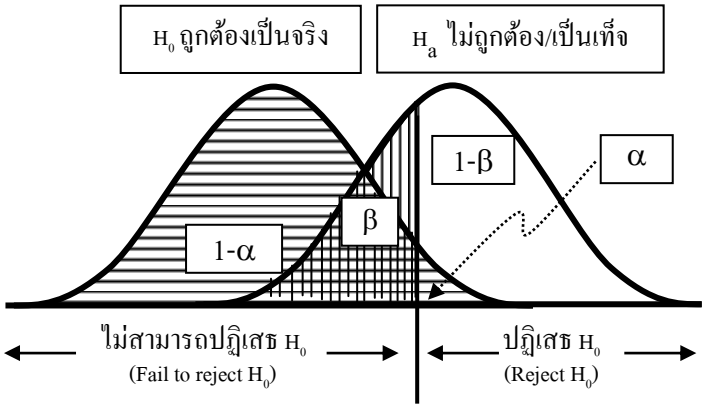
3.3 ประเภทของการทดสอบสมมุติฐาน แบ่งเป็น 2 ประเภท ตามลักษณะที่แตกต่างกันตามสมมุติฐานเลือก ในกรณีที่มีการกำหนดสมมุติฐานเลือกหรือสมมุติฐานวิจัยแบบไม่มีทิศทาง การทดสอบสมมุติฐานเป็น **การทดสอบสองทาง (two tailed test)** โดยมีพื้นที่เขตวิกฤติสองด้าน ด้านซ้ายและด้านขวาเท่ากับ $\alpha/2$ ส่วนกรณีที่มีการกำหนดสมมุติฐานเลือกหรือสมมุติฐานวิจัยแบบมีทิศทาง การทดสอบสมุติฐาน

เป็น**การทดสอบทางเดียว (one tailed test)** โดยมีพื้นที่เขตวิกฤติ เท่ากับ α ซึ่งอาจอยู่ด้านซ้าย หรือด้านขวา ตามลักษณะของสมมุติฐานเลือก

การกำหนดสมมุติฐานทางสถิติ 3 แบบ และประเภทการทดสอบสมมุติฐาน



การตัดสินใจในการการทดสอบสมมุติฐาน



		สถานการณ์ที่เป็นจริง (True Situation)	
		H_0 ถูกต้องเป็นจริง	**H_0 ไม่ถูกต้อง/เป็นเท็จ**
การตัดสินใจของนักวิจัย	ไม่สามารถปฏิเสธ H_0 (Fail to reject H_0)	ตัดสินใจถูกต้อง (Confidence level = $1-\alpha$)	ความคลาดเคลื่อนประเภทที่ II (Type II Error = β)
	ปฏิเสธ H_0 (Reject H_0)	ความคลาดเคลื่อนประเภทที่ I (Type I Error = α)	ตัดสินใจถูกต้อง (Power of Test = $1-\beta$)

ภาพ 5 การกำหนดสมมุติฐานทางสถิติ ประเภท และการตัดสินใจในการทดสอบสมมุติฐาน

ตอนที่ 2 วิธีการกำหนดขนาดตัวอย่างจากตารางสำเร็จ

การกำหนด หรือการประมาณค่าขนาดตัวอย่าง สำหรับการทดสอบสมมุติฐาน ขึ้นอยู่กับประเภทของสถิติทดสอบและลักษณะการแจกแจงของค่าสถิติจากตัวอย่าง นอกจากนี้ขนาดตัวอย่าง (n) มีความสัมพันธ์เกี่ยวข้องกับ ก) ระดับนัยสำคัญ (α) และอำนาจการทดสอบ ($1-\beta$) ข) ระดับความถูกต้องในการทดสอบ หรือความแตกต่างระหว่างค่าสถิติกับพารามิเตอร์ที่ต้องการทดสอบ ($|\bar{X}_i - \mu|$) และ ค) ค่าส่วนเบี่ยงเบนมาตรฐานของตัวแปรของประชากร (σ) ดังนั้นโดยหลักการ การกำหนด/การประมาณค่าขนาด

ตัวอย่าง ต้องใช้สูตรซึ่งแตกต่างกันตามประเภทของสถิติทดสอบ จากนั้นจึงกำหนดค่าอื่นๆ ได้แก่ α, (1-β), $|\bar{x}_i - \mu|$, σ เพื่อแทนค่าลงในสูตร ให้เหลือตัวไม่ทราบค่า คือ n เพียงตัวเดียว เพื่อ แก้สมการหาค่าขนาดตัวอย่าง (n) ที่เหมาะสม ได้ แต่ในทางปฏิบัติ มีนักสถิติหลายคน เช่น Yamane; Cochran; Krejcie and Morgan; ศิริชัย กาญจนวาสี ; Smith และ Cohen ได้คำนวณขนาดตัวอย่างโดยใช้สูตรแตกต่างกัน มีทั้งสูตรที่ง่ายและซับซ้อน และพัฒนาเป็นตารางสำเร็จ แยกได้เป็น 3 ประเภท ดังนี้

1. ตารางสำเร็จที่สร้างตามข้อกำหนดของข้อมูลจากสัดส่วนตัวแปรทวินาม

นักสถิติสร้างตารางสำเร็จสำหรับการประมาณค่า/คำนวณค่าขนาดตัวอย่าง ในช่วงคริสต์ทศวรรษ 1960s โดยมีข้อตกลงเบื้องต้นว่าตัวแปรเป็นตัวแปรทวิวิภาค (dichotomous variable) มีเพียงสองค่า กรณีตัวแปรมีหลายค่าใช้วิธีรวมค่าตัวแปรให้เหลือเพียงสองค่า ตารางสำเร็จที่นิยมใช้กันมากในช่วงนั้นมีดังนี้

1.1 ตารางสำเร็จของ Cochran

Cocharan (1963) พัฒนาตารางสำเร็จสำหรับการประมาณค่าขนาดตัวอย่าง กรณีประชากรขนาดใหญ่ เมื่อกำหนดให้ z = ค่าสถิติ z ที่ระดับความเชื่อมั่น (1 - α) % เช่น ที่ระดับความเชื่อมั่น 95 % ใช้ z = $\pm$ 1.96 เป็นต้น e = ระดับความถูกต้องในการทดสอบสมมุติฐาน p = สัดส่วนของหน่วยตัวอย่างที่มีคุณสมบัติตามที่นักวิจัยสนใจศึกษา และ q = (1 – p) โดยนำสูตรการทดสอบสมมุติฐานทางสถิติเกี่ยวกับสัดส่วน มาสร้างสูตรใหม่ จากสูตร z-test: $z = \frac{p - P}{\sqrt{(p)(q)}/\sqrt{n}}$ ยกกำลังสองทั้งสองข้าง: $z^2 = \frac{e^2}{(p)(q)/n}$ ได้สูตร $n = \frac{(z^2)(p)(q)}{(e^2)}$ กรณีประชากรขนาดเล็ก ต้องปรับแก้ขนาดกลุ่มตัวอย่างตามสูตร adjusted $n = \frac{n}{1+\frac{(n-1)}{N}}$ ตารางสำเร็จนี้มีจุดอ่อนคือ ตารางชุดเดียวใช้สำหรับสถิติทดสอบทุกชนิด และต้องเป็นตัวแปรทวิวิภาค

1.2 ตารางสำเร็จของ Yamane

ตารางสำเร็จของ Yamane (1973) พัฒนาเมื่อปี 1967 เป็นตารางสำเร็จที่นักวิจัยนิยมใช้ในระยะแรก มีการคำนวณขนาดตัวอย่าง เมื่อกำหนดให้ n = ขนาดตัวอย่าง N = ขนาดประชากร และ e = ระดับความแตกต่างของค่าสถิติกับพารามิเตอร์ (error = e) มีสูตรการคำนวณ คือ $n = \frac{N}{1+(N)(e^2)}$

ต่อมา Yamane (1973) ปรับปรุงสูตรการคำนวณให้ถูกต้องมากขึ้น โดยเพิ่มค่า π = ความแปรปรวนของประชากร โดยกำหนดจากตัวแปรทวิวิภาคเท่ากับ .50 และ z = ค่าประมาณคะแนน z ที่ระดับนัยสำคัญ α (ใช้ z = 2 ที่ α = .05 และ ใช้ z = 3 ที่ α = .01) โดยคำนวณตามสูตรดังนี้ $n = \frac{(z^2)(\pi)(1-\pi)(N)}{(z^2)(\pi)(1-\pi)+(N)(e^2)}$ ตารางสำเร็จของ Yamane มี 2 ตาราง สำหรับกรณี α = .05 และ .01

Butkovith, Smith & Hoffman (2004) เสนอว่า ตารางสำเร็จประเภทนี้มีตารางเดียว คำนวณจากขนาดประชากร และความคลาดเคลื่อนเป็นหลัก แต่ใช้กับสถิติทดสอบทุกประเภท จึงมีความเหมาะสมน้อย กล่าวโดยสรุปจุดอ่อนของตารางประเภทนี้มี 3 ประการ ก) ตารางชุดเดียวใช้สำหรับการกำหนดขนาดตัวอย่างที่ใช้สถิติทดสอบทุกชนิด ข) ค่า z ใช้ค่าประมาณ 2 แทน 1.96 และใช้ 3 แทน 2.58 ทำให้ผลการคำนวณคลาดเคลื่อน และ ค) มีข้อกำหนดว่าตัวแปรเป็นตัวแปรทวิวิภาค มีความแปรปรวนสูงสุด = π = 0.50 ซึ่งไม่เหมาะสมกับลักษณะตัวแปรชนิดอื่นที่ไม่ใช้ตัวแปรทวิวิภาค

1.3 ตารางสำเร็จของ Krejzie and Mogan

Krejzie and Morgan (1970) พัฒนาตารางสำเร็จรูป เมื่อกำหนดให้ s = ขนาดตัวอย่างที่ต้องการ N = ขนาดประชากร p = ความแปรปรวนของประชากรเมื่อตัวแปรเป็นตัวแปรทวินาม กำหนดให้ = .50, ค่าสถิติไค-สแควร์ = χ^2 ที่องศาอิสระ (df) = 1 ซึ่งกำหนดให้เป็นค่าคงที่ = 3.841 และ d = ระดับความถูกต้องในการทดสอบสมมุติฐาน โดยใช้สูตรดังนี้ $s = \frac{(\chi^2)(N)(p)(1-p)}{(e^2)(N-1)+(\chi^2)(p)(1-p)}$ ตารางที่ได้มี 2 ตาราง ตามระดับนัยสำคัญ .01 และ .05 ซึ่งมีจุดอ่อนเช่นเดียวกับกับจุดอ่อนของตารางสำเร็จของ Yamane แต่ดีขึ้นเมื่อ Krejcie & Morgan ได้ปรับปรุงในช่วงต่อมาโดยใช้ค่า χ^2 มีค่าต่างกันตามที่นักวิจัยกำหนด

2. ตารางสำเร็จที่สร้างตามข้อกำหนดของข้อมูลว่าเป็นตัวแปรต่อเนื่อง

นักสถิติสร้างตารางสำเร็จสำหรับการประมาณค่า/คำนวณค่าขนาดกลุ่มตัวอย่าง ตามข้อกำหนดของข้อมูลว่าเป็นตัวแปรต่อเนื่อง ใช้สูตรการทดสอบสมมุติฐานทางสถิติเกี่ยวกับค่าเฉลี่ย มาสร้างเป็นสูตรใหม่ จากสูตร z-test: $Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ ยกกำลังสองทั้งสองข้าง: $z^2 = \frac{e^2}{\sigma^2/n}$ ได้สูตร $n = \frac{(z^2)(\sigma^2)}{e^2}$ Israel (2009) อธิบายว่าสูตรที่ได้นี้คล้ายคลึงกับสูตรของ Cochran (1963) และเหมาะสมกับกรณีมีตัวแปรต่อเนื่อง จุดอ่อนของสูตรนี้อยู่ที่การกำหนดค่าประมาณพารามิเตอร์ความแปรปรวนของประชากร (σ^2) ซึ่งเป็นค่าประมาณที่มีความลำเอียง ทำให้ผลการประมาณค่าขนาดตัวอย่างไม่ใคร่ถูกต้อง ตารางสำเร็จที่พัฒนาโดย ศิริชัย กาญจนวาสี (2550) เป็นการนำแนวคิดในการพัฒนาตารางสำเร็จตามข้อกำหนดของข้อมูลว่าเป็นตัวแปรต่อเนื่องมาใช้พัฒนาสูตรของ Krejcie & Morgan และมีจุดอ่อนในลักษณะเดียวกัน

3. ชุดของตารางสำเร็จสำหรับใช้ในการทดสอบค่าสถิติหลายวิธี ของ Cohen

หนังสือเรื่อง “Statistical power analysis for the behavioral sciences” ของ Cohen ฉบับพิมพ์ครั้งที่ 1, 2 (1977, 1988) ได้รับการยอมรับว่าเป็นการพัฒนาชุดตารางสำเร็จสำหรับการวิเคราะห์เพื่อประมาณค่าขนาดกลุ่มตัวอย่าง อำนาจการทดสอบ เมื่อทราบค่าของขนาดอิทธิพล และระดับนัยสำคัญ ที่อำนวยความสะดวกให้นักวิจัยสามารถกำหนดขนาดตัวอย่างได้ถูกต้องมากขึ้น และประมาณค่าอำนาจการทดสอบได้อย่างเหมาะสม หนังสือแสดงตารางการอ่านค่าขนาดตัวอย่าง อำนาจการทดสอบ สำหรับสถิติวิเคราะห์แต่ละแบบ รวม 5 ประเภท ได้แก่ t-test, z-test, F-test, χ^2- test, และ exact test ซึ่งใช้งานได้ในการทดสอบตั้งแต่กรณีตัวแปรตามตัวเดียง และตัวอย่างกลุ่มเดียว ไปจนถึงกรณีการทดสอบสถิติวิเคราะห์ตัวแปรพหุนาม (multivariate statistical analysis) โดยมีตารางแต่ละตารางสำหรับการทดสอบแต่ละแบบ หนังสือจึงมีความหนามากโดยเฉพาะในฉบับพิมพ์ครั้งที่สอง และเป็นหนังสือที่นักวิจัยทั่วโลกรู้จักและใช้ประโยชน์กันอย่างกว้างขวาง แต่การใช้งานยังไม่สะดวกมากเท่ากับการกำหนดขนาดตัวอย่างด้วยโปรแกรมคอมพิวเตอร์ อย่างไรก็ดีผู้เขียนเชื่อว่าการได้ศึกษาหนังสือของ Cohen จะช่วยให้นักวิจัยเข้าถึงหลักการและสูตรในการกำหนดขนาดตัวอย่าง และการประมาณค่าอำนาจการทดสอบสำหรับการวิจัย ซึ่งต้องใช้การแจกแจงค่าสถิติจากตัวอย่าง (sampling distribution) ที่แตกต่างกันตามสถิติทดสอบที่นักวิจัยต้องการ จะเห็นได้ว่า จำนวนสถิติทดสอบยิ่งมีมากเท่าไร ตารางสำเร็จที่ต้องใช้มีจำนวนมากเท่านั้นด้วย

ตอนที่ 3 วิธีการกำหนดขนาดตัวอย่างโดยใช้โปรแกรมคอมพิวเตอร์

การกำหนดขนาดตัวอย่างเกี่ยวข้องสัมพันธ์กันกับการวิเคราะห์อำนาจการทดสอบ อย่างแยกกันไม่ออก วงการสถิติรับรู้ว่าการวิเคราะห์อำนาจการทดสอบแยกได้เป็น 3 ยุค ยุคแรกนักสถิติรับรู้ความสำคัญของอำนาจการทดสอบว่ามีผลต่อการทดสอบสมมุติฐาน แต่ไม่มีการนำไปใช้ในการปฏิบัติจริง ยุคที่สองช่วงต้นคริสตศตวรรษที่ 1980s เมื่อมีการพัฒนาแผนภูมิอำนาจการทดสอบ (power chart) เช่น แผนภูมิที่พัฒนาโดย Scheffe โดย Tang และมีการพัฒนาชุดตารางสำเร็จที่พัฒนาโดย Cohen (1977) สำหรับการวิเคราะห์หาอำนาจการทดสอบ ส่วนยุคที่สามเริ่มต้นประมาณปี ค.ศ. 1988 เมื่อมีการพัฒนาโปรแกรมคอมพิวเตอร์จำนวนมากทั้งโปรแกรมให้เปล่า โปรแกรมร่วมใช้ และโปรแกรมเชิงพาณิชย์ (free-ware, shareware and commercial program) เพื่ออำนวยความสะดวกในการประมาณค่าอำนาจการทดสอบ เช่น โปรแกรม MS/PC-DOS computer ที่พัฒนาโดย Goldstein และโปรแกรม G*Power นับแต่นั้นมาการประมาณค่าอำนาจการทดสอบและการกำหนดขนาดตัวอย่างมีประสิทธิภาพมากขึ้น อย่างไรก็ดีโปรแกรมคอมพิวเตอร์ส่วนใหญ่เป็นผลการพัฒนาในเชิงธุรกิจ ส่วนน้อยเป็นผลการพัฒนาให้ใช้งานได้โดยไม่เสียค่าใช้จ่าย ในจำนวนนี้บางโปรแกรมก็มีประสิทธิภาพไม่สูงมากนัก ในทัศนะของ Howell (2010) และผู้เขียน พิจารณาเห็นว่าโปรแกรม G*Power เป็นโปรแกรมที่ดีมากและสมบูรณ์ ที่สร้างจากสูตรของ Cohen (1977) และผ่านการตรวจสอบ/รับรองคุณภาพโดยนักวิจัยหลายคน สมควรที่นักวิจัยไทยได้เรียนรู้และใช้ประโยชน์เพื่อให้การกำหนดตัวอย่างให้ถูกต้องสมบูรณ์และทันสมัยเหมาะสมกับการเผยแพร่ผลงานวิจัยในระดับสากล

การเสนอสาระในตอนนี้ ผู้วิจัยนำเสนอแยกเป็น 3 เรื่อง คือ ก) แนะนำโปรแกรม G*Power 3 และ ข) การใช้โปรแกรม G*Power และ ค) ตัวอย่างการใช้โปรแกรม G*Power ดังนี้

3.1 แนะนำโปรแกรม G*Power

โปรแกรม G*Power ได้รับการพัฒนาเมื่อปี 1996 เป็นโปรแกรมสำเร็จรูปที่ใช้งานได้ทั้ง Windows and Mac platform ต่อมามีการพัฒนา G*Power 2 และ G*Power 3 ในช่วงปี 2003-2007 และล่าสุดมีการพัฒนา G*Power 3.1 แต่การพัฒนาคู่มือการใช้โปรแกรมยังไม่เสร็จสมบูรณ์ ต้องใช้คู่มือการใช้โปรแกรม G*Power 2 ควบคู่กันไป ผู้อ่านสามารถ download โปรแกรม และคู่มือ ได้จาก web page ของมหาวิทยาลัย Heinrich-Heine-Universität ที่ web site: http://www.psycho.uni-duesseldorf.de/abteilungen/aap/gpower3/ (Erdfelder, Faul & Buchner, 1996; Buchner, Erdfelder, & Faul, 1997; Faul, Erdfelder, Buchner & Lang, 2009; Buchner, 2010) และ http://www.psycho.uni-duesseldorf.de/abteilungen/aap/gpower3. นอกจากนี้เว็บเพจของ G*Power ยังมี ‘Tutorial’ ที่แนะนำการใช้โปรแกรม และช่วยให้ผู้ใช้สามารถอับเดทโปรแกรมได้ รวมทั้งสามารถติดต่อสอบถามปัญหาหรือแจ้งปัญหาการใช้ทางอีเมล์ได้โดยตรงที่ email address ต่อไปนี้ gpower-feedback@uni-duesseldorf.de. (Buchner, 2010)

จากเอกสารของ Faul, Erdfelder, Lang & Buchner (2007); Faul, Erdfelder, Buchner & Lang (2009); Buchner, (2010) สรุปได้ว่า โปรแกรม G*Power 3.1 สามารถรองรับการวิเคราะห์อำนาจการทดสอบ

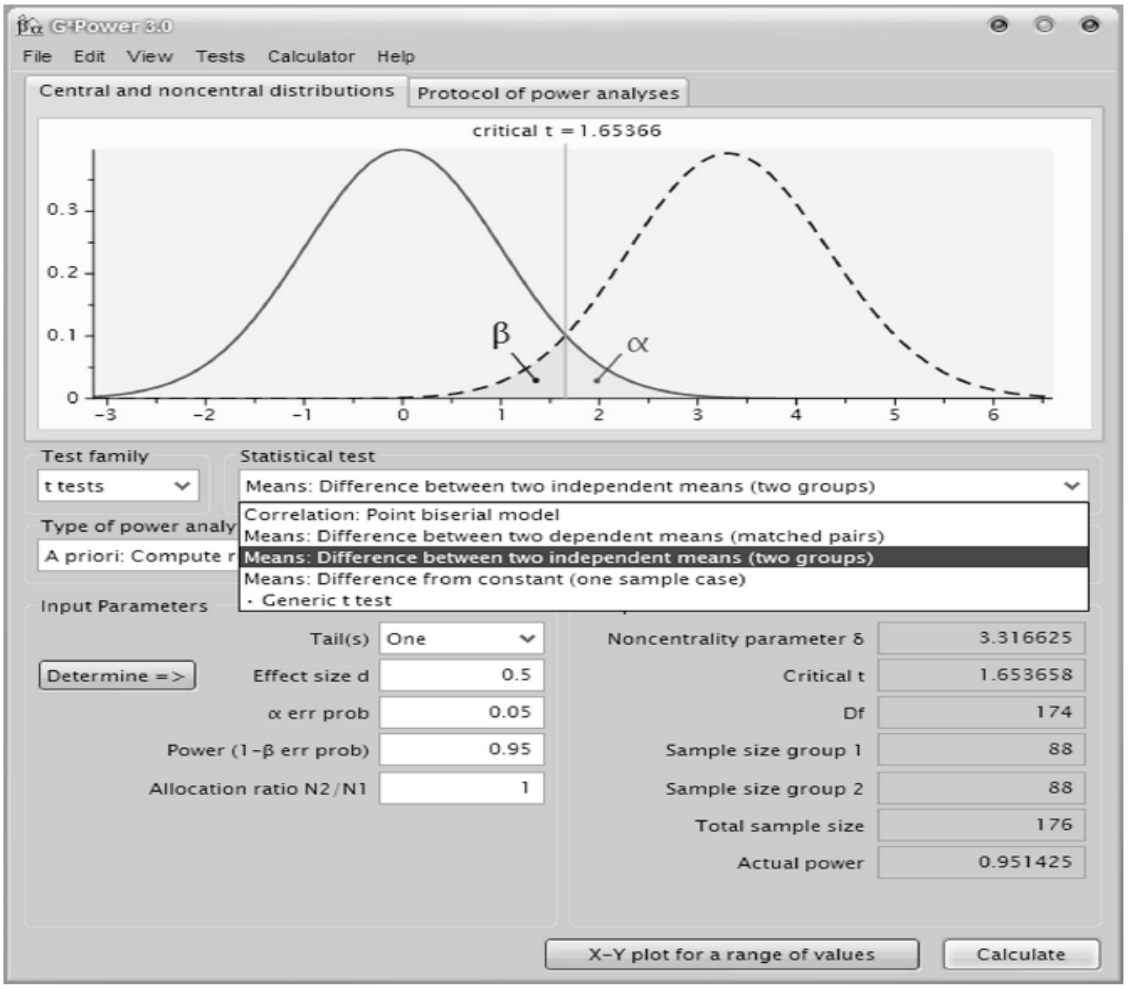
ได้ถึง 5 แบบ ตามที่ Cohen (1977, 1988) ระบุไว้ ในขณะที่โปรแกรม G*Power 2 ทำได้เพียง 3 แบบแรก การวิเคราะห์ทั้ง 5 แบบ คือ 1) การวิเคราะห์อำนาจการทดสอบก่อนการวิจัย (priori power analysis) เป็นการกำหนดขนาดตัวอย่างจากค่าอำนาจการทดสอบที่นักวิจัยต้องการ 2) การวิเคราะห์อำนาจการทดสอบหลังการวิจัย (post hoc power analysis) เป็นการประมาณค่าอำนาจการทดสอบจากขนาดตัวอย่างที่นักวิจัยใช้ในการวิจัย ให้ผลการวิเคราะห์ถูกต้องมากกว่าวิธี retrospective power analysis ที่ใช้ในโปรแกรม SPSS และโปรแกรมอื่นๆ ซึ่งมีข้อตกลงเบื้องต้นที่เป็นปัญหาหลายข้อ 3) การวิเคราะห์อำนาจการทดสอบแบบประนีประนอม (compromise power analysis) เป็นการวิเคราะห์ที่ทำได้ทั้งก่อนและหลังการวิจัย โดยมีการกำหนดอัตราส่วนความน่าจะเป็นของความคลาดเคลื่อน (error probability ratio) = q = α/β กรณีการใช้ q = 1 แสดงว่านักวิจัยประกันการทดสอบโดยยอมให้เกิดความคลาดเคลื่อนในการทดสอบทั้งสองประเภทมีค่าเท่ากัน วิธีนี้มีประโยชน์เมื่อผลการวิเคราะห์อำนาจการทดสอบก่อนการวิจัยได้ขนาดตัวอย่างที่มีจำนวนสูงกว่าประชากรที่มีอยู่จริง นักวิจัยอาจใช้การวิเคราะห์อำนาจการทดสอบแบบประนีประนอม โดยกำหนดขนาดตัวอย่างสูงสุดตามที่เป็นจริง และกำหนดค่า q = 4 ซึ่งเป็นการยอมให้มีความคลาดเคลื่อนประเภทที่หนึ่งสูงมากเมื่อเทียบกับประเภทที่สองเนื่องจากขนาดตัวอย่าง 4) การวิเคราะห์ความไว (sensitivity analysis) เป็นการวิเคราะห์ขนาดอิทธิพล (effect size) ที่เป็นจุดวิกฤติในรูปฟังก์ชั่นของขนาดตัวอย่าง ค่าอำนาจการทดสอบและระดับนัยสำคัญที่นักวิจัยกำหนด ซึ่งรับประกันได้ว่าผลการวิจัยมีอำนาจการทดสอบตามที่นักวิจัยต้องการ และ 5) การวิเคราะห์เกณฑ์ (criterion analysis) ซึ่งเป็นทางเลือกหนึ่งของการวิเคราะห์อำนาจการทดสอบหลังการวิจัย เมื่อนักวิจัยเห็นความสำคัญของระดับนัยสำคัญน้อยกว่าความสำคัญของอำนาจการทดสอบ เช่นในการทดสอบความกลมกลืน (goodness of fit test)

เนื่องจากนักวิจัยส่วนใหญ่ใช้ต้องการกำหนดขนาดตัวอย่างที่เหมาะสมก่อนการวิจัย และใช้การวิเคราะห์อำนาจการทดสอบก่อนการวิจัย (priori power analysis) กันมาก ในเอกสารฉบับนี้ผู้เขียนจึงเสนอวิธีการประมาณค่าขนาดตัวอย่างเฉพาะการวิเคราะห์อำนาจการทดสอบก่อนการวิจัย ส่วนการวิเคราะห์แบบอื่นมีหลักการคล่ายกัน ผู้อ่านที่สนใจอาจศึกษาได้จากคู่มือการใช้โปรแกรมของ Buchner (2010) ได้

3.2 การใช้โปรแกรม G*Power

การใช้โปรแกรม G*Power 3 และ 3.1 (Faul, Erdfelder, Lang & Buchner, 2007; Faul, Erdfelder, Buchner & Lang, 2009; Buchner, 2010) มี 4 ขั้นตอน คือ 1) เลือกวิธีการวิเคราะห์ทางสถิติที่เหมาะสมกับปัญหาวิจัยที่ต้องการจากเมนูหลัก ซึ่งมีสถิติให้เลือกใช้ 5 กลุ่ม (ได้แก่ t-test, z-test, F-test, χ^2- test, และ exact test) แต่ละกลุ่มมีสถิติทดสอบอีกหลายประเภท เช่น กลุ่ม t-test มีสถิติทดสอบในโปรแกรม G*Power ถึง 12 ประเภท 2) เลือกวิธีการวิเคราะห์อำนาจการทดสอบหนึ่งในห้าแบบข้างต้น เช่น เลือก t-independent samples จากกลุ่ม t-test 3) ป้อนข้อมูลสำหรับการวิเคราะห์ขนาดตัวอย่างตามที่กำหนดไว้ และ 4) คลิกปุ่ม “Calculate” ที่หน้าต่างโปรแกรมเพื่อให้ได้ผลลัพธ์ที่ต้องการ

เมื่อเสร็จสิ้นการใช้งาน โปรแกรมให้หน้าจอแสดงผลการวิเคราะห์ที่ต้องการ และภาพแสดงความคลาดเคลื่อนทั้งสองประเภทดังภาพต่อไปนี้



ภาพ 6 หน้าจอแสดงผลการใช้โปรแกรม G*Power 3

ที่มา: Buchner (2010)

ขั้นตอนการใช้โปรแกรม G*Power ทั้ง 4 ขั้นตอนนี้ ขั้นตอนที่ 3 คือ การป้อนข้อมูลสำหรับการวิเคราะห์ขนาดตัวอย่างตามที่กำหนดไว้รวม 4 ค่า ได้แก่ 1) ค่าขนาดอิทธิพล (effect size) 2) ความน่าจะเป็นของความคลาดเคลื่อนในการทดสอบประเภทที่หนึ่ง = α 3) อำนาจการทดสอบ = (1 – ความน่าจะเป็นของความคลาดเคลื่อนในการทดสอบประเภทที่สอง) = (1-β) และ 4) อัตราส่วนการจัดสรรขนาดตัวอย่างทั้งสองกลุ่ม (allocation ratio) = N2/N1 ซึ่งนิยมกำหนดให้มีขนาดตัวอย่างสองกลุ่มเท่ากัน โดย N2/N1 = 1

การกำหนดค่าขนาดอิทธิพลทำได้ 2 วิธี คือ ก) การป้อนข้อมูลดิบโดยตรงและให้โปรแกรมประมาณค่าขนาดอิทธิพลให้ และ ข) การป้อนข้อมูลขนาดอิทธิพลโดยตรง วิธีนี้นักวิจัยต้องมีข้อค้นพบจากงานวิจัยในอดีต รู้ค่าขนาดอิทธิพล และป้อนข้อมูลโดยใส่ค่าขนาดอิทธิพลที่ทราบ หรือใส่ค่าประมาณขนาดอิทธิพลที่ Cohen (1977) กำหนดไว้แยกเป็น 3 ระดับ คือ ขนาดเล็ก ปานกลาง และใหญ่ ตามประเภทของค่าขนาดอิทธิพลที่สำคัญรวม 6 ประเภท คือ 1) ขนาดอิทธิพลสำหรับการทดสอบเปรียบเทียบค่าเฉลี่ยสองกลุ่ม = d 2) ขนาดอิทธิพลสำหรับการทดสอบเปรียบเทียบค่าเฉลี่ยมากกว่าสองกลุ่ม = f 3) ขนาดอิทธิพลสำหรับการทดสอบไค-สแควร์ = w 4) ขนาดอิทธิพลสำหรับการทดสอบสัมประสิทธิ์สหสัมพันธ์ = ρ 5) ขนาด

อิทธิพลสำหรับการทดสอบเปรียบเทียบสัมประสิทธิ์สหสัมพันธ์ = q และ 6) ขนาดอิทธิพลสำหรับการทดสอบสัมประสิทธิ์สหสัมพันธ์พหุคูณในการวิเคราะห์ถดถอยเชิงเส้นตรง = f^2 ดังตารางต่อไปนี้

ตาราง 1 ค่าขนาดอิทธิพล 3 ระดับ จำแนกตามประเภทของค่าขนาดอิทธิพล 4 ประเภท

Test	Effect size	Small	Medium	Large
1. Difference between two means	$d = (\mu_1 - \mu_2)/\sigma$	0.20	0.50	0.80
2. Difference between many means	$f = \sqrt{(\eta^2 /(1 - \eta^2))}$	0.10	0.25	0.40
หมายเหตุ: $f = \sigma_m/\sigma$ และ 'SPSS η^{2}' $= \eta^2 N/(N + k(\eta^2 - 1))$				
3. Chi-squared test	$w = \sqrt{\sum_{i=1}^{k} \frac{(\pi_{1i} - \pi_{0i})^2}{\pi_{0i}}}$	0.10	0.30	0.50
4. Pearson's correlation coefficient	ρ or r	0.10	0.30	0.50
5. Difference between correlation coeff.	q = Fisher's z_1- Fisher's z_2	0.10	0.30	0.50
6. Linear multiple correlation coefficient	$f^2 = R^2_{Y \cdot B}/(1 - R^2_{Y \cdot B})$	0.02	0.15	0.35

ที่มา: Buchner (2010); Cohen (1977)

3.3 ตัวอย่างการใช้โปรแกรม G*Power

การเสนอตัวอย่างในตอนนี้ เสนอเฉพาะตัวอย่างการวิเคราะห์ข้อมูลที่นิยมใช้กันมาก 3 ตัวอย่าง ดังนี้

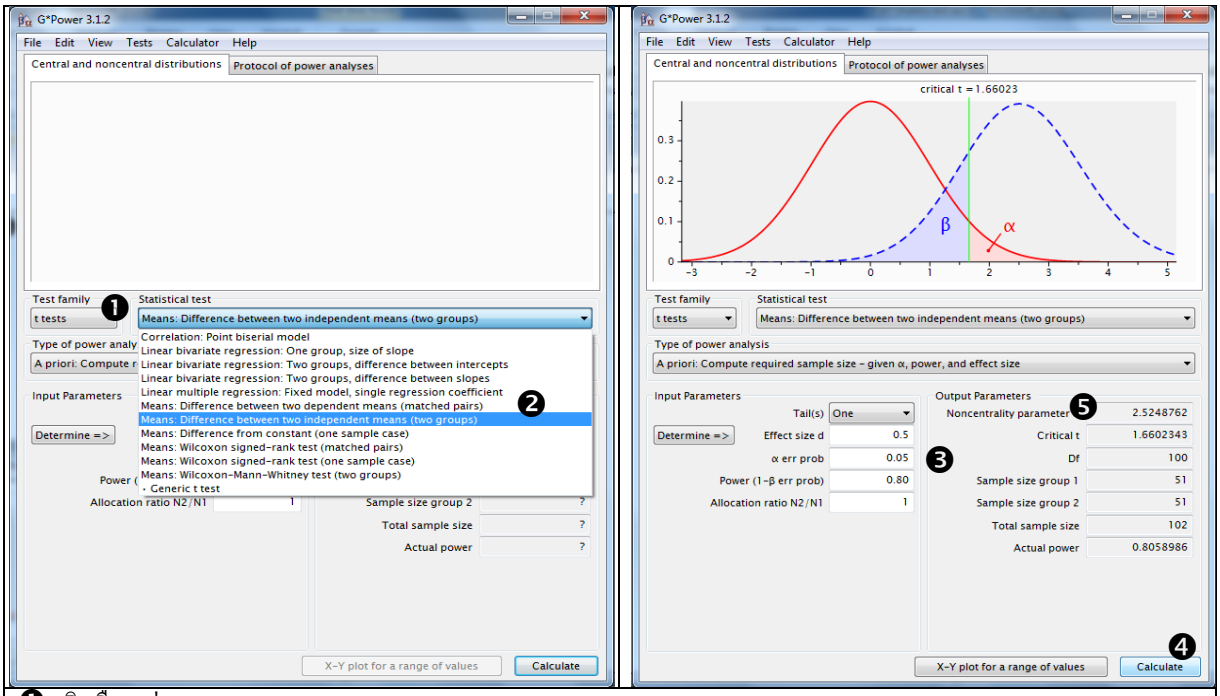
ตัวอย่าง 1 การประมาณค่าขนาดตัวอย่างกรณี t-test independent samples

โจทย์ นักวิจัยวางแผนการทดลองให้มีตัวอย่างสองกลุ่ม มีขนาดตัวอย่างเท่ากัน เพื่อทดสอบอิทธิพลของวิธีการเรียนการสอนสองวิธี ที่มีต่อผลสัมฤทธิ์ทางการเรียน โดยใช้การทดสอบแบบทางเดียว และกำหนดให้มีค่า α =.05 และ (1-β) = .80 เมื่อมีข้อมูลสองแบบ **แบบที่หนึ่ง** รู้ค่าขนาดอิทธิพล ว่ามีระดับปานกลาง = 0.50 **แบบที่สอง** ไม่รู้ค่าขนาดอิทธิพล แต่รู้ค่าสถิติจากผลการวิจัยในอดีตว่าค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน กลุ่มทดลอง = 32.50 และ 13.52 คะแนน กลุ่มควบคุม = 25.90 และ 12.97 คะแนน ตามลำดับ

วิธีการประมาณค่าขนาดตัวอย่าง

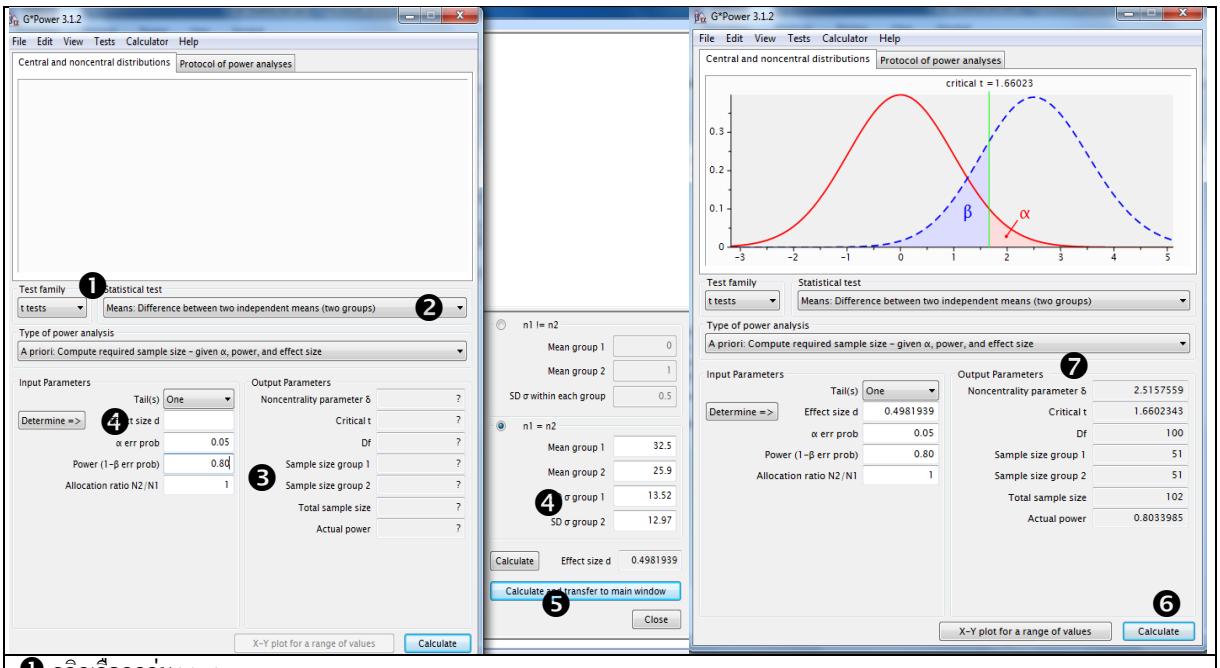
แบบที่หนึ่ง รู้ค่าขนาดอิทธิพล นักวิจัยประมาณค่าขนาดตัวอย่างแต่ละกลุ่มตามขั้นตอนดังนี้ 1) เลือกกลุ่มการทดสอบ 2) เลือกประเภทของการทดสอบ 3) ป้อนข้อมูล 4) ประมาณค่าขนาดตัวอย่าง และ 5) อ่านผลการวิเคราะห์จากโปรแกรม G*Power ดังผลการประมาณค่าขนาดตัวอย่าง การเปรียบเทียบค่าเฉลี่ย กรณีตัวอย่างเป็นกลุ่มอิสระไม่สัมพันธ์กัน แบบรู้ค่าขนาดอิทธิพล ภาพ 7

แบบที่สอง ไม่รู้ค่าขนาดอิทธิพล แต่รู้ค่าสถิติ นักวิจัยประมาณค่าขนาดตัวอย่างแต่ละกลุ่มตามขั้นตอนดังนี้ 1) เลือกกลุ่มการทดสอบ 2) เลือกประเภทของการทดสอบ 3) ป้อนข้อมูลแบบเดียวกับแบบที่หนึ่ง ยกเว้นค่าขนาดอิทธิพล ซึ่งโปรแกรมคำนวณให้ 4) ประมาณค่าขนาดอิทธิพลโดยการป้อนค่าสถิติ ได้แก่ ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของตัวอย่างทั้งสองกลุ่ม ผลการประมาณค่าขนาดอิทธิพลซึ่งได้มาจากการศึกษาเอกสารและงานวิจัยในอดีต และอัตราส่วนการจัดสรรขนาดตัวอย่างทั้งสองกลุ่ม (allocation ratio) = N2/N1 5) ประมาณค่าขนาดตัวอย่าง และ 6) อ่านผลการวิเคราะห์จากโปรแกรม G*Power ดังภาพการประมาณค่าขนาดตัวอย่าง ในการเปรียบเทียบค่าเฉลี่ยกรณีตัวอย่างเป็นกลุ่มอิสระ แบบไม่รู้ค่าขนาดอิทธิพล ดังภาพ 8 จะเห็นได้ว่าผลจากการวิเคราะห์จะได้ค่าใกล้เคียงกันเพราะข้อมูลทั้งสองแบบมาจากฐานข้อมูลวิจัยเดียวกัน



❶ คลิกเลือกกลุ่ม t-test
❷ คลิกเลือก Means: Difference between two independent means (two groups)
❸ เลือกรูปแบบการทดสอบ tailed = One และพิมพ์ป้อนข้อมูล ค่า d = 0.50, α =.05, (1-β) = .80 และ N2/N1 = 1
❹ คลิกปุ่ม ‘Calculate
❺ อ่านผลการวิเคราะห์จาก ‘Output Parameters’ ได้ขนาดตัวอย่างกลุ่มละ 51 คน รวมสองกลุ่ม = 102 คน

ภาพ 7 การประมาณค่าขนาดตัวอย่าง การเปรียบเทียบค่าเฉลี่ยกรณีตัวอย่างกลุ่มอิสระ แบบรู้ค่าขนาดอิทธิพล



❶ คลิกเลือกกลุ่ม t-test
❷ คลิกเลือก Means: Difference between two independent means (two groups)
❸ เลือกรูปแบบการทดสอบ tailed = One และพิมพ์ป้อนข้อมูล ค่า α =.05, (1-β) = .80 และ N2/N1 = 1 ยกเว้นค่า d ซึ่งไม่รู้ค่า
❹ คลิกปุ่ม ‘Determine’ เพื่อเปิดหน้าต่างให้โปรแกรมคำนวณค่า d และป้อนข้อมูล
❺ คลิกปุ่ม’Calculate and transfer to main window’
❻ คลิกปุ่ม ‘Calculate
❼ อ่านผลการวิเคราะห์จาก ‘Output Parameters’ ได้ขนาดตัวอย่างกลุ่มละ 51 คน รวมสองกลุ่ม = 102 คน

ภาพ 8 การประมาณค่าขนาดตัวอย่าง การเปรียบเทียบค่าเฉลี่ยกรณีตัวอย่างอิสระ แบบไม่รู้ค่าขนาดอิทธิพล

ตัวอย่าง 2 การประมาณค่าขนาดตัวอย่างกรณีการวิเคราะห์ความแปรปรวน

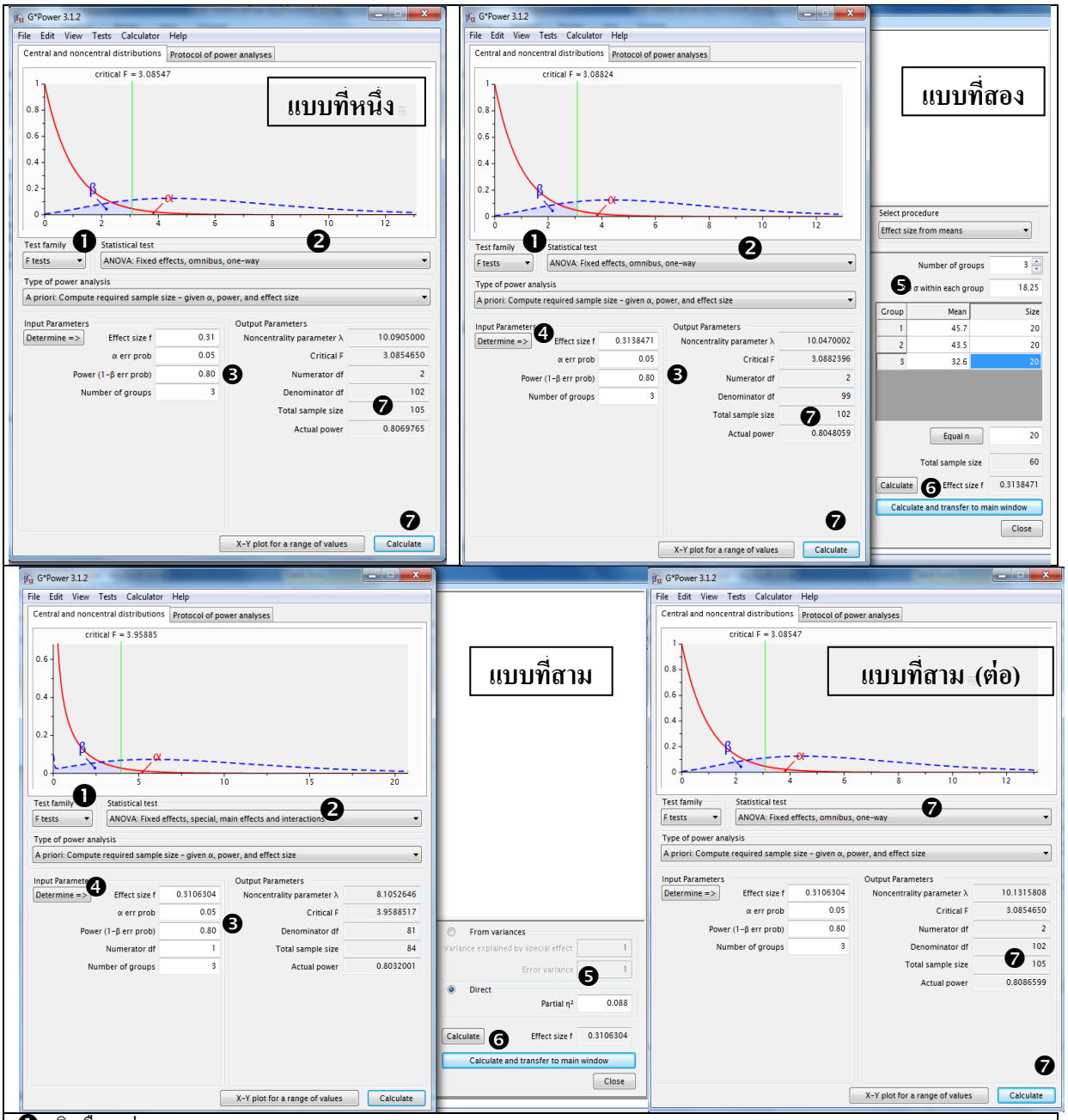
การประมาณค่าขนาดตัวอย่าง สำหรับการวิเคราะห์ความแปรปรวน แบ่งเป็นสองกรณี คือ **1) กรณีรู้ค่าขนาดอิทธิพล** $f = \sqrt{(\eta^2 /(1 - \eta^2))}$ หรือประมาณค่าได้ว่าขนาดอิทธิพลต่ำ (.10) ปานกลาง (.25) หรือสูง (.40) นักวิจัยประมาณค่าขนาดตัวอย่างแยกเป็น 2 แบบ คือ 1.1) การวิเคราะห์ความแปรปรวนแบบทางเดียว เลือกใช้การทดสอบ ANOVA: fixed effect, omnibus, oneway แล้วป้อนข้อมูลค่าขนาดอิทธิพล (f), α, (1-β), และจำนวนกลุ่มที่ต้องการ และ 1.2) การวิเคราะห์ความแปรปรวนแบบหลายทาง เลือกใช้การทดสอบ ANOVA: fixed effects, special main effects and interaction แล้วป้อนข้อมูลค่าขนาดอิทธิพล (f), α, (1-β), df ของตัวเศษ (numerater df) และจำนวนกลุ่มที่ต้องการ สำหรับการทดสอบแต่ละชุด **2) กรณีไม่รู้ค่าขนาดอิทธิพล** นักวิจัยให้โปรแกรมคำนวณค่าขนาดอิทธิพลให้ ทำได้เป็น 2 วิธี คือ 2.1) รู้ค่าเฉลี่ย ขนาดตัวอย่างของตัวอย่างแต่ละกลุ่ม และส่วนเบี่ยงเบนมาตรฐานภายในกลุ่ม (withing group sd) ทำได้โดยเลือกใช้การทดสอบ ANOVA: fixed effect, omnibus, oneway ANOVA: fixed effects, special main effects and interaction และ 2.2) รู้ค่า η^2 ทำได้โดยเลือกใช้การทดสอบ ANOVA: fixed effects, special main effects and interaction ในที่นี้แสดงตัวอย่างการประมาณค่าขนาดตัวอย่างการวิเคราะห์ความแปรปรวนแบบทางเดียวกรณีรู้ค่าและไม่รู้ค่าขนาดอิทธิพล ดังนี้

โจทย์ นักวิจัยวางแผนการทดลองให้มีกลุ่มตัวอย่างสามกลุ่ม มีขนาดตัวอย่างเท่ากัน เพื่อทดสอบอิทธิพลของวิธีการเรียนการสอนสามวิธี ที่มีต่อผลสัมฤทธิ์ทางการเรียน โดยใช้การทดสอบแบบทางเดียว และกำหนดให้มีค่า α =.05 และ (1-β) = .80 เมื่อมีข้อมูลสองแบบ **แบบที่หนึ่ง** รู้ค่าขนาดอิทธิพลว่าอยู่ในระดับค่อนข้างสูง (f = 0.31) เมื่อ $f = \sqrt{(\eta^2 /(1 - \eta^2))}$ **แบบที่สอง** ไม่รู้ค่าขนาดอิทธิพล แต่รู้ค่าสถิติจากผลการวิจัยในอดีตว่าตัวอย่างทั้งสามกลุ่มมีค่าเฉลี่ย = 45.7, 43.5, 32.6 ขนาดตัวอย่าง = 20, 20, 20 และส่วนเบี่ยงเบนมาตรฐานภายในกลุ่ม (withing group sd) = 18.25 และ**แบบที่สาม**ไม่รู้ค่าขนาดอิทธิพล แต่รู้ค่าสถิติจากผลการวิจัยในอดีตว่า $\eta^2 = 0.088$

วิธีการประมาณค่าขนาดตัวอย่าง

แบบที่หนึ่ง รู้ค่าขนาดอิทธิพล นักวิจัยประมาณค่าขนาดตัวอย่างแต่ละกลุ่มตามขั้นตอนดังนี้ 1) เลือกกลุ่มการทดสอบ F-test 2) เลือกประเภทของการทดสอบ ANOVA: fixed effect, omnibus, oneway 3) ป้อนข้อมูล 4) ประมาณค่าขนาดตัวอย่าง และ 5) อ่านผลการวิเคราะห์จากโปรแกรม G*Power ดัง ภาพ 9

แบบที่สองไม่รู้ค่าขนาดอิทธิพล แต่รู้ค่าสถิติจากผลการวิจัยในอดีตว่าค่าเฉลี่ย ขนาดตัวอย่าง และ ส่วนเบี่ยงเบนมาตรฐานภายในกลุ่ม นักวิจัยประมาณค่าขนาดอิทธิพลก่อนโดยใช้โปรแกรม แล้วจึงประมาณค่าขนาดตัวอย่างตามขั้นตอนดังนี้ 1) เลือกกลุ่มการทดสอบ F-test 2) เลือกประเภทของการทดสอบแบบ ANOVA: fixed effect, omnibus, oneway 3) ป้อนข้อมูลแบบเดียวกับแบบที่หนึ่ง ยกเว้นค่าขนาดอิทธิพล ซึ่งจะให้โปรแกรมคำนวณให้ 4) ประมาณค่าขนาดอิทธิพลโดยการป้อนค่าสถิติได้แก่ค่า η^2 และป้อนผลการประมาณค่าขนาดอิทธิพลใส่โปรแกรม 5) ประมาณค่าขนาดตัวอย่าง และ 6) อ่านผลการวิเคราะห์จากโปรแกรม G*Power ดังแสดงในภาพต่อไปนี้



❶ คลิกเลือกกลุ่ม F-test

❷ คลิกเลือกประเภทการทดสอบ

<u>แบบที่หนึ่ง และสอง</u> เลือก ANOVA: fixed effect, omnibus, oneway

<u>แบบที่สาม</u> เลือก ANOVA: fixed effects, special main effects and interaction

❸ พิมพ์ป้อนข้อมูลที่ทราบ ค่า α =.05, (1-β) = .80 และ

<u>แบบที่หนึ่ง</u> ป้อนข้อมูล ค่าขนาดอิทธิพล f

<u>แบบที่สอง และสาม</u> ไม่ป้อนข้อมูลค่าขนาดอิทธิพล f

❹ คลิกปุ่ม 'Determine' เพื่อเปิดหน้าต่างให้โปรแกรมคำนวณค่าขนาดอิทธิพล f

❺ ป้อนข้อมูลให้โปรแกรมคำนวณค่าขนาดอิทธิพล

<u>แบบที่สอง</u> ป้อนข้อมูล ค่าเฉลี่ย = 45.7, 43.5, 32.6 ขนาดตัวอย่าง = 20, 20, 20 และ withing group sd = 18.25

<u>แบบที่สาม</u> ป้อนข้อมูล η^2 = 0.088

❻ คลิกปุ่ม'Calculate and transfer to main window'

❼ การประมาณค่าขนาดตัวอย่าง

<u>แบบที่หนึ่งและสอง</u> คลิกปุ่ม 'Calculate' และอ่านผลการวิเคราะห์จาก' Output Parameters' ได้ขนาดตัวอย่าง = 105, 102

<u>แบบที่สาม</u> เลือก ANOVA: fixed effect, omnibus, oneway ก่อน แล้วคลิกปุ่ม 'Calculate' อ่านผลการวิเคราะห์ได้ขนาดตัวอย่าง = 105

ภาพ 9 การประมาณค่าขนาดตัวอย่าง การวิเคราะห์ความแปรปรวนแบบทางเดียว แบบรู้และไม่รู้ค่า f

แบบที่สามไม่รู้ค่าขนาดอิทธิพล แต่รู้ค่าสถิติจากผลการวิจัยในอดีตว่า $\eta^2 = 0.088$ นักวิจัยประมาณค่าขนาดอิทธิพลก่อนโดยใช้โปรแกรม แล้วจึงประมาณค่าขนาดตัวอย่างตามขั้นตอนดังนี้ 1) เลือกกลุ่มการทดสอบ F-test 2) เลือกประเภทของการทดสอบแบบ ANOVA: fixed effects, special main effects and interaction 3) ป้อนข้อมูลแบบเดียวกับแบบที่หนึ่ง ยกเว้นค่าขนาดอิทธิพล ซึ่งโปรแกรมคำนวณให้ 4) ประมาณค่าขนาดอิทธิพลโดยการป้อนค่าสถิติได้แก่ค่า $\eta^2 = 0.088$ และป้อนผลการประมาณค่าขนาดอิทธิพล 5) ประมาณค่าขนาดตัวอย่าง และ 6) อ่านผลการวิเคราะห์จากโปรแกรม G*Power ดังแสดงในภาพ 9 จะเห็นได้ว่าผลจากการวิเคราะห์ทั้งสามแบบจะได้ค่าใกล้คียงกัน เพราะข้อมูลทั้งสองแบบมาจากฐานข้อมูลวิจัยเดียวกัน

ตัวอย่าง 3 การประมาณค่าขนาดตัวอย่างกรณีการวิเคราะห์ถดถอยพหุคูณ

การประมาณค่าขนาดตัวอย่าง สำหรับการวิเคราะห์ความแปรปรวน แบ่งเป็นสองกรณี คือ **1) กรณีรู้ค่าขนาดอิทธิพล** $f^2 = R^2_{Y \cdot B}/(1 - R^2_{Y \cdot B})$ หรือประมาณค่าได้ว่าขนาดอิทธิพล f^2 มีค่าต่ำ (.02) ปานกลาง (.15) หรือสูง (.35) นักวิจัยประมาณค่าขนาดกลุ่มตัวอย่าง โดยเลือกใช้การทดสอบ ANOVA: fixed effects, special main effects and interaction แล้วป้อนข้อมูลค่าขนาดอิทธิพล (f), α, $(1-\beta)$, df ของตัวเศษ (numerater df), df ของตัวส่วน (denominater df) และจำนวนกลุ่มตัวอย่าง สำหรับการทดสอบแต่ละชุด **2) กรณีไม่รู้ค่าขนาดอิทธิพล** แต่รู้ค่าสถิติ R^2 นักวิจัยให้โปรแกรมคำนวณค่าขนาดอิทธิพล f^2 ให้ และนำไปใช้ในการประมาณค่าขนาดตัวอย่างตามกรณีรู้ค่าขนาดอิทธิพล ดังตัวอย่างต่อไปนี้

โจทย์ นักวิจัยต้องการประมาณค่าขนาดตัวอย่างสำหรับการศึกษาปัจจัย (factors) รวม 5 ตัวแปร คือ เจตคติ (X1) แรงจูงใจใฝ่สัมฤทธิ์ (X2) ความเชื่อประสิทธิภาพในตน (X3) ความคาดหวัง (X4) และเป้าหมายการทำงาน (X5) ที่มีอิทธิพลต่อพฤติกรรมการทำงาน (Y) โดยใช้การวิเคราะห์ถดถอยพหุคูณ และกำหนดให้มีค่า α =.05 และ $(1-\beta)$ = .90 เมื่อมีข้อมูลสองแบบ **แบบที่หนึ่ง** รู้ค่าขนาดอิทธิพล f^2 = 0.204 **แบบที่สอง** ไม่รู้ค่าขนาดอิทธิพล แต่รู้ค่าสถิติจากผลการวิจัยในอดีตว่า ค่า R^2 = .17 และ**แบบที่สาม**ไม่รู้ค่าขนาดอิทธิพล แต่รู้ค่าเมทริกซ์สหสัมพันธ์ของตัวแปรตามภาพขวามือ

	X1	X2	X3	X4	X5	Y
X1	1.0					
X2	.30	1.0				
X3	.40	.50	1.0			
X4	.20	.20	.30	1.0		
X5	.10	.30	.20	.20	1.0	
Y	.34	.32	.20	.09	.12	1.0

วิธีการประมาณค่าขนาดตัวอย่าง

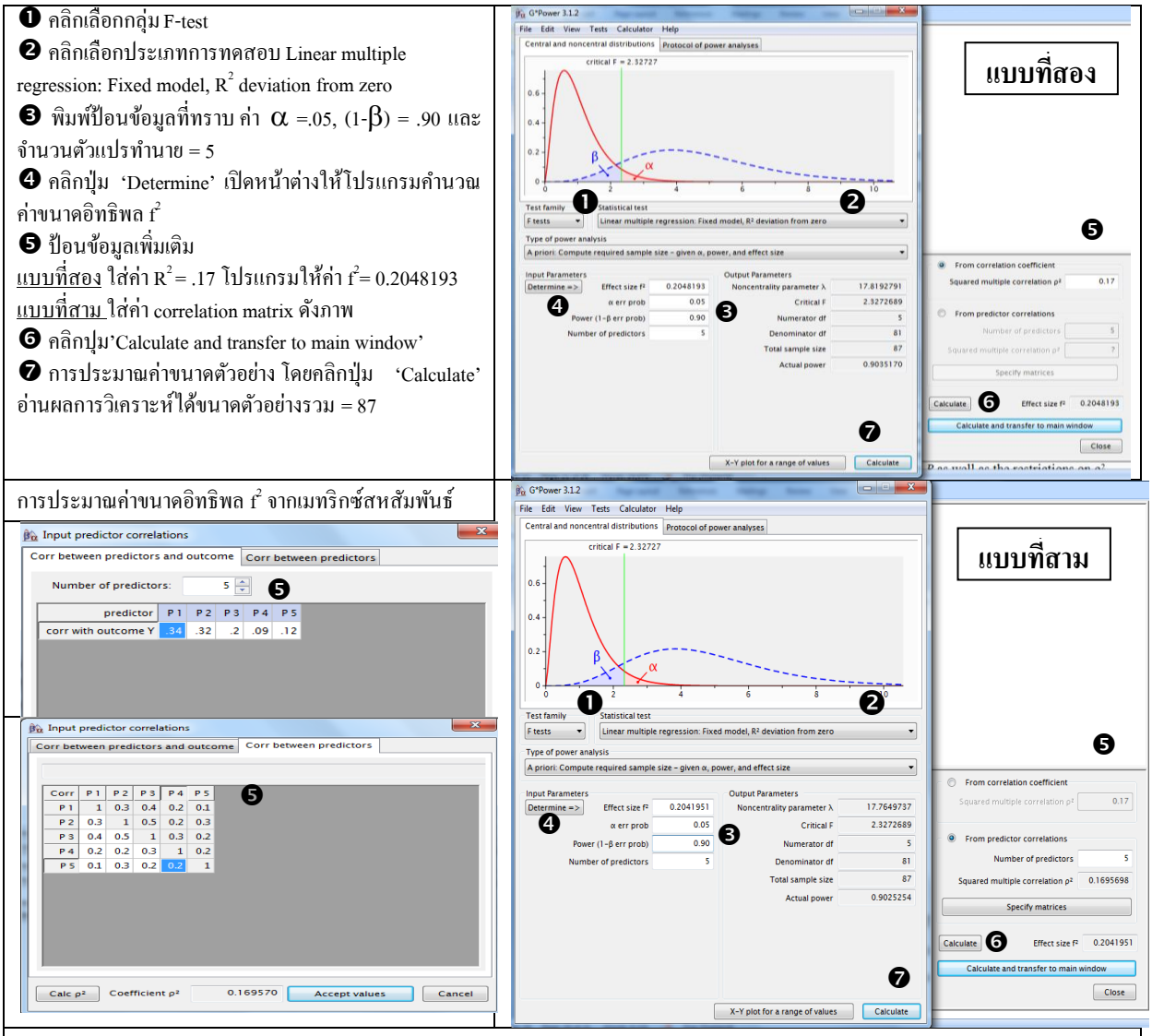
แบบที่หนึ่ง รู้ค่าขนาดอิทธิพล นักวิจัยประมาณค่าขนาดตัวอย่างแต่ละกลุ่มตามขั้นตอนดังนี้ 1) เลือกกลุ่มการทดสอบ F-test 2) เลือกประเภทของการทดสอบ Linear multiple regression: Fixed model, R^2 deviation from zero 3) ป้อนข้อมูล α =.05; $(1-\beta)$ = .90; จำนวนตัวแปรทำนาย = 5 และค่าขนาดอิทธิพล f^2 = 0.204 4) ประมาณค่าขนาดตัวอย่าง โดยคลิกปุ่ม ‘Calculate’ และ 5) อ่านผลการวิเคราะห์จาก Output parameters ได้ขนาดตัวอย่าง = 87

แบบที่สองไม่รู้ค่าขนาดอิทธิพล แต่รู้ค่าสถิติจากผลการวิจัยในอดีตว่าค่า R^2= .17 นักวิจัยประมาณค่าขนาดอิทธิพลก่อนโดยใช้โปรแกรม แล้วจึงประมาณค่าขนาดตัวอย่างตามขั้นตอนดังนี้ 1) เลือกกลุ่มการทดสอบ F-test 2) เลือกประเภทของการทดสอบแบบ Linear multiple regression: Fixed model, R^2 deviation from

zero 3) ป้อนข้อมูลแบบเดียวกับแบบที่หนึ่ง ยกเว้นค่าขนาดอิทธิพล ซึ่งโปรแกรมคำนวณให้ 4) ประมาณค่าขนาดอิทธิพลแบบใช้ตัวเลือก 'From correlation coefficient' โดยป้อนค่าสถิติได้แก่ค่า R^2 = .17 และป้อนผลการประมาณค่าขนาดอิทธิพลใส่โปรแกรม ซึ่งได้ค่า = 0.2048193 5) ประมาณค่าขนาดตัวอย่างโดยคลิก 'Calculate' และ 6) อ่านผลการวิเคราะห์จาก Output parameters ได้ขนาดตัวอย่าง = 87 ดังภาพ 10

แบบที่สอง ไม่รู้ค่าขนาดอิทธิพล แต่รู้ค่าเมทริกซ์สหสัมพันธ์ระหว่างตัวแปรจากผลการวิจัยในอดีต นักวิจัยประมาณค่าขนาดอิทธิพลก่อนโดยใช้โปรแกรม แล้วจึงประมาณค่าขนาดตัวอย่างตามขั้นตอนดังนี้ ขั้นตอน 1-3 ดำเนินการเหมือนแบบที่สอง 4) ประมาณค่าขนาดอิทธิพลแบบใช้ตัวเลือก 'From predictors correlation' โดยป้อนค่าสัมประสิทธ์สหสัมพันธ์ และป้อนผลการประมาณค่าขนาดอิทธิพลใส่โปรแกรม ซึ่งได้ค่า f^2 = 0.1695698 5) ประมาณค่าขนาดตัวอย่างโดยคลิก 'Calculate' และ 6) อ่านผลการวิเคราะห์จาก Output parameters ได้ขนาดตัวอย่าง = 87 ดังภาพ 10 จะเห็นได้ว่าผลจากการวิเคราะห์ทั้งสามแบบจะได้ค่าใกล้เคียงกัน เพราะข้อมูลทั้งสองแบบมาจากฐานข้อมูลวิจัยเดียวกัน

❶ คลิกเลือกกลุ่ม F-test
❷ คลิกเลือกประเภทการทดสอบ Linear multiple regression: Fixed model, R^2 deviation from zero
❸ พิมพ์ป้อนข้อมูลที่ทราบ ค่า α =.05, (1-β) = .90 และจำนวนตัวแปรทำนาย = 5
❹ คลิกปุ่ม 'Determine' เปิดหน้าต่างให้โปรแกรมคำนวณค่าขนาดอิทธิพล f^2
❺ ป้อนข้อมูลเพิ่มเติม
แบบที่สอง ใส่ค่า R^2 = .17 โปรแกรมให้ค่า f^2= 0.2048193
แบบที่สาม ใส่ค่า correlation matrix ดังภาพ
❻ คลิกปุ่ม'Calculate and transfer to main window'
❼ การประมาณค่าขนาดตัวอย่าง โดยคลิกปุ่ม 'Calculate' อ่านผลการวิเคราะห์ได้ขนาดตัวอย่างรวม = 87



ภาพ 10 การประมาณค่าขนาดตัวอย่าง การวิเคราะห์ถดถอยพหุคูณ แบบไม่รู้ค่าขนาดอิทธิพล 2 แบบ

ความส่งท้าย

บทความนี้มุ่งนำเสนอสาระรวม 3 เรื่อง คือ ความรู้พื้นฐานเกี่ยวกับการกำหนดขนาดตัวอย่าง วิธีการกำหนดขนาดตัวอย่างจากตารางสำเร็จ และตัวอย่างการกำหนดขนาดตัวอย่างโดยใช้โปรแกรมคอมพิวเตอร์ G*Power ซึ่งเป็นโปรแกรมคอมพิวเตอร์ให้ใช้ฟรี ที่ได้รับความนิยมสูงมากในวงการวิจัยทางกลุ่มประเทศยุโรปและอเมริกา โดยมีวัตถุประสงค์ให้นักวิจัยไทยได้ใช้วิธีการกำหนดขนาดตัวอย่างที่ทันสมัย และถูกต้องสมบูรณ์ ทัดเทียมกับงานวิจัยระดับนานาชาติ การใช้โปรแกรม G*Power มีประเด็นที่นักวิจัยควรปฏิบัติ 3 ประการ คือ **ประการที่ 1 ควรมีการปรับขนาดตัวอย่างให้เหมาะสมมากขึ้น** เมื่อนักวิจัยได้ผลลัพธ์จากการประมาณค่าขนาดตัวอย่างด้วยโปรแกรมคอมพิวเตอร์แล้ว นักวิจัยควรต้องมีการปรับตัวเลขขนาดตัวอย่างให้เหมาะสมกับงานวิจัยของตนด้วย เช่น การปรับให้เป็นตัวเลขกลมๆ การปรับเพิ่มขนาดตัวอย่างเพื่อชดเชยอัตราตอบกลับแบบสอบถามต่ำ การปรับเพิ่ม/ลดขนาดตัวอย่างเพื่อชดเชยกรณีตัวแปรในประชากรที่ศึกษามีความแปรปรวนต่ำ/สูงกว่าปกติ การปรับเพิ่ม/ลดกรณีที่การเลือกตัวอย่างมิได้ใช้การสุ่มแบบง่ายแต่ใช้การสุ่มแบบอื่น เช่นการสุ่มแบบแบ่งชั้น เป็นต้น **ประการที่สอง ควรประมาณค่าขนาดตัวอย่างทุกชุดกรณีที่มีการทดสอบหลายชุด** การวิจัยส่วนใหญ่มีการทดสอบหลายชุด นักวิจัยควรประมาณค่าขนาดตัวอย่างทุกชุดเพื่อให้ได้ช่วงของขนาดตัวอย่างที่เหมาะสมสำหรับทุกการทดสอบ และนำค่าสูงสุดมากำหนดเป็นขนาดตัวอย่างในภาพรวม และ **ประการที่สาม ไม่ควรใช้โปรแกรมเป็นเครื่องมือช่วยให้นักวิจัยเลือกขนาดตัวอย่างตามใจตน** การใช้โปรแกรม G*Power มีข้อควรระวัง คือ ความสะดวกในการประมาณค่าขนาดตัวอย่าง ทำให้นักวิจัยบางคนใช้โปรแกรม G*Power เป็นเครื่องมือกำหนดขนาดตัวอย่างตามความต้องการของตน โดยการปรับเพิ่ม/ลดตัวเลขข้อมูลที่ใช้ในการประมาณค่าขนาดตัวอย่าง ซึ่งเป็นสิ่งที่ไม่ควรทำเพราะขาดจรรยาการวิจัย สุดท้ายผู้เขียนคาดหวังว่า บทความเรื่องนี้จะเป็นประโยชน์ ให้ความรู้แก่นักวิจัยผู้สนใจพัฒนางานวิจัยให้มีมาตรฐานระดับนานาชาติ ได้นำไปใช้ในการวิจัย และช่วยกันเผยแพร่แนวทางที่ถูกต้องทันสมัยในการกำหนดขนาดตัวอย่างสำหรับการวิจัยต่อไป

บรรณานุกรม

นงลักษณ์ วิรัชชัย (2552). *การเลือกใช้สถิติที่เหมาะสมกับการวิจัย* เอกสารประกอบการบรรยาย 'Twilight Program' การนำเสนอผลงานวิจัยแห่งชาติ 2552 (Thailand Research EXPO 2009) ณ ศูนย์ประชุมบางกอกคอนเวนชั่นเซ็นเตอร์ ห้อง Lotus Suite 12 เซ็นทรัลเวิร์ล ราชประสงค์ กรุงเทพมหานคร 27 สิงหาคม 2552 เวลา 18.00-20.00 น.

นงลักษณ์ วิรัชชัย (2553). *ชุดวิชา 21701 การวิจัยหลักสูตรและการเรียนการสอน หน่วยที่ 10 สถิติวิเคราะห์เชิงปริมาณ: สถิติบรรยายและสถิติพารามเตริก* มหาวิทยาลัยสุโขทัยธรรมาธิราช กรุงเทพมหานคร: โรงพิมพ์มหาวิทยาลัยสุโขทัยธรรมาธิราช.

ศิริชัย กาญจนวาสี (2550). *สถิติประยุกต์สำหรับการวิจัย.* พิมพ์ครั้งที่ 5. กรุงเทพมหานคร: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

Butkovich, N.J., Smith, H. F. & Hoffman, C. E. (2004). A Comparison of updating frequency between web of science and current contents connect. *Issues in Science and Technology Librarianship.* (Electronic Journal).

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences.* New York: Academic Press.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences,* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cochran, W. G. 1963. *Sampling techniques*, (Second Edition). New York: John Wiley and Sons, Inc

Erdfelder, E., Faul, F., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavioral Research Method.* 39, 175-191.

Howell, D.C. (2010). *Statistical methods for psychology, (Seventh Edition).* Belmont: Wadsworth, Cengage Learning.

Israel, G. D. (2009). Determining Sample Size. Program Evaluation and Organizational Development, IFAS, University of Florida. PEOD-6.

Kerlinger, F.N. & Lee, H.B. (2000). *Foundations of behavioral research,* (Fourth Edition). New York: Harcourt College Publishers.

Kirk, R.E. (1995). *Experimental design: Procedure for the behavioral sciences,* (Third Edition). Pacific Groves: Brooks/Cole Publishing Company.

Krejcie, R. V. & Morgan, D. W. (1970). Determining sample sizes for research activities. *Educational and Psychological Measurement*. 30, 607-610.

Smith, M. F. (1983). *Sampling considerations in evaluating cooperative extension programs*. Florida Cooperative Extension Service Bulletin PE-1. Institute of Food and Agricultural Sciences. University of Florida.

Yamane, T. (1967). *Statistics: An introductory statistics,* (Second Edition). New York: Harper & Row.