บทที่

แนวคิดและหลักการ

ในการวิเคราะห์ความแปรปรวนร่วม

2

2.1 ความนำ

2.2 การทบทวนค่าสถิติและทฤษฎีสำคัญที่เกี่ยวข้อง

2.3 แนวคิดและหลักการในการวิเคราะห์ความแปรปรวนร่วม

2.4 วัตถุประสงค์หลักและการนำมาใช้

2.5 บทสรุป

2.6 เอกสารอ้างอิง

2.1 ความนำ

การวิเคราะห์ความแปรปรวนร่วม (analysis of covariance, ANCOVA) เป็นวิธีการทางสถิติที่ได้รับการยอมรับและถูกนำมาใช้ในงานวิจัยทางวิทยาศาสตร์สุขภาพ เพื่อประเมินผลกระทบจากสิ่งแทรกแซง *(Wan, 2021)* โดยเฉพาะภายใต้แบบแผนงานวิจัยเชิงทดลองสำหรับข้อมูลก่อน-หลังแบบวัดซ้ำสองกลุ่ม เมื่อข้อมูลก่อนและหลังทดลองที่มีลักษณะแบบต่อเนื่อง ถูกกำหนดให้เป็นตัวแปรร่วม (covariate) และตัวแปรผลลัพธ์ (outcome) ตามลำดับ ครอบคลุมทั้งงานวิจัยที่มีการจัดสรรผู้เข้าร่วมเข้ากลุ่มแบบสุ่มและไม่สุ่ม เช่น งานวิจัยเชิงทดลองแบบสุ่มที่มีกลุ่มควบคุม (randomized controlled trial, RCT) เกี่ยวกับประสิทธิผลของการใช้ยาลดไข้ชนิดหนึ่ง ด้วยการเปรียบเทียบระดับอุณหภูมิก่อนและหลังการทดลองระหว่างกลุ่มที่ได้รับยาลดไข้ (สิ่งแทรกแซง) กับกลุ่มที่ได้รับยาหลอก (placebo) หรือ งานวิจัยแบบกึ่งทดลอง (quasi-experimental study) เกี่ยวกับประสิทธิผลของโปรแกรมควบคุมระดับความดันเลือดในกลุ่มผู้ป่วยเบาหวาน ด้วยการเปรียบเทียบระดับความดันเลือดก่อนและหลังการทดลองระหว่างกลุ่มผู้ป่วยเบาหวานที่ได้รับโปรแกรม (กลุ่มทดลอง) กับกลุ่มผู้ป่วยเบาหวานที่ไม่ได้รับโปรแกรม (กลุ่มควบคุม) เป็นต้น แต่อย่างไรก็ตาม ในการนำวิธีการวิเคราะห์ ANCOVA มาใช้ในงานวิจัยภายใต้แบบแผนดังกล่าวที่ผ่านมา ยังพบว่า โดยส่วนใหญ่ มีขั้นตอนและแนวปฏิบัติที่ไม่ชัดเจนและครอบคลุม ซึ่งสามารถสังเกตได้จากหลายส่วนของผลการศึกษาที่ถูกรายงานในบทความวิจัยผ่านวารสารบนระบบออนไลน์ของประเทศ โดยเฉพาะประเด็นการตรวจสอบข้อตกลงเบื้องต้น พบว่า ส่วนใหญ่ไม่ให้ความสำคัญและนำเสนอขั้นตอนปฏิบัติอย่างไม่เป็นระบบที่สามารถนำไปสู่การตรวจสอบได้ หรือ ในประเด็นข้อค้นพบที่ได้จากงานวิจัย พบว่า ส่วนใหญ่ให้ความสำคัญกับข้อสรุปที่บ่งชี้เกี่ยวกับความแตกต่างระหว่างสองกลุ่มที่นำมาเปรียบเทียบภายใต้แนวทางการทดสอบสมมติฐานทางสถิติด้วยค่า p-value เพียงอย่างเดียวมากกว่า โดยละเลยการให้ความสำคัญกับข้อค้นพบ ซึ่งเป็นขนาดผลกระทบของวิธีการ หรือ treatment effect อย่างครอบคลุม ทั้งค่าประมาณและ 95%CI รวมถึงค่า effect size ได้แก่ ค่า eta-square (η^2) หรือ omega-square (ω^2) เป็นต้น นอกจากการนำวิธีการวิเคราะห์ ANCOVA มาใช้งานที่ขาดความชัดเจนและไม่ครอบคลุมดังกล่าวแล้ว ในทางปฏิบัติ อาจพบได้ว่า องค์ความรู้ หรือ เอกสารสนับสนุนเชิงวิชาการที่มีอยู่ในประเทศและเกี่ยวข้อง โดยเฉพาะภาษาไทย ยังมีจำนวนจำกัดและไม่ทันสมัย จึงอาจเป็นผลให้นักวิจัยส่วนหนึ่งมีการนำวิธีการดังกล่าวมาใช้งานด้วยแนวปฏิบัติเดิม ซึ่งผ่านการเรียนรู้แบบสะสมจากรุ่นสู่รุ่น ดังนั้นเพื่อให้นักวิจัยสามารถทบทวนและฟื้นฟูองค์ความรู้เกี่ยวกับประเด็นข้อจำกัดดังกล่าวให้มีความทันสมัย ในบทนี้จึงได้กล่าวถึงเนื้อหาที่เกี่ยวข้องกับแนวคิดและหลักการในการวิเคราะห์ความแปรปรวนร่วม ประกอบด้วย การทบทวนค่าสถิติและทฤษฎีสำคัญที่เกี่ยวข้อง แนวคิดและหลักการในการวิเคราะห์ความแปรปรวนร่วม ข้อควรคำนึงถึงในการวิเคราะห์ความแปรปรวนร่วม บทสรุปและเอกสารอ้างอิง

2.2 การทบทวนค่าสถิติและทฤษฎีสำคัญที่เกี่ยวข้อง

การวิเคราะห์ความแปรปรวนร่วม หรือ analysis of covariance (ANCOVA) เป็นวิธีการทางสถิติที่ถูกพัฒนาและขยายผลต่อเนื่องมาจากการวิเคราะห์ความแปรปรวน หรือ analysis of variance (ANOVA) โดยจากชื่อเรียกของทั้งสองวิธีการดังกล่าว เกี่ยวข้องโดยตรงกับค่าสถิติที่สำคัญ ประกอบด้วย “ความแปรปรวน (variance)” “ความแปรปรวนร่วม (covariance)” และ “สัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน (Pearson’s coefficient)” ดังนั้นการทบทวนและฟื้นฟูแนวคิดพื้นฐาน วิธีการคำนวณของค่าสถิติและทฤษฎีสำคัญที่เกี่ยวข้อง จึงมีความสำคัญและจำเป็นสำหรับนักวิจัย ทั้งนี้เพื่อนำไปใช้ต่อยอดและเรียนรู้เพิ่มเติมเกี่ยวกับแนวคิดและหลักการในการวิเคราะห์ความแปรปรวนร่วม ภายใต้แบบแผนงานวิจัยเชิงทดลองสำหรับข้อมูลก่อน-หลังแบบวัดซ้ำสองกลุ่มในหัวข้อถัดไป โดยมีรายละเอียด ดังนี้

2.2.1 ความแปรปรวน (variance)
2.2.2 ทฤษฎีเกี่ยวกับคุณสมบัติการเป็นตัวประมาณค่าที่ไม่เอนเอียงของพารามิเตอร์
2.2.3 ความแปรปรวนร่วม (covariance)
2.2.4 สัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน (Pearson’s correlation coefficient)

2.2.1 ความแปรปรวน (variance)

ความแปรปรวน (variance) ถือเป็นค่าสถิติที่แสดงความผันแปรของค่าข้อมูลภายในตัวแปรเดียวว่า

“โดยเฉลี่ยแล้ว ค่าข้อมูลแต่ละค่า (x_i) ภายใต้ตัวแปรเดียวกัน มีการกระจายห่างจากค่าเฉลี่ย ($\bar{x}$) ของชุดข้อมูลภายในตัวแปรเดียวกันนั้น มาก หรือ น้อย”

หากค่าความแปรปรวน “มาก” ก็แสดงว่า “โดยเฉลี่ยแล้ว ค่าข้อมูลแต่ละค่าภายใต้ตัวแปรเดียวกันมีการกระจายห่างจากค่าเฉลี่ยของชุดข้อมูลภายในตัวแปรเดียวกันมาก”

หากมีค่าความแปรปรวน “น้อย” ก็แสดงว่า “โดยเฉลี่ยแล้ว ค่าข้อมูลแต่ละค่าภายใต้ตัวแปรเดียวกันมีการกระจายห่างจากค่าเฉลี่ยของชุดข้อมูลภายในตัวแปรเดียวกันน้อย”

ดังนั้นวิธีการพิจารณา จึงใช้ข้อมูลเพียงตัวแปรเดียวมาประกอบการคำนวณ ซึ่งในกรณีพิจารณาข้อมูลระดับตัวอย่าง สามารถคำนวณได้จากสูตรดังนี้

$$S^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$$

เมื่อ S^2 = ค่าความแปรปรวนของตัวอย่าง
x_i = ค่าสังเกตที่ i ของตัวแปร X
$\bar{x}$ = ค่าเฉลี่ยของตัวแปร X

หรือ สามารถพิจารณาได้ดังเช่น **กรณีตัวอย่างที่ 2.1**

กรณีตัวอย่างที่ 2.1

สมมติให้ คะแนนพฤติกรรมการดูแลสุขภาพตนเอง (คะแนนเต็ม 30) ของผู้ป่วยเบาหวาน จำนวน 5 ราย เป็นดังนี้

20 12 18 9 11

จากกรณีตัวอย่างข้างต้น นักวิจัยสามารถพิจารณาคำนวณค่าเฉลี่ยได้ ดังนี้

$$\overline{X} = \frac{\sum X_i}{n} = \frac{20+12+18+9+11}{5} = 14$$

จากข้อมูลคะแนนพฤติกรรมการดูแลสุขภาพตนเอง (คะแนนเต็ม 30) ของผู้ป่วยเบาหวาน จำนวน 5 ราย และค่าเฉลี่ยของชุดข้อมูลทั้งหมด สามารถนำมาใช้ประกอบการพิจารณาค่าความแปรปรวนได้ดังเช่น **ตารางที่ 2.1**

ตารางที่ 2.1 แสดงแนวทางการพิจารณาค่าความแปรปรวนของข้อมูลจากชุดตัวอย่าง

❶ **คนที่ (i)**	❷ **ค่าข้อมูล** (x_i)	❸ **ความห่างระหว่างค่าข้อมูลกับค่าเฉลี่ย** $(x_i - \overline{x})$	❺ **ยกกำลังสองของความห่างระหว่างค่าข้อมูลกับค่าเฉลี่ย** $(x_i - \overline{x})^2$
1	20	(20 - 14) = 6	$(20 - 14)^2$ =36
2	12	(12 - 14) = -2	$(12 - 14)^2$ = 4
3	18	(18 - 14) = 4	$(18 - 14)^2$ =16
4	9	(9 - 14) = -5	$(9 - 14)^2$ = 25
5	11	(11 - 14) = -3	$(11 - 14)^2$ = 9
ผลรวม		$\sum(x_i - \overline{x})$ = 0 ❹	$\sum(x_i - \overline{x})^2$ = 90 ❻

จากตารางที่ 2.1 ข้างต้น พบว่า คอลัมน์ที่ ❶ แสดงจำนวนลำดับของผู้ป่วยเบาหวานแต่ละคน (i) และคอลัมน์ที่ ❷ แสดงคะแนนพฤติกรรมการดูแลสุขภาพตนเอง (คะแนนเต็ม 30) ของผู้ป่วยเบาหวานแต่ละราย, คอลัมน์ที่ ❸ แสดงความห่างที่เกิดขึ้นระหว่างค่าข้อมูลผู้ป่วยเบาหวานแต่ละราย (x_i) กับค่าเฉลี่ย $(\overline{x})$ ของคะแนนพฤติกรรมการดูแลสุขภาพตนเองจากผู้ป่วยเบาหวานทั้งหมดในตัวอย่างและเนื่องจากหลักการในการคำนวณค่าความแปรปรวน เราต้องการทราบผลรวมของความห่างที่เกิดขึ้นระหว่างค่าข้อมูลผู้ป่วยเบาหวานแต่ละราย (x_i) กับค่าเฉลี่ยของผู้ป่วยเบาหวานในชุดตัวอย่างนี้ $(\overline{x})$ เพื่อนำค่าผลรวมที่ได้ ไปใช้ในการพิจารณาคำนวณหาค่าเฉลี่ยของความห่างที่เกิดขึ้นในภาพรวมอีกครั้งหนึ่ง ทำให้ในตำแหน่งหมายเลข ❹ จึงเป็นการแสดงผลรวมของความห่างที่เกิดขึ้นระหว่างค่าข้อมูลผู้ป่วยเบาหวานแต่ละราย (x_i) กับค่าเฉลี่ยของผู้ป่วยเบาหวานในชุดตัวอย่างนี้ $(\overline{x})$ และพบว่า ผลรวมเท่ากับ ศูนย์ หรือ $\sum(x_i - \overline{x}) = 0$ ทั้งนี้เนื่องจากมีการนำค่าจำนวนที่ติดลบ อันเป็นผลที่เกิดขึ้นจาก

การตั้งลบที่วางค่าจำนวนน้อยอยู่ในตำแหน่งตัวตั้งและค่าจำนวนมากอยู่ในตำแหน่งตัวลบ จึงทำให้ ค่าผลลัพธ์ที่ได้มีค่าติดลบ เช่น กรณีผู้ป่วยเบาหวานรายที่ 4 พบว่า ได้ค่าคะแนนฯ (x_4) เท่ากับ 9 ขณะที่ค่าเฉลี่ยของผู้ป่วยเบาหวานในชุดตัวอย่างนี้ ($\bar{x}$) เท่ากับ 14 ดังนั้นเมื่อมีการนำมาตั้งลบ เพื่อพิจารณาความห่างที่เกิดขึ้น จึงมีผลลัพธ์ที่ได้เป็น 9 – 14 = -5 เป็นต้น ซึ่งจากผลลัพธ์ที่ได้ดังกล่าว เมื่อนำค่าผลลัพธ์ที่เกิดจากการตั้งลบของผู้ป่วยเบาหวานแต่ละรายที่มีอยู่ทั้งหมดมารวมกัน จึงทำให้ผลลัพธ์รวมที่ได้เท่ากับ ศูนย์ ซึ่งถือว่า เป็นไปตามหลักการทางคณิตศาสตร์ปกติที่มีอยู่

ในทางปฏิบัติ ประเด็นสำคัญที่สนใจจากการตั้งลบดังกล่าว ก็คือ ต้องการทราบเฉพาะของขนาดความห่างที่เกิดขึ้นระหว่างค่าข้อมูลผู้ป่วยเบาหวานแต่ละราย (x_i) กับค่าเฉลี่ยของผู้ป่วยเบาหวานในชุดตัวอย่างนี้ ($\bar{x}$) เท่านั้น ดังนั้นเพื่อเป็นการแก้ปัญหาผลลัพธ์ที่มีค่าติดลบดังกล่าว จึงสามารถนำหลักการทางคณิตศาสตร์มาใช้ด้วยการยกกำลังสอง เพื่อให้ค่าติดลบของผลลัพธ์ที่ได้หายไป ดังแสดงในคอลัมน์ที่ ❺ และทำให้ผลรวมของการยกกำลังค่าความห่างที่เกิดขึ้นระหว่างค่าข้อมูลผู้ป่วยเบาหวานแต่ละราย (x_i) กับค่าเฉลี่ยของผู้ป่วยเบาหวานในชุดตัวอย่างนี้ ($\bar{x}$) พบว่า มีค่าเท่ากับ 90 หรือ $\sum(x_i - \bar{x})^2 = 90$ ดังแสดงในตำแหน่งหมายเลข ❻ ซึ่งจากค่าผลรวมดังกล่าวที่คำนวณได้จากตารางที่ 2.1 จึงสามารถนำมาแทนค่าในสูตรการคำนวณค่าความแปรปรวนของระดับตัวอย่างได้ดังนี้

$$S^2 = \frac{\sum(x_i - \bar{x})^2}{n-1} = \frac{90}{4} \approx 22.5$$

อย่างไรก็ตาม เนื่องจากผลลัพธ์ที่ได้จากสูตรการคำนวณค่าความแปรปรวนของตัวอย่างอยู่ในรูปยกกำลังสอง ($\sum(x_i - \bar{x})^2$) ซึ่งค่อนข้างมีข้อจำกัดในการนำไปใช้เพื่อแปลผลเกี่ยวกับการกระจายของข้อมูลร่วมกับค่าเฉลี่ย ซึ่งอยู่ในรูปปกติไม่ยกกำลังสอง ($\sum x_i$) ดังนั้นในทางปฏิบัติสำหรับการนำไปใช้ประโยชน์สำหรับค่าความแปรปรวนดังกล่าว จึงมักถูกแปลงค่าผลลัพธ์ที่คำนวณได้ ให้อยู่ในรูปปกติไม่ยกกำลังสอง ด้วยการถอดรากกำลังที่สอง (square root) ของค่าความแปรปรวนและเรียกค่าใหม่นี้ว่า ***“ค่าส่วนเบี่ยงเบนมาตรฐาน”*** หรือ ***“standard deviation”*** แทนด้วย “***S***” หรือ “***SD***” โดยมีสูตรในการพิจารณาดังนี้

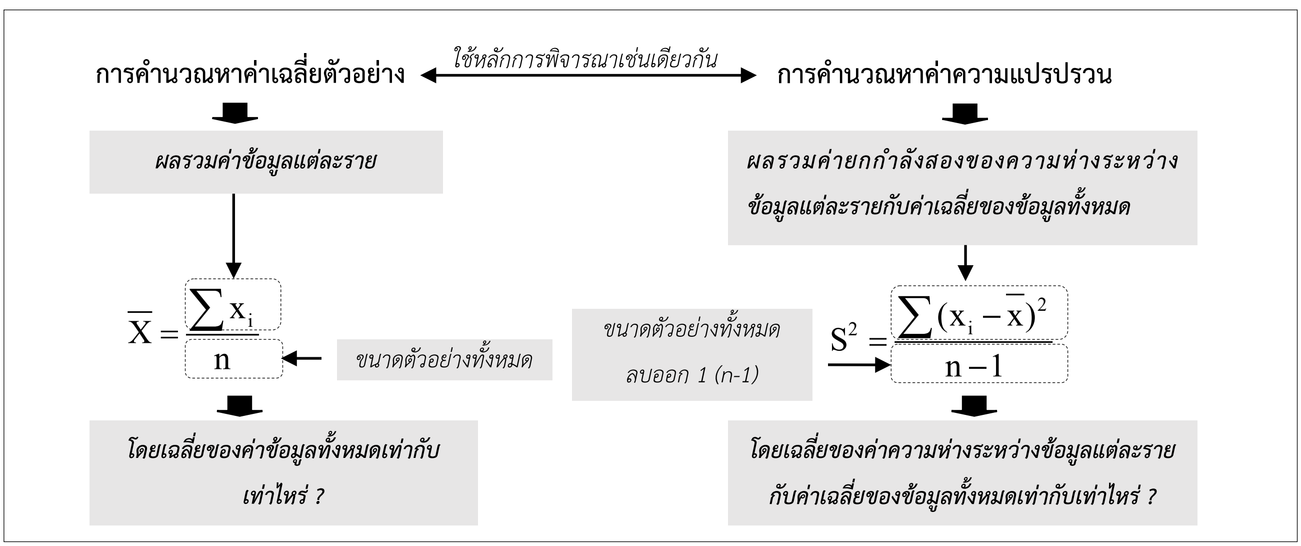
$$S = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n-1}}$$

จากนั้นจึงนำค่าผลลัพธ์ใหม่ที่ได้ มาใช้ในการแปลผลร่วมกับค่าเฉลี่ย โดยมีแนวทางการแปลผลลัพธ์ที่ได้ไปในทิศทางเช่นเดียวกันกับค่าเฉลี่ย ดังเช่นกรณีตัวอย่างที่ 2.1 ข้างต้น พบว่า คะแนนพฤติกรรมการดูแลสุขภาพฯ ของผู้ป่วยเบาหวาน มีค่าเฉลี่ยเท่ากับ 14 และค่าความแปรปรวนเท่ากับ 22.5 ซึ่งสามารถแปลงค่าความแปรปรวนดังกล่าวมาเป็นค่าส่วนเบี่ยงเบนมาตรฐาน จะได้ $\sqrt{22.5} \cong 4.74$ โดยในทางปฏิบัติ นักวิจัยสามารถแปลความหมายของค่าส่วนเบี่ยงเบนมาตรฐาน ได้เช่นเดียวกับความแปรปรวน ดังนี้

“ โดยเฉลี่ยแล้ว คะแนนพฤติกรรมการดูแลสุขภาพฯ ของผู้ป่วยเบาหวานแต่ละราย (x_i) ห่างจากค่าเฉลี่ย ($\bar{x}$) ของคะแนนพฤติกรรมการดูแลสุขภาพฯ ที่พิจารณาจากผู้ป่วยเบาหวานที่มีอยู่ของตัวอย่างทั้งหมด ประมาณ 4.74 ”

ซึ่งค่าส่วนเบี่ยงเบนมาตรฐานนี้ หากมีค่ามาก แสดงได้ว่า ข้อมูลแต่ละรายที่มีอยู่ในตัวอย่างมีการกระจายห่างจากค่าเฉลี่ยมาก ขณะเดียวกันนักวิจัยจำเป็นต้องตระหนักไว้เสมอว่า ค่าส่วนเบี่ยงเบนมาตรฐานของตัวอย่างที่คำนวณได้ ***ไม่สามารถมีค่าติดลบได้ ทั้งนี้เนื่องจากวิธีการได้มาของค่าผลลัพธ์ ถูกพิจารณาต่อมาจากรูปยกกำลังสองในสูตรการคำนวณความแปรปรวนของตัวอย่าง ดังนั้นผลลัพธ์ที่ได้ในรูปยกกำลังสองนั้น จึงเป็นไปไม่ได้ที่จะมีค่าติดลบ*** แต่อย่างไรก็ตาม ก็มีความเป็นไปได้ ที่ค่าส่วนเบี่ยงเบนมาตรฐานจะมีค่าเป็นศูนย์ หากข้อมูลที่มีอยู่ทุกค่าของชุดตัวอย่างทั้งหมดนั้นเป็นค่าเดียวกัน

แต่อย่างไรก็ตาม เพื่อให้นักวิจัยสามารถอธิบายและแปลความหมายของผลลัพธ์ที่ได้จากการคำนวณค่าความแปรปรวน ได้อย่างถูกต้องและเหมาะสม จึงสามารถแสดงผลการเทียบเคียงสูตรในการหาค่าความแปรปรวนกับสูตรในการหาค่าเฉลี่ยของตัวอย่าง ให้เห็นความคล้ายกันของแนวคิดและวิธีในการพิจารณา ทั้งตำแหน่งตัวตั้ง (เศษ) และตำแหน่งตัวหาร (ส่วน) ดังเช่น **แผนภาพที่ 2.1**



แผนภาพที่ 2.1 แสดงหลักการพิจารณาค่าเฉลี่ยและความแปรปรวนในระดับตัวอย่าง

จากแผนภาพที่ 2.1 เป็นการแสดงหลักการพิจารณาค่าเฉลี่ยและความแปรปรวนในระดับตัวอย่างซึ่งมีลักษณะเช่นเดียวกัน นั่นคือ กรณีตัวตั้งสำหรับการพิจารณาค่าเฉลี่ย ประกอบด้วย ผลรวมของค่าข้อมูลแต่ละราย หรือ $\sum x_i = x_1 + x_2 +, ..., + x_n$ ขณะที่ตัวตั้งสำหรับการพิจารณาคำนวณค่าของความแปรปรวน ประกอบด้วยผลรวมค่ายกกำลังสองของความห่างระหว่างข้อมูลแต่ละรายกับค่าเฉลี่ยข้อมูลทั้งหมดของชุดตัวอย่าง หรือ

$\sum(x_i-\bar{x})^2=(x_1-\bar{x})^2+(x_2-\bar{x})^2+,...,+(x_n-\bar{x})^2$ ส่วนกรณีตัวหารสำหรับการพิจารณาค่าเฉลี่ย ได้แก่ ขนาดตัวอย่าง (n) หรือ จำนวนข้อมูลที่มีทั้งหมด ขณะที่ตัวหารสำหรับการพิจารณาค่าความแปรปรวน ได้แก่ ขนาดตัวอย่างที่ถูกหักลบออกด้วยค่า 1 หรือ n-1 จากตัวเลขดังกล่าว อาจสังเกตได้ว่า มีความแตกต่างกันและขณะเดียวกันอาจสร้างความสงสัยให้กับนักวิจัยหลายคนเกี่ยวกับที่มาของการนำหนึ่งมาลบออกจากจำนวนขนาดตัวอย่างในกรณีที่เกิดขึ้นดังกล่าว ซึ่งคำตอบและคำอธิบายสามารถศึกษารายละเอียดเพิ่มเติมได้ในหัวข้อถัดไป

2.2.2 ทฤษฏีเกี่ยวกับคุณสมบัติการเป็นตัวประมาณค่าที่ไม่เอนเอียงของพารามิเตอร์

ในการดำเนินงานวิจัย นักวิจัยส่วนใหญ่ต้องการทราบค่าจริงที่ถูกคำนวณได้จากประชากร เพราะถือเป็นค่าคงที่และในหนึ่งประชากร จะมีได้เพียงค่าเดียว หรือ เรียกว่า **ค่าพารามิเตอร์ (parameter)** โดยหากนักวิจัยสามารถเก็บรวบรวบข้อมูลได้ครบถ้วนทั้งหมดจากประชากร ค่าพารามิเตอร์ดังกล่าว ก็จะถูกนำมาพิจารณาและหาค่าได้ แต่อย่างไรก็ตาม ในทางปฏิบัติ การเก็บรวบรวบข้อมูลทั้งหมดดังกล่าว อาจมีข้อจำกัดทั้งด้านทรัพยากร เวลา และงบประมาณ จึงทำให้นักวิจัยไม่สามารถเก็บรวบรวมข้อมูลได้ครบทั้งหมดจากประชากร ดังนั้นแนวทางในการเก็บรวบรวมข้อมูลบางส่วน จึงถูกนำมาพิจารณาผ่านวิธีการสุ่มและขนาดตัวอย่างที่ถูกพิจารณาอย่างถูกต้องและเพียงพอ จากนั้นจึงนำข้อมูลบางส่วนดังกล่าวมาพิจารณาและคำนวณหาค่าที่ต้องการทราบ ซึ่งค่าที่คำนวณได้จากตัวอย่าง จะถูกเรียกว่า **ค่าสถิติ (statistics)** โดยค่าสถิติดังกล่าวนี้ สามารถมีค่าที่เป็นไปได้มากกว่าหนึ่งค่า ขึ้นกับลักษณะตัวอย่างที่ถูกสุ่มมาได้ในแต่ละครั้ง

ในทางปฏิบัติ แม้นักวิจัยจะทำการเก็บรวบรวมข้อมูลจากตัวอย่าง ขณะเดียวกันกลับต้องการสรุปผลไปยังค่าพารามิเตอร์ของประชากร ซึ่งก็สามารถนำค่าสถิติที่ได้จากตัวอย่างดังกล่าวไปอ้างอิงถึงค่าพารามิเตอร์ของประชากรได้ ด้วยวิธีการทางสถิติที่เรียกว่า **สถิติอนุมาน (inferential statistics)** ประกอบด้วย วิธีการประมาณค่าพารามิเตอร์ (parameter estimation) และการทดสอบสมมติฐานทางสถิติ (statistical hypothesis) โดยเฉพาะวิธีการประมาณค่าพารามิเตอร์ ซึ่งค่าความแปรปรวนของตัวอย่าง (S^2) ถือเป็นอีกหนึ่งตัวประมาณค่าของตัวอย่างเช่นเดียวกับค่าเฉลี่ย ($\bar{x}$) เพื่อนำไปใช้ประมาณค่าพารามิเตอร์ของประชากร โดยการพิจารณาเพื่อตัดสินใจว่าตัวประมาณค่าของตัวอย่างใด เป็นตัวประมาณค่าพารามิเตอร์ของประชากรได้ดีที่สุด ในทางทฤษฏีมีหลายคุณสมบัติที่ถูกนำมาใช้ในการประเมินและตัดสินใจ โดยหนึ่งในคุณสมบัติสำคัญนั้น ได้แก่ ***คุณสมบัติการเป็นตัวประมาณค่าที่ไม่เอนเอียง หรือ unbiased estimators*** โดยมีหลักเกณฑ์ในการพิจารณาดังนี้

> **“เมื่อมีการพิจารณาตัวประมาณค่าของตัวอย่าง ด้วยค่าคาดหวังคณิตศาสตร์ หรือ mathematical expectation (E)” แล้ว หากพบว่า มีค่าเท่ากับค่าพารามิเตอร์ของประชากร นั่นแสดงว่า ตัวประมาณค่าของตัวอย่างนั้นเป็นตัวประมาณค่าที่ไม่เอนเอียงของค่าดังกล่าว หรือ การบ่งชี้ได้ว่า $E(\hat{\theta})=\theta$ เมื่อ $\hat{\theta}$ = ตัวประมาณค่าของค่าพารามิเตอร์ประชากร θ”**