3.1 ความนำ

ในการวิเคราะห์ข้อมูลภายใต้บริบทของงานวิจัยเชิงทดลอง หรือ แบบกึ่งทดลอง โดยเฉพาะแบบแผนงานวิจัยสำหรับข้อมูลก่อน–หลังแบบวัดซ้ำสองกลุ่ม การทำความเข้าใจลักษณะของข้อมูลในเบื้องต้นถือเป็นขั้นตอนที่มีความสำคัญอย่างยิ่ง เนื่องจากข้อมูลที่ได้จากการศึกษา อาจมีความแตกต่างกันในด้านคุณลักษณะพื้นฐานของกลุ่มตัวอย่าง เช่น อายุ เพศ หรือ ค่าพื้นฐานของตัวแปรผลลัพธ์ (baseline) ซึ่งปัจจัยเหล่านี้ อาจส่งผลกระทบต่อค่าประมาณ หรือ ผลลัพธ์ที่นักวิจัยสนใจ หากไม่ได้รับการพิจารณาจัดการ หรือ ควบคุมอย่างเหมาะสม ซึ่งอาจนำไปสู่การแปลความหมายและสรุปผลที่ผิดพลาดได้ ดังนั้นการวิเคราะห์ข้อมูลด้วยสถิติพรรณนา (descriptive statistics) จึงเข้ามามีบทบาทสำคัญและถือเป็นแนวทางในการตรวจสอบและการทำความเข้าใจข้อมูลเบื้องต้นในภาพรวม ครอบคลุมทั้งสถิติที่เกี่ยวข้องกับการวัดแนวโน้มเข้าสู่ส่วนกลาง การวัดการกระจายของข้อมูล รวมถึงการตรวจสอบความแตกต่างของข้อมูลพื้นฐานระหว่างกลุ่ม ก่อนที่จะขยับไปสู่การวิเคราะห์ข้อมูลด้วยวิธีการทางสถิติหลักที่ได้กำหนดไว้ดังเช่น การวิเคราะห์ความแปรปรวนร่วม (ANCOVA) ต่อไป โดยเฉพาะในงานวิจัยเชิงทดลองแบบสุ่ม เช่น RCT การบ่งชี้ความสมดุลระหว่างกลุ่มที่เกิดขึ้นจากกระบวนการสุ่ม (randomization) อย่างถูกต้อง จะนำไปสู่การเลือกวิธีการสถิติเชิงอนุมานมาใช้เพื่อตอบคำถามวิจัยได้อย่างชัดเจนและเหมาะสมมากขึ้น ***(Ooms, Waldron et al. 2025)*** เช่นเดียวกัน ในงานวิจัยเชิงทดลองแบบไม่สุ่ม เช่น งานวิจัยแบบกึ่งทดลอง หากนักวิจัยสามารถการันตีได้ว่า ลักษณะข้อมูลพื้นฐานระหว่างกลุ่มมีความใกล้เคียงกันตั้งแต่ก่อนเริ่มทดลอง ความแตกต่างที่เกิดขึ้นระหว่างกลุ่มภายหลังการทดลอง จะช่วยเพิ่มความน่าเชื่อถือในการตีความผลลัพธ์ภายหลังการทดลองได้ ***(Capili and Anastasi 2024)*** ดังนั้นการทบทวนและทำความเข้าใจรายละเอียดของวัตถุประสงค์ในการวิเคราะห์ข้อมูลด้วยสถิติพรรณนาสำหรับวิธี ANCOVA รวมถึงการอธิบายลักษณะข้อมูลและการเปรียบเทียบข้อมูลพื้นฐานระหว่างกลุ่มจึงมีความสำคัญและจำเป็นสำหรับนักวิจัย เพราะถือเป็นกระบวนการที่จะช่วยสนับสนุนทำให้การวิเคราะห์ข้อมูลด้วยวิธี ANCOVA มีความถูกต้อง แม่นยำและสามารถตีความผลลัพธ์ได้อย่างมีความหมายในบริบทของงานวิจัยที่กำหนด โดยเฉพาะในกรณีที่มีการนำตัวแปรร่วม (covariate) มาใช้ในการปรับค่าสำหรับการวิเคราะห์ข้อมูล แต่อย่างไรก็ตาม ในการวิเคราะห์ข้อมูล เครื่องมือในการคำนวณค่าทางสถิติที่ครอบคลุมและทันสมัย ยังมีความจำเป็น โดยในตำราเล่มนี้ได้นำโปรแกรม STATA ซึ่งเป็นโปรแกรมที่ถูกนำมาใช้ค่อนข้างแพร่หลาย โดยเฉพาะทางด้านการแพทย์และสาธารณสุข ***(Wang and Leeflang 2019)*** และเป็นโปรแกรมที่มีทิศทางและแนวโน้มการนำมาใช้เพิ่มขึ้นในทุก ๆ 10 ปี อย่างต่อเนื่องตั้งแต่ปี 1997, 2007 และ 2017 เท่ากับ ร้อยละ 1.2, 1.8 และ 15.1 ตามลำดับ ***(Masuadi, Mohamud et al. 2021)*** ด้วยศักยภาพของโปรแกรมที่ครอบคลุมความสามารถทั้งด้านการบริหารจัดการข้อมูล (data management) และด้านการวิเคราะห์ทางสถิติ (statistical analysis) ซึ่งประกอบด้วยการนำเสนอข้อมูลด้วยวิธีการที่หลากหลายจากกราฟและแผนภูมิ การวิเคราะห์ข้อมูลด้วยวิธีการทางสถิติขั้นพื้นฐานและสถิติขั้นสูงอย่างครอบคลุม ดังนั้นเพื่อให้นักวิจัยสามารถทบทวนสถิติพรรณนาที่เกี่ยวข้องและมีทักษะในการนำโปรแกรม STATA มาใช้ ในบทนี้ จึงได้นำเสนอแนวทางการวิเคราะห์ข้อมูลด้วยโปรแกรม STATA โดยมุ่งเน้นการวิเคราะห์ข้อมูลด้วยสถิติพรรณนา ซึ่งประกอบด้วย วัตถุประสงค์ของการวิเคราะห์ข้อมูลด้วยสถิติพรรณนาสำหรับวิธี ANCOVA , การอธิบายลักษณะข้อมูล, การเปรียบเทียบข้อมูลพื้นฐานระหว่างกลุ่ม บทสรุปและเอกสารอ้างอิง

3.2 วัตถุประสงค์ของการวิเคราะห์ข้อมูลด้วยสถิติพรรณนาสำหรับวิธี ANCOVA

การวิเคราะห์ข้อมูลด้วยสถิติพรรณนา (descriptive statistics) ถือเป็นขั้นตอนเริ่มต้นที่มีความสำคัญในการวิเคราะห์ข้อมูลเชิงปริมาณ เนื่องจากช่วยให้นักวิจัยสามารถอธิบาย สรุป และทำความเข้าใจลักษณะพื้นฐานของข้อมูลก่อนนำไปสู่การวิเคราะห์ด้วยสถิติเชิงอนุมาน *(Bulanov, Suvorov et al. 2021)* โดยเฉพาะในบริบทของการวิเคราะห์ความแปรปรวนร่วม (ANCOVA) สำหรับงานวิจัยเชิงทดลอง หรือ แบบกึ่งทดลองที่มีข้อมูลก่อน–หลังแบบวัดซ้ำสองกลุ่ม ซึ่งตัวแปรผลลัพธ์และตัวแปรร่วมมักมีลักษณะเป็นข้อมูลแบบต่อเนื่อง โดยในทางปฏิบัติการวิเคราะห์ข้อมูลด้วยสถิติพรรณนาในขั้นตอนนี้ ไม่ได้มีบทบาทเพียงเพื่อรายงานลักษณะของข้อมูลเท่านั้น หากแต่ยังช่วยตรวจสอบคุณภาพของข้อมูล ประเมินความสมดุลของข้อมูลพื้นฐานระหว่างกลุ่มและพิจารณาความเหมาะสมของตัวแปรร่วมที่จะนำไปใช้ในตัวแบบ ANCOVA สำหรับขั้นตอนในบทถัดไป ดังนั้นเพื่อให้นักวิจัยสามารถเข้าใจบทบาทของสถิติพรรณนาในบริบทของวิธี ANCOVA ได้อย่างชัดเจนและเป็นระบบมากขึ้น ในหัวข้อนี้จึงจำแนกวัตถุประสงค์สำคัญของการวิเคราะห์ข้อมูลด้วยสถิติพรรณนา ออกได้เป็น 5 ประเด็น ดังนี้

3.2.1 เพื่ออธิบายลักษณะทั่วไปของข้อมูล
3.2.2 เพื่อตรวจสอบความสมดุลของข้อมูลพื้นฐานระหว่างกลุ่ม
3.2.3 เพื่อตรวจสอบลักษณะการกระจายของข้อมูลและค่าผิดปกติ
3.2.4 เพื่อตรวจสอบความสัมพันธ์ระหว่างตัวแปรร่วมและตัวแปรผลลัพธ์
3.2.5 เพื่อเตรียมความพร้อมของข้อมูลสำหรับการวิเคราะห์ด้วยวิธี ANCOVA

3.2.1 เพื่ออธิบายลักษณะทั่วไปของข้อมูล

การอธิบายลักษณะทั่วไปของข้อมูล เป็นการสรุปข้อมูลในภาพรวม โดยพิจารณาครอบคลุม ทั้งค่าของการวัดแนวโน้มเข้าสู่ส่วนกลาง (measures of central tendency) และค่าวัดการกระจายของข้อมูล (measures of variability) โดยมีค่าสถิติที่นิยมใช้ ได้แก่ ค่าเฉลี่ย (mean) ค่ามัธยฐาน (median) ค่าส่วนเบี่ยงเบนมาตรฐาน (standard deviation) พิสัย (range) และพิสัยระหว่างควอไทล์ (interquartile range, IQR) ซึ่งค่าดังกล่าวเหล่านี้จะช่วยทำให้นักวิจัยได้ทำความเข้าใจลักษณะของข้อมูลทั้งในระดับภาพรวมและระดับรายกลุ่มได้ค่อนข้างดีและเป็นระบบ โดยเฉพาะในงานวิจัยเชิงทดลอง หรือ แบบกึ่งทดลองสำหรับข้อมูลก่อน–หลังแบบวัดซ้ำสองกลุ่ม ซึ่งการพิจารณาค่าข้อมูลพื้นฐานของตัวแปรผลลัพธ์ (baseline outcome) ถือว่า มีความสำคัญเป็นพิเศษ เนื่องจากความแตกต่าง หรือ ไม่แตกต่างที่เกิดขึ้นของค่าเริ่มต้นระหว่างกลุ่ม สามารถส่งผลกระทบโดยตรงต่อผลลัพธ์ หรือค่าประมาณที่ได้ภายหลังการติดตาม หรือ หลังการได้รับวิธีการ หรือ สิ่งแทรกแซงอย่างชัดเจน

3.2.2 เพื่อตรวจสอบความสมดุลของข้อมูลพื้นฐานระหว่างกลุ่ม

การตรวจสอบความสมดุลของข้อมูลพื้นฐานระหว่างกลุ่ม (baseline comparability) ถือเป็นหนึ่งในข้อพิจารณาที่สำคัญของงานวิจัยเชิงทดลอง โดยเฉพาะในกรณีที่ไม่มีการจัดสรรผู้เข้าร่วมเข้ากลุ่มแบบสุ่ม หรือแบบกึ่งทดลอง ซึ่งความแตกต่าง หรือ ความไม่สมดุลของลักษณะพื้นฐานระหว่างกลุ่ม เช่น อายุ เพศ ระดับ

ความรุนแรงของโรค หรือ ค่าพื้นฐานของตัวแปรผลลัพธ์ เป็นต้น สามารถส่งผลกระทบต่อผลลัพธ์ภายหลังการติดตามได้ ดังนั้นการใช้สถิติพรรณนาในการเปรียบเทียบข้อมูลพื้นฐานระหว่างกลุ่ม เช่น ค่าเฉลี่ย ค่าส่วนเบี่ยงเบนมาตรฐาน จำนวน ร้อยละ หรือ สัดส่วนของตัวแปรต่าง ๆ จะช่วยทำให้นักวิจัยสามารถประเมินระดับความสมดุลของกลุ่มตัวอย่างในระยะเริ่มต้นได้ และอีกทั้งยังสามารถนำมาใช้เป็นข้อมูลประกอบในการพิจารณาตัวแปรร่วม (covariate) สำหรับการวิเคราะห์ด้วยวิธี ANCOVA ให้มีความเหมาะสมมากขึ้นได้อีกด้วย

แต่อย่างไรก็ตาม การประเมินความสมดุลของข้อมูลพื้นฐาน นักวิจัยอาจต้องพึงระวัง โดยไม่ควรมุ่งเน้นเฉพาะการทดสอบนัยสำคัญทางสถิติ หรือ การพิจารณาจากค่า p-value เพียงอย่างเดียว เนื่องจากในงานวิจัยแบบสุ่ม ความแตกต่างของข้อมูลพื้นฐานระหว่างกลุ่ม อาจเกิดขึ้นได้จากโอกาส แม้กระบวนการสุ่มจะดำเนินการอย่างถูกต้องแล้วก็ตาม ในขณะที่งานวิจัยแบบกึ่งทดลอง ค่า p-value ก็อาจได้รับอิทธิพลจากขนาดตัวอย่างและไม่สามารถนำมาใช้เพื่อแสดงความสำคัญเชิงคลินิก หรือ ความสำคัญเชิงปฏิบัติของความแตกต่างได้อย่างเพียงพอ ดังนั้นการประเมิน baseline comparability ในทางปฏิบัติ นักวิจัยจึงควรพิจารณาจากขนาดของความแตกต่าง ทิศทางของความแตกต่าง ความสำคัญเชิงเนื้อหาและบทบาทของตัวแปรนั้นต่อผลลัพธ์ที่ศึกษา มากกว่าการตัดสินจากค่า p-value เพียงอย่างเดียว ***(Stang and Baethge 2018)***

3.2.3 เพื่อตรวจสอบลักษณะการกระจายของข้อมูลและค่าผิดปกติ

การตรวจสอบลักษณะการกระจายของข้อมูลและค่าผิดปกติ ถือเป็นอีกขั้นตอนที่สำคัญในการเลือกใช้ค่าสถิติพรรณนาให้มีความเหมาะสมกับลักษณะของข้อมูล ซึ่งหากข้อมูลมีการแจกแจงใกล้เคียงปกติ การรายงานค่าเฉลี่ยร่วมกับส่วนเบี่ยงเบนมาตรฐาน จะมีความเหมาะสม ขณะที่หากข้อมูลมีการแจกแจงแบบเบ้ หรือ มีค่าผิดปกติที่ส่งผลต่อค่ากลางของข้อมูล การรายงานค่ามัธยฐานร่วมกับพิสัยระหว่างควอไทล์ จะมีความเหมาะสมมากกว่า ***(Smeltzer and Ray 2022)*** นอกจากนี้ ค่าผิดปกติ (outliers) อาจส่งผลกระทบต่อค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และค่าประมาณของพารามิเตอร์ในตัวแบบได้ โดยเฉพาะเมื่อค่าดังกล่าว มีความเบี่ยงเบนสูงผิดปกติจากข้อมูลส่วนใหญ่ ดังนั้นการตรวจสอบข้อมูลด้วยกราฟ เช่น histogram, boxplot หรือ scatter plot ร่วมกับการพิจารณาค่าสถิติพรรณนาที่เกี่ยวข้อง จึงช่วยให้นักวิจัยสามารถประเมินลักษณะการกระจายของข้อมูล ตรวจสอบค่าผิดปกติและตัดสินใจเลือกวิธีการสรุปข้อมูลได้อย่างถูกต้องและเหมาะสมก่อนนำข้อมูลเข้าสู่การวิเคราะห์ด้วยวิธี ANCOVA ต่อไป

3.2.4 เพื่อตรวจสอบความสัมพันธ์ระหว่างตัวแปรร่วมและตัวแปรผลลัพธ์

หลักการสำคัญของวิธี ANCOVA คือ การปรับค่าผลลัพธ์ด้วยตัวแปรร่วม (covariate) เพื่อช่วยลดความผันแปรของค่าความคลาดเคลื่อนและเพิ่มความแม่นยำของการประมาณค่าผลกระทบของวิธีการ หรือ treatment effect ของตัวแปรกลุ่ม ดังนั้นตัวแปรร่วมที่นำมาใช้ในตัวแบบ จึงควรมีความสัมพันธ์กับตัวแปรผลลัพธ์ในระดับที่เหมาะสม ***(Tsiatis, Davidian et al. 2008)*** โดยเฉพาะค่าพื้นฐานของตัวแปรผลลัพธ์ ซึ่งมักมีความสัมพันธ์กับค่าผลลัพธ์ภายหลังการติดตาม ดังนั้นการใช้แผนภาพ scatter plot หรือ การคำนวณค่าสัมประสิทธิสหสัมพันธ์ (correlation coefficient) จึงอาจช่วยทำให้นักวิจัยสามารถประเมินความสัมพันธ์เบื้องต้นระหว่าง

ตัวแปรร่วมและตัวแปรผลลัพธ์ รวมถึงพิจารณาความเหมาะสมของตัวแปรร่วมก่อนนำเข้าสู่ตัวแบบ ANCOVA ได้อย่างเป็นระบบมากยิ่งขึ้น

3.2.5 เพื่อเตรียมความพร้อมของข้อมูลสำหรับการวิเคราะห์ด้วยวิธี ANCOVA

การวิเคราะห์ข้อมูลด้วยสถิติพรรณนา ถือเป็นขั้นตอนสำคัญในการเตรียมความพร้อมของข้อมูลก่อนนำเข้าสู่การวิเคราะห์ด้วยวิธี ANCOVA โดยช่วยให้นักวิจัยสามารถประเมินความเหมาะสมของข้อมูล พิจารณาค่าสถิติที่ควรใช้ในการสรุปข้อมูล และระบุตัวแปรที่มีความเหมาะสมในการนำมาใช้เป็นตัวแปรร่วม (covariate) ภายใต้ตัวแบบได้อย่างเป็นระบบ ซึ่งประเด็นดังกล่าวเหล่านี้ ล้วนมีผลโดยตรงต่อความถูกต้อง ความน่าเชื่อถือ และความสามารถในการตีความผลการวิเคราะห์ด้วยวิธี ANCOVA

ดังนั้นจากวัตถุประสงค์ทั้งหมดที่กล่าวมาข้างต้น จึงสามารถสรุปได้ว่า

> ***การวิเคราะห์ข้อมูลด้วยสถิติพรรณนาในบริบทของวิธี ANCOVA ไม่ได้มีบทบาทเพียงแค่เพื่อสรุปลักษณะทั่วไปของข้อมูลเท่านั้น หากแต่ยังเป็นขั้นตอนสำคัญในการตรวจสอบคุณภาพของข้อมูล ประเมินความสมดุลของกลุ่มตัวอย่าง พิจารณาความเหมาะสมของตัวแปรร่วม และเตรียมข้อมูลให้พร้อมสำหรับการวิเคราะห์ในระดับที่สูงขึ้น ซึ่งจะมีส่วนสำคัญในการช่วยสนับสนุนให้ผลการวิเคราะห์ที่ได้มีความถูกต้อง น่าเชื่อถือ และสามารถตีความได้อย่างมีความหมายภายใต้บริบทของแบบแผนงานวิจัยที่กำหนด***

3.3 การอธิบายลักษณะข้อมูล

การอธิบายลักษณะของข้อมูล (descriptive data analysis) เป็นขั้นตอนสำคัญในการวิเคราะห์ข้อมูลทางสถิติ เนื่องจากมีบทบาทในการช่วยให้นักวิจัยสามารถทำความเข้าใจโครงสร้างของข้อมูลในเบื้องต้น ก่อนนำไปสู่การวิเคราะห์ในระดับที่สูงขึ้น โดยเฉพาะการวิเคราะห์ด้วยวิธี ANCOVA ในงานวิจัยเชิงทดลองทั้งแบบสุ่มและแบบกึ่งทดลองสำหรับข้อมูลก่อน–หลังแบบวัดซ้ำสองกลุ่ม การอธิบายลักษณะข้อมูล ไม่ได้มีวัตถุประสงค์เพียง เพื่อสรุปค่าทางสถิติพื้นฐานเท่านั้น หากแต่ยังสามารถช่วยให้นักวิจัยได้มีการตรวจสอบคุณภาพของข้อมูล การประเมินลักษณะการกระจายของข้อมูล การตรวจสอบค่าผิดปกติ และรวมถึงพิจารณาความเหมาะสมของตัวแปรต่าง ๆ ที่อาจจะนำมาใช้ในการปรับค่าภายใต้ตัวแบบ ANCOVA อีกด้วย ดังนั้นเพื่อให้นักวิจัยสามารถทบทวนและทำความเข้าใจในรายละเอียดของแต่ละประเด็นได้อย่างชัดเจน รวมถึงสามารถนำคำสั่งที่เกี่ยวข้องในโปรแกรม STATA มาใช้ในการวิเคราะห์ข้อมูลสถิติพรรณนาได้อย่างถูกต้องและเหมาะสม ในหัวข้อนี้ จึงได้นำเสนอประเด็นสำคัญของการอธิบายลักษณะข้อมูล จำแนกเป็น 3 ส่วน ได้ดังนี้

3.3.1 การวัดแนวโน้มเข้าสู่ส่วนกลาง (measures of central tendency)
3.3.2 การวัดการกระจายของข้อมูล (measures of variability)
3.3.3 คำสั่งโปรแกรม STATA ที่เกี่ยวข้อง

3.3.1 การวัดแนวโน้มเข้าสู่ส่วนกลาง (measures of central tendency)

ถือเป็นวิธีการที่ถูกนำมาใช้ เพื่อบ่งชี้ค่ากลางหนึ่งค่าที่สามารถเป็นตัวแทนที่ดีของคุณลักษณะ หรือตัวแปรที่สนใจได้ *(Ali, Bhaskar et al. 2019)* ซึ่งโดยทั่วไป ประกอบด้วย ค่าเฉลี่ย (mean) ค่ามัธยฐาน (median) และค่าฐานนิยม (mode) แต่อย่างไรก็ตาม แม้ค่าฐานนิยม จะถือเป็นค่าบ่งชี้หนึ่งที่ถูกนำมาใช้ในการพิจารณาค่ากลางภายใต้การวัดแนวโน้มเข้าสู่ส่วนกลาง แต่ในทางปฏิบัติ เมื่อนำมาใช้กับข้อมูลบางชุด โดยเฉพาะในงานวิจัยที่มีตัวแปรแบบต่อเนื่อง ค่าฐานนิยม ยังมีข้อจำกัด ภายใต้หลักการคิด ซึ่งถูกพิจารณาจากจำนวนความถี่ของค่าข้อมูลที่มีการซ้ำกันมากสุด นั่นคือ ผลลัพธ์ที่ได้จากค่าฐานนิยม อาจมีความเป็นไปได้มากกว่าหนึ่งค่า เช่น จากกรณีการประเมินคะแนนพฤติกรรมการรับรู้ความรุนแรงของโรคเบาหวานในกลุ่มผู้สูงอายุ จำนวน 10 ราย (คะแนนเต็ม 10 คะแนน) โดยคะแนนที่ได้ของคนที่ 1 ถึงคนที่ 10 เท่ากับ 8, 4, 5, 2, 7, 6, 5, 4, 7, 9 ตามลำดับและจากข้อมูลดังกล่าว เมื่อนำมาพิจารณาค่าฐานนิยม จะพบว่า มีค่าข้อมูลที่ซ้ำและมีความถี่เท่ากันจำนวน 2 ค่า คือ 5 และ 7 ดังนั้นจึงสามารถสรุปได้ว่า ในข้อมูลชุดนี้ มีค่าฐานนิยมสองค่า เท่ากับ 5 และ 7 เป็นต้น หรือ ขณะเดียวกันในบางชุดข้อมูล อาจไม่สามารถหาค่าฐานนิยมได้ เนื่องจากข้อมูลไม่มีจำนวนความถี่ของการซ้ำกัน เช่น จากกรณีการวัดระดับความดันเลือด systolic (mmHg) ของผู้ป่วยเบาหวาน จำนวน 6 ราย โดยมีระดับค่าความดันเลือด systolic ที่วัดได้ของคนที่ 1 ถึงคนที่ 6 เท่ากับ 128, 132, 155, 130, 147, 160 ตามลำดับ ซึ่งจากข้อมูลดังกล่าว เมื่อพิจารณาค่าฐานนิยม จะพบว่า ไม่สามารถหาค่าได้ เนื่องจากไม่มีค่าข้อมูลที่ซ้ำกัน เป็นต้น *(Manikandan 2011)* ดังนั้นในการวิเคราะห์ข้อมูลภาพรวม ด้วยสถิติพรรณนา สำหรับการวิเคราะห์ความแปรปรวนร่วม (ANCOVA) ภายใต้แบบแผนงานวิจัยเชิงทดลองสำหรับข้อมูลก่อน-หลังแบบวัดซ้ำสองกลุ่มในบทนี้ จึงไม่ได้กล่าวถึงการนำเอาค่าฐานนิยมมาใช้ แต่จะมุ่งเน้นวิธีการวัดแนวโน้มเข้าสู่ส่วนกลางที่ถูกนำมาใช้อย่างแพร่หลายและเป็นที่ยอมรับ 2 วิธี ประกอบด้วย

(1) ค่าเฉลี่ย (mean)
(2) ค่ามัธยฐาน (median)

(1) ค่าเฉลี่ย (mean)

เป็นค่าเฉลี่ยทางคณิตศาสตร์ของชุดข้อมูล โดยมีสัญลักษณ์ในระดับประชากรเป็น μ (“มิว”) และในระดับตัวอย่างเป็น $\overline{X}$ (“เอกซ์บาร์”) และเป็นวิธีการที่ถูกนำมาใช้อย่างแพร่หลายในงานวิจัย เนื่องจากมีวิธีการคำนวณที่ง่ายและแปลผลได้อย่างตรงไปตรงมา ซึ่งสามารถพิจารณาตัวบ่งชี้ได้อย่างชัดเจนหนึ่งค่า จากการนำเอาค่าข้อมูลทุกค่าที่มีอยู่มาบวก หรือ รวมกันและหารด้วยจำนวนข้อมูลทั้งหมด *(Mishra, Pandey et al. 2019)* โดยมีสูตรในการพิจารณาทั้งในระดับประชากรและระดับตัวอย่างได้ **ดังแสดงในแผนภาพที่ 3.1**

ระดับประชากร	ระดับตัวอย่าง
$\mu = \frac{X_1 + X_2 + X_3 + ... + X_N}{N}$	$\overline{X} = \frac{X_1 + X_2 + X_3 + ... + X_n}{n}$
เมื่อ μ = ค่าเฉลี่ยในระดับประชากร	$\overline{X}$ = ค่าเฉลี่ยในระดับตัวอย่าง
$X_1 + X_2 + X_3 + ... + X_N$ = ผลรวมของข้อมูล	$X_1 + X_2 + X_3 + ... + X_N$ = ผลรวมของข้อมูล
N = จำนวนข้อมูลทั้งหมดในระดับประชากร	n = จำนวนข้อมูลทั้งหมดในระดับตัวอย่าง

แผนภาพที่ 3.1 แสดงสูตรในการคำนวณค่าเฉลี่ยทั้งระดับประชากรและระดับตัวอย่าง

จากแผนภาพที่ 3.1 ข้างต้น ซึ่งแสดงถึงวิธีการคำนวณค่าเฉลี่ย ภายใต้สูตรและสัญลักษณ์ที่นำมาใช้พิจารณาแตกต่างกัน ตามระดับประชากรและตัวอย่าง ดังนั้นจึงสามารถแสดงข้อดี/ข้อจำกัดของค่าเฉลี่ย ได้ดังนี้

- ***ข้อดีของค่าเฉลี่ย***

1. ค่าเฉลี่ย เป็นตัวบ่งชี้ที่มีการนำข้อมูลทุกค่ามาพิจารณา ทำให้สามารถการันตีได้ว่า ค่าเฉลี่ยเป็นค่าที่มีความเป็นตัวแทนของข้อมูลทั้งหมดที่มีอยู่อย่างแท้จริง ดังนั้นในทางปฏิบัติ จึงถูกนำไปใช้ขยายผลต่อในการวิเคราะห์ข้อมูลทางสถิติอื่นได้อย่างไม่ลำเอียง

2. ในประชากร หรือ ตัวอย่างหนึ่งกลุ่ม สามารถพิจารณาค่าเฉลี่ยภายใต้คุณลักษณะ หรือตัวแปรที่สนใจ ได้เพียงค่าเดียว ดังนั้นจึงทำให้ค่าเฉลี่ยถูกนำไปใช้ประโยชน์ในการเปรียบเทียบระหว่างกลุ่มได้อย่างชัดเจนและเป็นรูปธรรม

- ***ข้อจำกัดของค่าเฉลี่ย***

1. ค่าเฉลี่ย เหมาะสำหรับการนำมาใช้กับค่าข้อมูลที่มีระดับมาตรวัดแบบอัตราส่วนสเกล (ratio scale) และแบบช่วงสเกล (interval scale) เท่านั้น เช่น ค่าดัชนีมวลกาย (BMI) หรือ ระดับคะแนน เป็นต้น แต่มีข้อจำกัด เมื่อนำมาใช้กับค่าข้อมูลที่มีระดับมาตรวัดแบบนามสเกล (nominal scale) และแบบอันดับสเกล (ordinal scale) เช่น เพศ หรือ ระดับความรุนแรงของโรค เป็นต้น เนื่องจากไม่สามารถแปลความหมายและสรุปตามนัยของค่าเฉลี่ยได้ เช่น การคำนวณค่าเฉลี่ยของเพศ (1=ชาย, 2=หญิง) ได้เท่ากับ 1.32 ซึ่งค่าตัวเลข 1 หรือ 2 ที่มีในตัวแปรเพศ หมายถึง รหัส (code) ที่ใช้ในการบ่งชี้และจำแนกเพศที่แตกต่างกันเท่านั้น จึงไม่ใช่ค่าข้อมูลตัวเลขจริงที่เกิดขึ้น ส่งผลให้แม้สามารถคำนวณค่าเฉลี่ยได้ แต่ก็ไม่สามารถแปลและสรุปผลให้มีความหมายได้ เป็นต้น

2. จากการพิจารณานำทุกค่าของข้อมูลที่มีอยู่ มาคำนวณหาค่าเฉลี่ยตามสูตรดังกล่าวข้างต้น จะพบว่า หากในชุดข้อมูลนั้น มีค่าข้อมูลใดที่สูง หรือ ต่ำมากผิดปกติ (extreme outliers) ย่อมส่งผลกระทบโดยตรงต่อผลลัพธ์ของค่าเฉลี่ยที่ได้ ให้มีค่าเบี่ยงเบน หรือ คลาดเคลื่อนไปจากความเป็นจริง ***(Mishra, Pandey et al. 2019)*** หรือ อาจกล่าวได้ว่า ***ค่าเฉลี่ย ถือเป็นค่าที่ได้รับผลกระทบจากข้อมูลที่มีความเบี่ยงเบนมาก (Ali, Bhaskar et al. 2019)*** หรือ ***ค่าเฉลี่ย มีความไวค่อนข้างมากจากค่าข้อมูลที่เป็น outliers*** นั่นเอง