DAD8103 ภาคเรียนที่ 2/2567
สถิติชั้นสูงสำหรับการวิจัยในการบริหารการพัฒนา

ครั้งที่ 1 วันอาทิตย์ที่ 8 ธันวาคม 2567

ผศ.ดร.วิชิต สุรดินทร์กูร ดร.นลินี สุรดินทร์กูร และดร.ศุภากร สุรดินทร์กูร

โครงการปรัชญาดุษฎีบัณฑิต มหาวิทยาลัยราชภัฏสวนสุนันทา

สถิติพื้นฐานสำหรับการวิจัย
BASIC STATISTICS

สถิติกับชีวิตประจำวัน

ด้านการพัฒนาประเทศ

- สถิติประชากร สถิติแรงงาน
- สถิติการเกิด-การตาย
- สถิติการส่งสินค้าออกหรือการนำสินค้าเข้า
- เพื่อวางแผนหรือกำหนดนโยบายในการพัฒนาประเทศให้ดียิ่งขึ้น

ด้านธุรกิจ

- การวางแผนการผลิต
- การจำหน่าย
- การบริหารจัดการ

ด้านการเกษตร

- สถิติภูมิอากาศ ฯลฯ
- ทำให้ผลิตผลมากที่สุด หรือ สูงกว่าเดิมในระยะเวลาที่ต้องการ

ด้านการศึกษา/วิจัย

- พัฒนาคุณภาพการศึกษา
- การศึกษาวิจัย

สถิติ (Statistics): 3 ความหมาย

1. **สถิติ** หมายถึง **ตัวเลข**ที่ใช้บรรยายเหตุการณ์หรือข้อเท็จจริง (facts) ของเรื่องต่างๆ ที่เราต้องการศึกษา เช่น สถิติจำนวนผู้ป่วย สถิติจำนวนคนเกิด สถิติจำนวนคนตาย เป็นต้น

2. **สถิติ** หมายถึง **ศาสตร์หรือวิชา**ที่ว่าด้วยหลักการและระเบียบวิธีทางสถิติ สถิติใน ความหมาย นี้มักเรียกว่า สถิติศาสตร์ (Statistics)

3. **สถิติ** หมายถึง **ค่าที่คำนวณขึ้นมาจากตัวอย่าง** เพื่อแสดงถึงคุณลักษณะบางอย่างของข้อมูลชุดนั้น โดยทั่วไปจะนำค่าสถิติไปใช้ในการประมาณค่าพารามิเตอร์ ตัวอย่างเช่น ถ้าเราสนใจรายได้เฉลี่ยของคนในหมู่บ้าน แล้วเราสามารถนำรายได้ของทุกคนมา รวมกันแล้วหาค่าเฉลี่ยของรายได้ ค่าเฉลี่ยที่คำนวณได้นี้ถือว่าเป็นค่าพารามิเตอร์ แต่ถ้าเราสุ่มตัวอย่างคนในหมู่บ้านมาจำนวนหนึ่งแล้วคำนวณรายได้เฉลี่ย ค่าเฉลี่ยที่ได้นี้จะเป็นค่าสถิติ

การวิจัย (RESEACH)

Research = “Re + search”

การค้นคว้าหาความรู้หรือคำตอบที่ได้จากการกระทำซ้ำแล้ว
ซ้ำอีก จนหระทั่งได้ผลสรุปหรือข้อยุติ

กระบวนการหาความรู้ที่ถูกต้อง เชื่อถือได้
อย่างเป็นระบบ

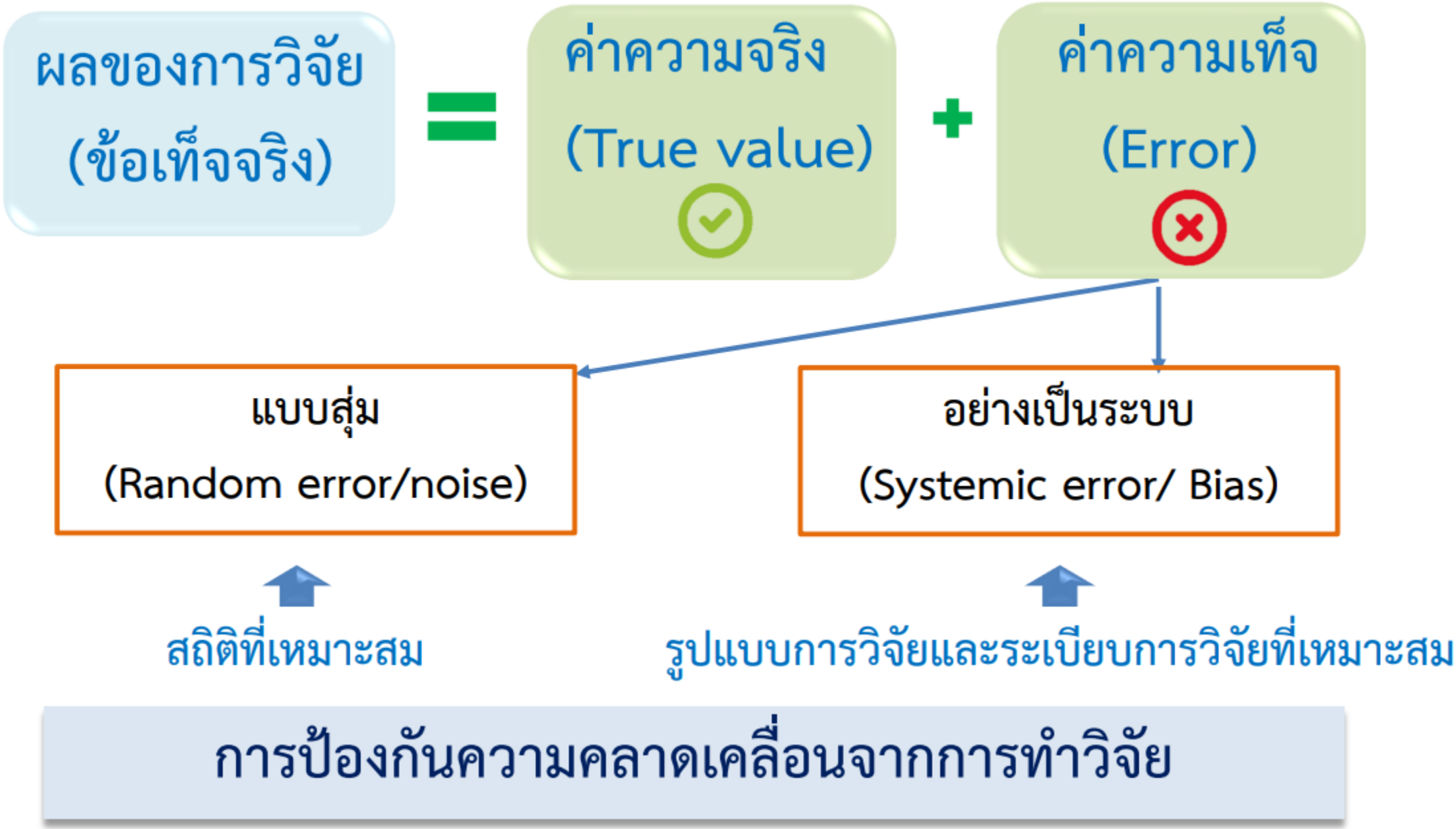
การวิจัย (RESEACH)

คำถามการวิจัย

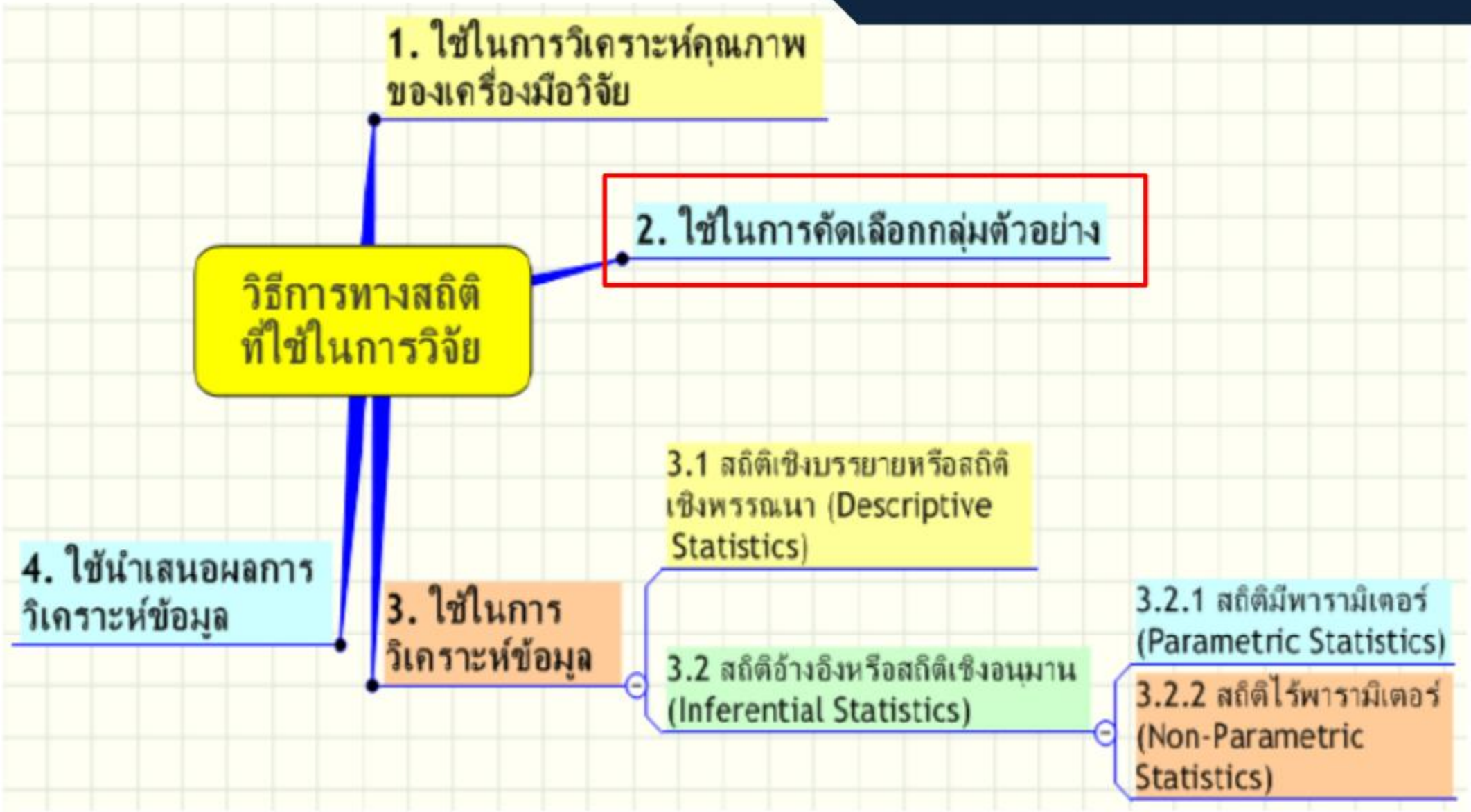
ดำเนินการวิจัยเพื่อค้นหาความจริง

ผลของการวิจัย (ข้อเท็จจริง) = ค่าความจริง (True value) + ค่าความเท็จ (Error)

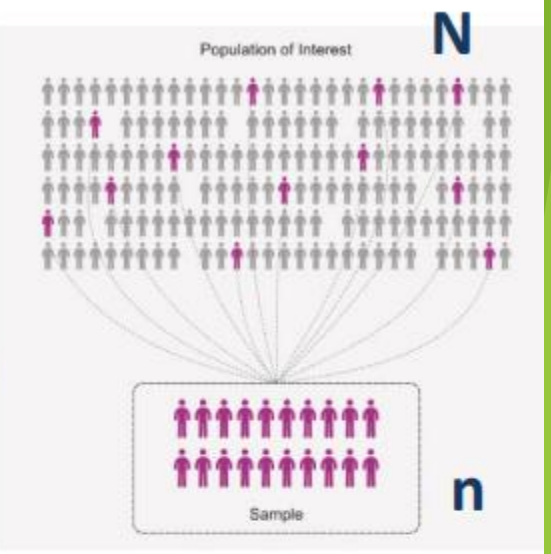
ความคลาดเคลื่อน (error) ของผลการวิจัย


ผลของการวิจัย
(ข้อเท็จจริง)
=
ค่าความจริง
(True value)
+
ค่าความเท็จ
(Error)
แบบสุ่ม
(Random error/noise)
อย่างเป็นระบบ
(Systemic error/ Bias)
สถิติที่เหมาะสม
รูปแบบการวิจัยและระเบียบการวิจัยที่เหมาะสม
การป้องกันความคลาดเคลื่อนจากการทำวิจัย

วิธีการทางสถิติที่ใช้ในการวิจัย



คำศัพท์ต่าง ๆ เกี่ยวกับสถิติ (1)



ประชากร (Population)

สมาชิกทุกหน่วย(ทั้งหมด)ที่ต้องการศึกษาหรือต้องการนำผลการวิจัยไปใช้อ้างอิงถึง คน สัตว์ พืช สิ่งของ หน่วยงานหรือสถาบัน (N=ขนาดประชากร) แบ่งเป็น

- Finite Population (ประชากรจำนวนจำกัดหรือนับได้)
- Infinite Population (ประชากรจำนวนไม่จำกัดหรือนับไม่ได้)

ตัวอย่าง (Sample)

สมาชิกบางส่วนของประชากรที่ผู้วิจัยสุ่มหรือเลือกมาศึกษาแทนประชากร (n=ขนาดตัวอย่าง) ตัวอย่างที่เป็นตัวแทนที่ดี (Representative) ของประชากร ขึ้นอยู่กับ

(1) Sample Size: n ขนาดตัวอย่างพอหรือไม่ -- > ใช้การคำนวณขนาดตัวอย่างจากสูตรที่เหมาะสมกับรูปแบบ วัตถุประสงค์ ระดับการวัดข้อมูล จำนวนกลุ่ม ความเป็นอิสระต่อกันของกลุ่ม

(2) Sampling technique วิธีการสุ่มตัวอย่างเหมาะสมหรือไม่

การคำนวณขนาดตัวอย่าง: Case-Control study

ตัวอย่างสิ่งส่งตรวจผู้ป่วยโรคมะเร็งลำไส้ใหญ่และทวารหนัก (กลุ่มทดลอง) และกลุ่มควบคุม

จากสูตร $$n = \left(\frac{r+1}{r}\right)\frac{(\bar{p})(1-\bar{p})(Z_{\beta}+Z_{\alpha/2})^2}{(p_1-p_2)^2}$$

- n = sample size of stool sample
- For 80% power, $Z\beta = .84$
- For 0.05 significance level, $Z\alpha = 1.96$
- r = 1 (equal number of cases and controls)
- The proportion exposed in the cases exposed is 0.43% (3)
- To get proportion of control group:

$$p_{control} = \frac{ORp_{caseexp}}{p_{caseexp}(OR-1)+1} \qquad p_{control} = \frac{2.0(.43)}{(.43)(2.0-1)+1} = \frac{.86}{1.43} = .060$$

- Average proportion exposed = (.60 + .43)/2 = .515

$$n = 2\frac{(.515)(1-.515)(.84+1.96)^2}{(.60-.43)^2}$$

ดังนั้น, n = 135 (67 cases, 67 controls)

การเก็บตัวอย่างอุจจาระจากอาสาสมัครสุขภาพดี

จากสูตร $$n = \frac{Z^2 P (1-P)}{d^2}$$

n = required sample size

Z = confidence level at 95% (standard value of 1.96)

P = estimated prevalence of previous report of ESBL producing *E. coli* in healthy adults ($p = 0.7$)

d = allowable error (0.1 or 10%)

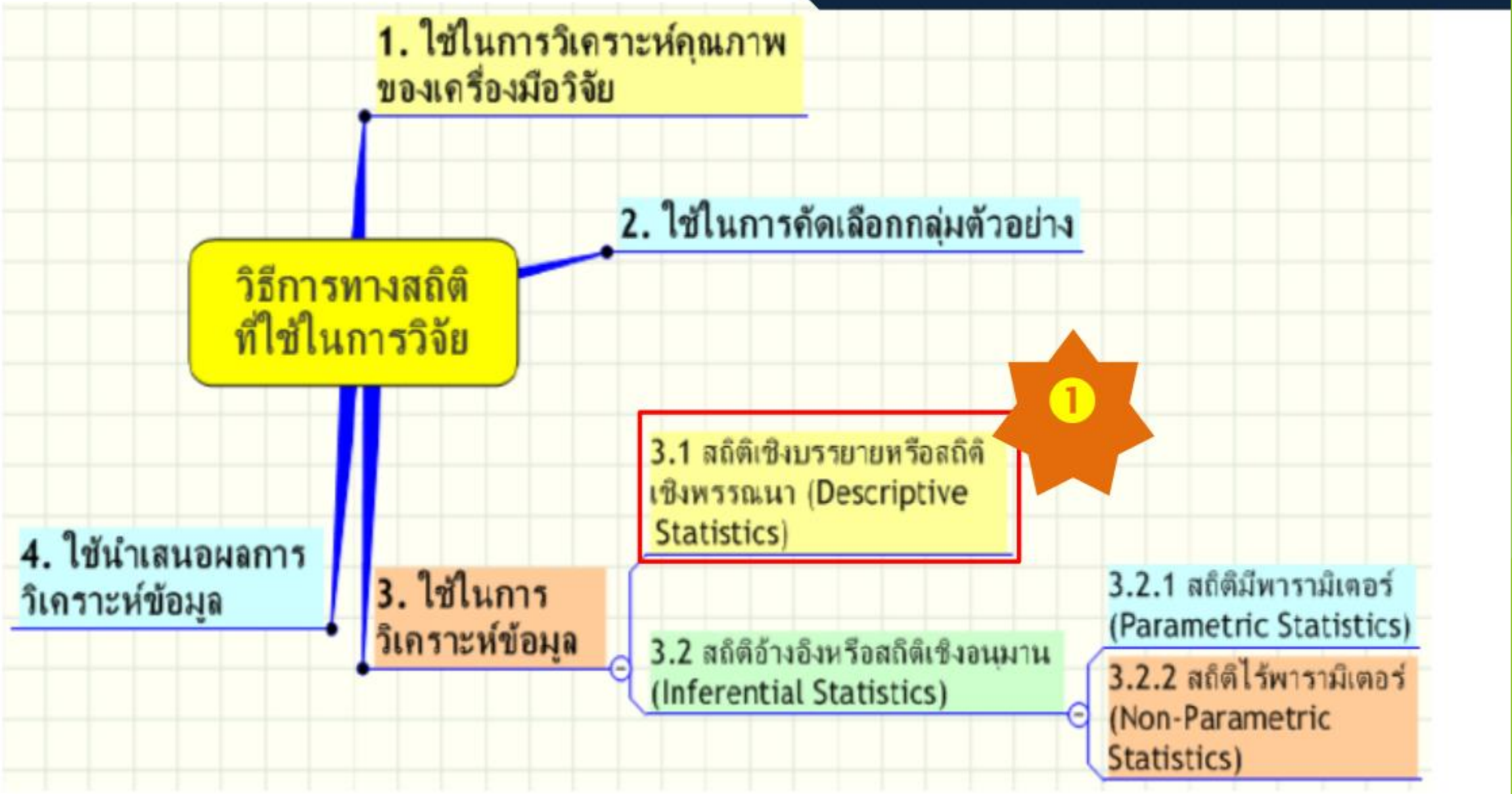
ดังนั้น, $n = (1.96)^2 \times 0.7 (1-0.7)/(0.1)^2$

$= 3.8416 \times 0.21/0.01$

$= 81$ sample

จากการแทนค่าตามสูตร จะได้ว่า การวิจัยนี้ต้องเก็บตัวอย่างอุจจาระจากอาสาสมัคร จำนวนรวมทั้งสิ้น 81 ราย ทั้งนี้ คำนวณจำนวนตัวอย่างเพิ่ม 10% เพื่อป้องกันความเสี่ยงในกรณีไม่สามารถติดตามอาสาสมัครได้ คิดเป็นจำนวนอย่างน้อย 90 ราย

วิธีการทางสถิติที่ใช้ในการวิจัย



ประเภทของสถิติในการวิจัย

สถิติบรรยายหรือสถิติเชิงพรรณนา (Descriptive Statistics)

เป็นวิธีการทางสถิติเพื่อใช้ในการสรุป บรรยาย อธิบายลักษณะข้อมูลที่ศึกษา จะพรรณนาภายในขอบเขตของข้อมูลที่เก็บรวบรวมมาเท่านั้น ไม่สามารถจะคาดคะเนลักษณะต่าง ๆ ออกไปนอกเหนือจากข้อมูลที่มีอยู่ได้หรือไม่มีการอ้างอิงหรืออนุมานไปถึงกลุ่มอื่น อาจพรรณนาโดย

- การแจกแจงความถี่ เป็นจำนวน ร้อยละ ตารางทางเดียวหรือตารางไขว้
- คำนวณค่าอัตรา อัตราส่วน
- คำนวณค่าการวัดแนวโน้มเข้าสู่ส่วนกลาง(ค่าเฉลี่ย มัธยฐาน และฐานนิยม)
- การวัดการกระจายของข้อมูล(ส่วนเบี่ยงเบนมาตรฐาน ค่าพิสัย เป็นต้น)
- การนำเสนอข้อมูลโดยใช้กราฟ หรือแผนภูมิต่าง ๆ

วิธีการทางสถิติที่ใช้ในการวิจัย

สถิติเชิงอนุมานหรือสถิติอ้างอิง (Inferential Statistics)

เป็นการใช้ระเบียบวิธีการทางสถิติในการอนุมานค่าพารามิเตอร์จากค่าสถิติ ที่คำนวณได้จากข้อมูลของกลุ่มตัวอย่างที่สุ่มมาศึกษา โดยอาศัยทฤษฎีความน่าเป็น ประกอบด้วย 2 องค์ประกอบ คือ

- การประมาณค่าพารามิเตอร์ (**Parameter Estimation**) นำเสนอในลักษณะของช่วงเชื่อมั่น (95% **Confidence Interval;95% CI**)
- การทดสอบสมมติฐาน (**Hypothesis Testing**) นำเสนอค่า **p-value** เพื่อสรุปผลการทดสอบสมมติฐานว่ามีนัยสำคัญทางสถิติ หรือ ไม่มีนัยสำคัญทางสถิติ(**Significance** หรือ **Non-Significance**) การทดสอบสมมติฐานทางสถิติจะใช้วิเคราะห์ในกรณีที่งานวิจัยที่มีการตั้งสมมติฐานไว้

เป็นสถิติที่มีความสำคัญและถูกใช้มากกว่า ลักษณะที่สำคัญคือ เป็นการศึกษาข้อมูลของกลุ่มตัวอย่าง แล้วนำผลสรุปไปประมาณหรือคาดการณ์ประชากร หรือเหตุการณ์ที่จะเกิดขึ้นในอนาคต

ลักษณะข้อมูลที่ใช้ในงานวิจัย

ข้อมูลเชิงคุณภาพ เป็นข้อมูลที่แสดงถึงสถานภาพ คุณลักษณะ หรือคุณสมบัติ เช่น เพศ ตำแหน่งหรือจำแนกตัวแปรตามระดับการวัด ได้แก่

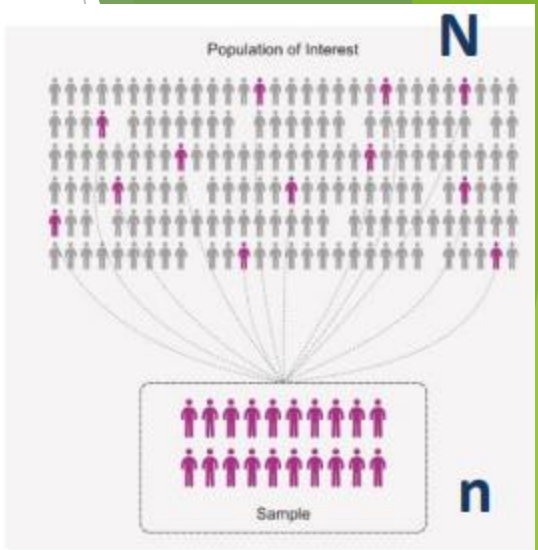
❶ นามบัญญัติ (nominal scale) จำแนกความแตกต่างของสิ่งที่ต้องการวัดออกเป็นกลุ่ม เช่น 1=ชาย 2=หญิง ตัวเลขไม่สามารถนำมาบวก ลบ คูณ หาร ได้

❷ เรียงอันดับ (ordinal scale) ใช้สำหรับจัดอันดับที่หรือตำแหน่งของสิ่งที่ต้องการวัด เช่น ระดับการศึกษา ผลการเรียน ความเก่ง ตัวเลขอันดับที่แตกต่างกันไม่สามารถบ่งบอกถึงปริมาณความแตกต่างได้ เช่น ไม่สามารถบอกได้ว่าเก่งกว่ากันเท่าไหร่ ตัวเลขสามารถนำมาบวกหรือลบกันได้

❸ อันตรภาค หรือระดับช่วง (interval scale) กำหนดค่าตัวเลขโดยมีช่วงห่างระหว่างตัวเลขเท่าๆ กัน สามารถนำตัวเลขมาเปรียบเทียบกันได้ว่าว่ามีปริมาณมากน้อยเท่าใด แต่ไม่สามารถบอกได้ว่าเป็นกี่เท่าของกันและกัน เพราะไม่มีศูนย์ที่แท้จริง เช่น คะแนนสอบ อุณหภูมิ (อุณหภูมิ 0 องศา มิได้หมายความว่าจะไม่มีความร้อน) ตัวเลขสามารถนำมาบวก ลบ คูณ หรือหารกันได้

❹ อัตราส่วน (ratio scale) สามารถกำหนดค่าตัวเลขให้กับสิ่งที่ต้องการวัด มีศูนย์แท้ เช่น น้ำหนัก ความสูง อายุ สามารถนำตัวเลขมาบวก ลบ คูณ หาร หรือหาอัตราส่วนกันได้ (เช่น ถนน 50 กิโลเมตร ยาวเป็น 2 เท่าของถนน 25 กิโลเมตร)

คำศัพท์ต่าง ๆ เกี่ยวกับสถิติ (2)



ค่าพารามิเตอร์ (**Parameter**)

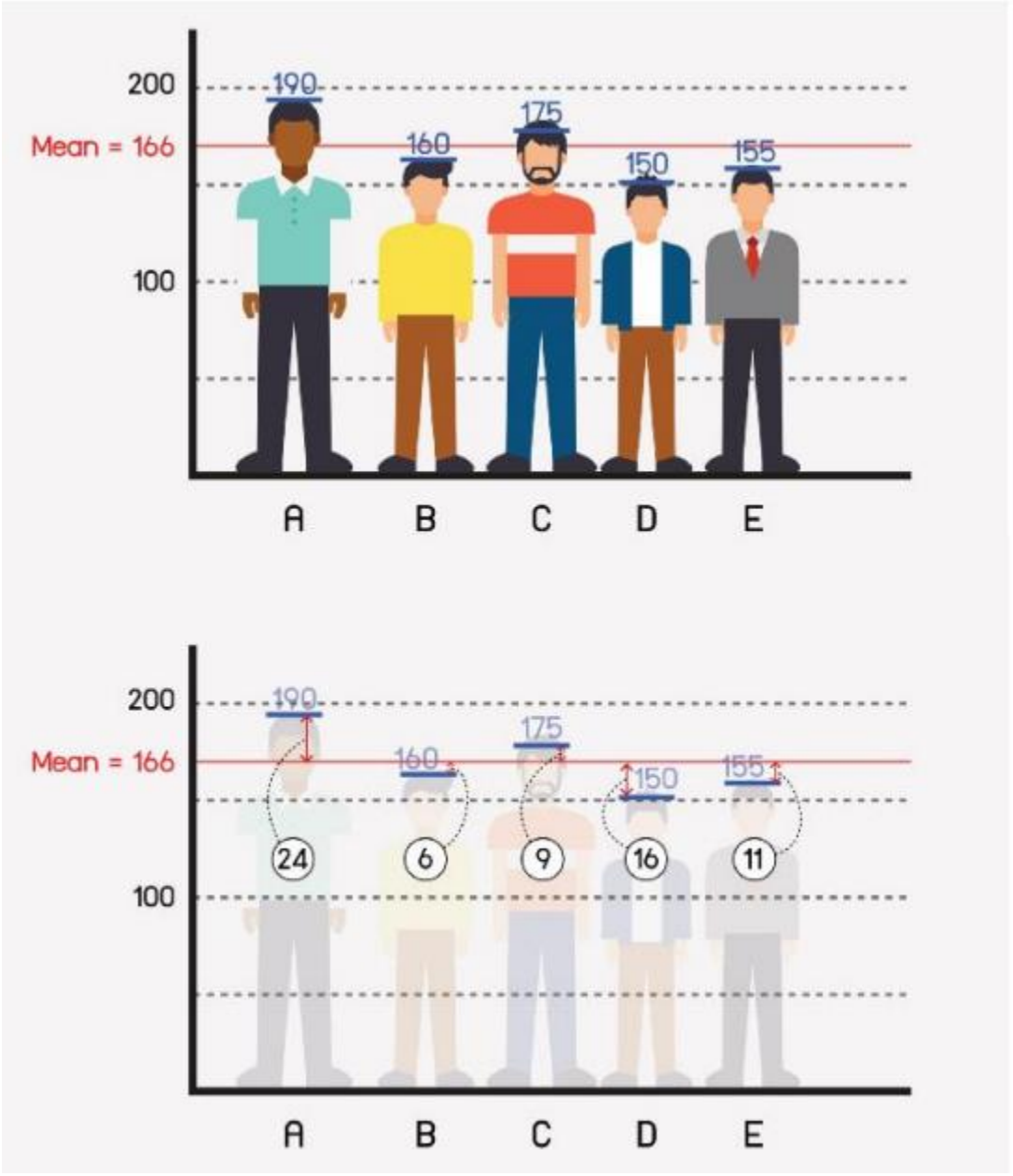
ค่าที่ใช้สรุป อธิบายลักษณะของประชากร หรือค่าที่คำนวณจากสมาชิกทั้งหมดของประชากร เขียนสัญลักษณ์ด้วย อักษรกรีก เช่น σ (ซิกม่า) คือ ส่วนเบี่ยงเบนมาตรฐานของประชากร ฯลฯ

ค่าสถิติ (Statistic)

ค่าที่ใช้สรุป อธิบายลักษณะของกลุ่มตัวอย่าง หรือค่าที่คำนวณจากกลุ่มตัวอย่างที่นำมาศึกษา เขียนสัญลักษณ์ด้วย อักษรภาษาอังกฤษ เช่น sd. คือ ส่วนเบี่ยงเบนมาตรฐานของตัวอย่าง ฯลฯ

เปรียบเทียบสัญลักษณ์พารามิเตอร์และค่าสถิติ P

	สัญลักษณ์ค่าพารามิเตอร์	สัญลักษณ์ค่าสถิติ
ค่าเฉลี่ย(Mean)	μ (มิว)	$\bar{x}$
ค่าสัดส่วน(Proportion)	π (พาย) หรือ P	p
ส่วนเบี่ยงเบนมาตรฐาน(Standard Deviation)	σ (ซิกม่า)	s หรือ s.d.
ความแปรปรวน(Variance)	σ^2	s^2 หรือ $s.d.^2$
สัมประสิทธิ์สหสัมพันธ์(Correlation Coefficient)	ρ (โร)	r
สัมประสิทธิ์การถดถอย(Regression Coefficient)	β (เบต้า)	b



ค่าเฉลี่ย คือ Mean

$$\frac{190+160+175+150+155}{5}=166$$

ค่าแปรปวน คือ Variance

$$\frac{24^2+(-6)^2+9^2+(-16)^2+(-11)^2}{5}=26.4$$

ค่าเบี่ยงเบนมาตรฐาน คือ Standard Deviation

$$\sqrt{26.4}=5.138$$

- Variance หรือ ค่าแปรปรวน

  ใช้เพื่อ**วัดการกระจายของข้อมูล**คิดจากค่าเฉลี่ยของความต่างจากค่าเฉลี่ยยกกำลัง 2

- Standard deviation หรือ ค่าเบี่ยงเบนมาตรฐาน

  เพื่อดูการ**กระจายข้อมูล**จากค่าเฉลี่ย คือ รากที่ 2 ของค่าแปรปรวน

คำศัพท์ต่าง ๆ เกี่ยวกับสถิติ (3)

ข้อมูล(Data)

ค่าที่เป็นตัวเลขหรือคุณลักษณะของตัวแปร เรียกว่า ข้อมูล(Data) ได้จากการ สังเกต ชั่ง ตวง วัด นับ หรือสอบถามจากหน่วยตัวอย่างที่ศึกษา โดยที่หน่วยที่ศึกษาอาจเป็นคน สัตว์ พืช และ สิ่งของ เมื่อหน่วยศึกษามีลักษณะที่แตกต่างกัน ข้อมูลที่ได้จึงแตกต่างกัน

ตัวแปร(Variable)

คุณลักษณะของสิ่งที่เราสนใจศึกษา อาจวัดเป็นตัวเลขเชิงปริมาณ เช่น น้ำหนัก ส่วนสูง อายุ เป็นต้น หรือบอกคุณลักษณะเชิงคุณภาพ เช่น เพศ สถานภาพสมรส มีค่าที่แปรเปลี่ยนไปตามหน่วยตัวอย่างที่ศึกษา

คำศัพท์ต่าง ๆ เกี่ยวกับสถิติ (4)

ตัวแปรต้น(Independent Variable)

ตัวแปรที่เกิดก่อนหรือเป็นสาเหตุของตัวแปรตามหรืออาจจะเรียกว่าตัวแปรอิสระ

ตัวแปรตาม(Dependent Variable)

ตัวแปรที่เกิดขึ้นหรือแปรผันไปตามตัวแปรอิสระ หรือกล่าวได้ว่าเป็นตัวแปรที่เป็นผลเมื่อตัวแปรอิสระเป็นเหตุ

ประเภทของตัวแปร

ตัวแปรเชิงคุณภาพ(Qualitative variable) **หรือตัวแปรแจงนับ** (Categorical Variable)

ตัวแปรที่แสดงลักษณะเชิงคุณภาพของหน่วยสังเกต โดยบ่งบอกหรือบรรยายลักษณะของตัวแปรนั้น เช่น ตัวแปรเพศ จะแสดงลักษณะเพศชาย เพศหญิง เป็นต้น ต้องทาการแจงข้อมูลออกตามคุณลักษณะเป็นกลุ่ม หรือพวกที่มีลักษณะเหมือนกันก่อน แล้วค่อยนับความถี่ออกมาเป็นจำนวน จึงเรียกว่าตัวแปรแจงนับ จะสรุปลักษณะของตัวแปรแบบนี้ โดยการแจกแจงความถี่ เป็นจำนวน คานวณค่าสัดส่วน ค่าร้อยละ หรือ ค่าอัตรา

ตัวแปรเชิงปริมาณ(Quantitative variable) **หรือตัวแปรต่อเนื่อง** (Continuous Variable)

ตัวแปรที่แสดงลักษณะของหน่วยสังเกตออกมาเป็นตัวเลขที่บอกขนาดหรือปริมาณได้ เช่น อายุ น้ำหนัก ส่วนสูง ระดับความดันโลหิต เป็นต้น ตัวแปรเชิงปริมาณ ได้จากการชั่ง ตวง วัด หรือนับ จะสรุป ลักษณะของตัวแปรแบบนี้ โดยการคำนวณการวัดแนวโน้มเข้าสู่ส่วนกลาง และค่าการกระจายของข้อมูล อาจจัด กลุ่มตัวแปรเหล่านี้เป็นช่วง แล้วค่อยแจกแจงความถี่ คำนวณค่าสัดส่วน ค่าร้อยละ หรือ ค่าอัตรา ก็ได้

การวิเคราะห์ข้อมูลในงานวิจัย

การวิเคราะห์ข้อมูลในงานวิจัย

ประกอบด้วยการใช้สถิติ 2 องค์ประกอบใหญ่ ๆ ได้แก่

- สถิติเชิงพรรณนา (Descriptive Statistics)
  - ->ใช้อธิบายลักษณะข้อมูลที่ศึกษา อยู่ภายในขอบเขตของข้อมูลที่เก็บรวบรวมมาเท่านั้น
- สถิติเชิงอนุมาน (Inferential Statistics)
  - ->นำค่าสถิติจากกลุ่มตัวอย่างไปใช้อนุมานค่าพารามิเตอร์ของประชากร

การวิเคราะห์ข้อมูลโดยใช้สถิติเชิงพรรณนา

1. กรณีข้อมูลแจงนับ (ตัวแปรเชิงคุณภาพ ประกอบด้วยข้อมูลที่มีระดับการวัดเป็น Nominal or Ordinal Scale) เช่น อาชีพ สถานภาพสมรส การศึกษา ตำแหน่ง การออกกำลังกาย การสูบบุหรี่ เป็นต้น

1.1) วิเคราะห์ข้อมูลโดยการแจกแจงความถี่ เป็นจำนวน ร้อยละ (**ตารางทางเดียว**)

1.2) วิเคราะห์ข้อมูลโดยการแจกแจงความถี่ เป็นจานวน ร้อยละ เพื่อดูความสัมพันธ์ ระหว่างตัวแปรหรือเปรียบเทียบค่าสัดส่วน (อัตรา) ตามตัวแปรแบ่งกลุ่ม (**ตารางทางสองทาง หรือ ตารางไขว้**)

2. กรณีข้อมูลต่อเนื่อง (ตัวแปรเชิงปริมาณ เป็นตัวเลขเชิงปริมาณหรือ มีระดับการวัดเป็น Interval or Ratio Scale) เช่น อายุ รายได้ ประสบการณ์ในการทางาน น้ำหนัก ส่วนสูง BMI เป็นต้น

2.1) วิเคราะห์ข้อมูลโดยการคำนวณค่าการวัดแนวโน้มเข้าสู่ส่วนกลางและค่าการกระจายของข้อมูล

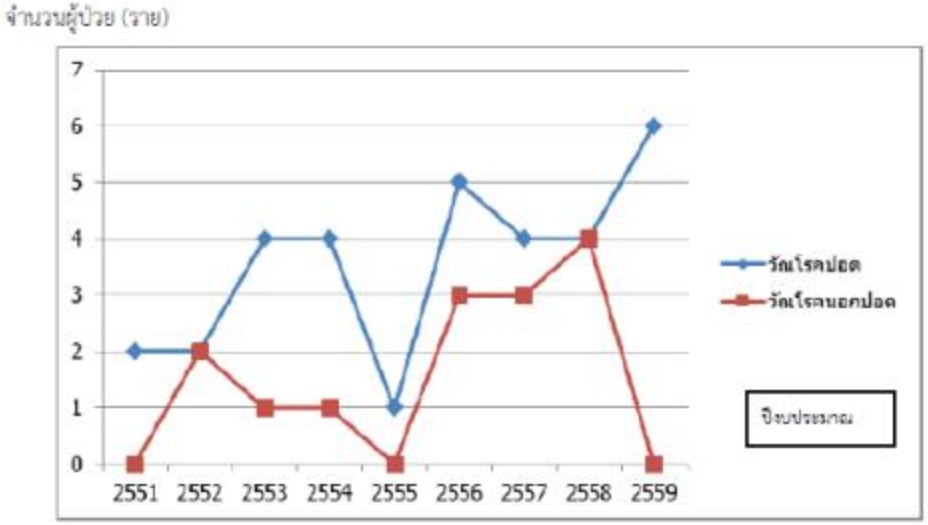
- หากข้อมูลมีการกระจายไม่มาก หรือข้อมูลมีการแจกแจงเป็นแบบโค้งปกติ
  - -> ค่าเฉลี่ย (Mean) คู่กับส่วนเบี่ยงเบนมาตรฐาน (S.D.)
- ในกรณีที่ข้อมูลมีการกระจายมาก (คือ 2*S.D.>Mean) หรือข้อมูลมีการแจกแจงไม่เป็นแบบปกติ
  - -> มัธยฐาน (Median) คู่กับพิสัยของควอไทล์ (Inter Quartile Range : IQR*)

*IQR =Q3-Q1

ตัวอย่าง 1 ประสิทธิผลการบำบัดรักษาและฟื้นฟูสมรรถภาพ ผู้เสพยาเสพติดในพื้นที่ตำบลนาแขม อำเภอเมือง จังหวัดเลย

ตารางที่ 2 รายละเอียดการใช้สารเสพติดก่อนเข้ารับการบำบัดรักษา (n = 261)

ตัวแปร	จำนวน	ร้อยละ
อายุที่ใช้สารเสพติดครั้งแรก		
น้อยกว่า 20 ปี	146	55.90
20 - 40 ปี	95	36.40
41 ปี ขึ้นไป	20	7.70
ประเภทสารเสพติดที่ใช้ครั้งแรก		
ยาบ้า	235	90.00
กัญชา	23	8.80
สารระเหย	2	0.80
อื่นๆ	1	0.40



การวิเคราะห์ข้อมูลโดยใช้สถิติอ้างอิงหรืออนุมาน (Inferential Statistics)

แนวทางการใช้สถิติเชิงอนุมานเพื่อทดสอบสมมติฐาน

การทดสอบสมมติฐานเป็นวิธีการของสถิติเชิงอนุมานที่เกี่ยวข้องกับการตัดสินใจสรุปลักษณะของประชากรที่สนใจ โดยใช้ข้อมูลจากตัวอย่าง มีขั้นตอนดังนี้

1.) การตั้งสมมติฐานทางสถิติ

2.) การกำหนดระดับนัยสำคัญทางสถิติ (α)

3.) การเลือกสถิติในการทดสอบ

4.) การหาค่าวิกฤติ/หาค่า p value

5.) ตัดสินใจและสรุปผลการทดสอบ

1.การประมาณค่าพารามิเตอร์(Parametric Estimation)

1.1 การประมาณค่าแบบจุด(Point Estimation)

เป็นการประมาณค่าพารามิเตอร์ (Parameter) เพื่อสรุปหรือบรรยาย ลักษณะของประชากรด้วยสถิติ (Statistic) ที่ได้จากคานวณข้อมูลจากตัวอย่างที่ศึกษาด้วยค่าตัวเลขเพียง ตัวเดียว ซึ่งในทางปฏิบัติไม่นิยมกระทำ เพราะไม่สามารถบอกโอกาสความผิดพลาดในการสรุปได้ เช่น การประมาณค่า μ จากค่า $\bar{x}$ ซึ่งในทางปฏิบัติจะเห็นได้ว่าค่า $\bar{x}$ ที่คำนวณได้จากกลุ่มตัวอย่าง อาจมีค่า ไม่เท่ากันพอดีกับค่า μ หากสามารถศึกษาทั้งประชากร ดังนั้นจึงทำการหาขอบเขต (Interval) ของค่า μ ผลทำให้เกิดการประมาณค่าพารามิเตอร์แบบช่วง (Interval Estimation)

1.2 การประมาณค่าแบบช่วง (Interval Estimation)

เป็นการประมาณค่าพารามิเตอร์ (Parameter) เพื่อสรุปหรือบรรยายลักษณะของประชากรด้วยสถิติ (Statistic) ที่ได้จากคำนวณข้อมูลจากตัวอย่างที่ศึกษา โดยวิธีการกำหนดขอบเขตหรือช่วงที่จะเป็นไปได้ของค่าพารามิเตอร์นั้น พร้อมทั้งระบุความเชื่อมั่นในการสรุปด้วยช่วงเชื่อมั่นเท่าใด หรือด้วยโอกาสความผิดพลาดของการสรุปผิดมากน้อยเพียงใด โอกาสผิดพลาดในการประมาณค่าจะแสดงด้วยค่า α (Alpha) หรือแสดงด้วยช่วงเชื่อมั่น เป็น (1- α)100 %

ระดับความเชื่อมั่น (Level of Confidence Interval) คือ โอกาสหรือความน่าจะเป็นที่พารามิเตอร์ของประชากรจะอยู่ในช่วงของค่าที่ประมาณได้ ที่ระดับความเชื่อมั่น (1- α) 100% หรือโอกาสความผิดพลาดในการสรุป (α) 100%

ปัจจัยที่มีผลต่อช่วงเชื่อมั่น

ช่วงเชื่อมั่นที่แคบจะมีความแม่นยำของการประมาณค่าสูง

1.) ระดับความเชื่อมั่น (1- α) x 100%

ระดับความเชื่อมั่นที่สูงกว่าช่วงของการประมาณค่าจะกว้างกว่าที่ระดับความเชื่อมั่นที่ต่ำกว่า

เช่น ที่ช่วงเชื่อมั่น 99 % ช่วงของการประมาณค่าจะกว้างกว่า ที่ช่วงเชื่อมั่น 90 %

2.) ความคาดเคลื่อนมาตรฐานของตัวประมาณค่า

2.1) ส่วนเบี่ยงเบนมาตรฐาน (การกระจายของข้อมูล)

ข้อมูลกระจายมาก ๆ (S.D.มาก) ช่วงของการประมาณค่าจะกว้างกว่าข้อมูลกระจายน้อย ๆ (S.Dน้อย)

2.2) ขนาดตัวอย่าง

ขนาดตัวอย่าง(n)มาก ๆ จะทำให้ความคาดเคลื่อนมาตรฐานของตัวประมาณค่าน้อยลง ทำให้ความแม่นยาในการประมาณค่ามากขึ้น ดังนั้นช่วงของการประมาณค่าจะแคบกว่าขนาดตัวอย่างน้อย ๆ ซึ่งทำให้ ความแม่นยาในการประมาณค่าน้อย

2.1. การเปรียบเทียบความแตกต่างของค่าเฉลี่ย

(หรือเปรียบเทียบค่ามัธยฐาน กรณีข้อมูลแจกแจงไม่เป็นโค้งแบบปกติ จะใช้สถิติ Non Parametric Statistics)

2.1.1) กรณีเปรียบเทียบค่าเฉลี่ย 1 กลุ่มกับเกณฑ์หรือค่ามาตรฐาน ใช้สถิติ One Sample t- test

2.1.2) กรณีเปรียบเทียบค่าเฉลี่ย (*หรือ ค่ามัธยฐาน*) 2 กลุ่ม

2.1.2.1) **กรณี 2 กลุ่ม เป็นอิสระต่อกัน** เช่น เปรียบเทียบความแตกต่างของค่าเฉลี่ยดัชนีมวลกายระหว่างเพศชายกับเพศหญิง ใช้สถิติ

- ❑ Independent t - test -- >Parametric Statistics
- ❑ *Mann-Whitney U Test -- >Non Parametric Statistics*

หรือ Wilcoxon Rank sum Test-- >Non Parametric Statistics

2.1.2.2) **กรณี 2 กลุ่ม ไม่เป็นอิสระต่อกัน** เช่น เปรียบเทียบค่าเฉลี่ยของความแตกต่างดัชนีมวลกายก่อนกับหลังการเข้าร่วมโครงการออกกำลังกายเพื่อสุขภาพ ใช้สถิติ

- ❑ Paired t- test -- >Parametric Statistics
- ❑ *Wilcoxon (Matched Paired) Signed Rank Test -- >Non Parametric Statistics*

2.1.3) **กรณีเปรียบเทียบค่าเฉลี่ย 3 กลุ่มขึ้นไปที่เป็นอิสระต่อกัน** เช่น เปรียบเทียบความแตกต่างของค่าเฉลี่ยดัชนีมวลกาย ระหว่างตำแหน่งพยาบาลวิชาชีพ นักวิชาการ และเจ้าพนักงานสาธารณสุขชุมชน ใช้สถิติ

- F-test (One Way ANOVA) -- >Parametric Statistics
  หากพบว่ามีความแตกต่างกันใช้สถิติทดสอบรายคู่ (Multiple Comparisons) โดยใช้สถิติ LSD หรือ Bonferroni หรือ Scheffe ฯลฯ
- Kruskal Wallis Test -- >Non Parametric Statistics

2.2 การหาความสัมพันธ์

2.2.1) **กรณีตัวแปร 2 ตัวแปรที่นำมาหาความสัมพันธ์กันเป็นตัวแปรต่อเนื่อง (เชิงปริมาณ)**
เช่น การหาความสัมพันธ์ระหว่างน้ำหนักกับระดับคลอเลสเตอรอล

- ใช้สถิติสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน (Person Product Moment Correlation Coefficient; r) -- >Parametric Statistics

2.2.2) **กรณีตัวแปร 2 ตัวแปรที่นำมาหาความสัมพันธ์กันเป็นข้อมูลอันดับ** (Ordinal scale)
หรือ **กรณีตัวแปรต่อเนื่องที่มีการแจกแจงไม่เป็นแบบโค้งปกติ หรือ จานวนตัวอย่างน้อย**

- ใช้ สัมประสิทธิ์สหสัมพันธ์ของสเปียร์แมน (Spearman Correlation Coefficient; rs) -- >*Non Parametric Statistics*

2.2.3) **กรณีตัวแปร 2 ตัวแปรที่นามาหาความสัมพันธ์กันเป็นตัวแปรแจงนับ (เชิงคุณภาพ) ที่เป็นอิสระต่อกัน** เช่น การหาความสัมพันธ์ระหว่างเพศกับการสวมหมวกนิรภัย ความสัมพันธ์ระหว่างอาชีพกับการดื่มเครื่องดื่มบำรุงกาลัง

- ใช้สถิติไคว์สแคร์ (Chi-Square test ; χ^2) หรือ
  Fisher's กรณีที่ค่าคาดหวัง (Expected Value) ในแต่ละเซลล์มีค่าน้อยกว่า 5 เกิน 20% ของจำนวนเซลล์ทั้งหมด -- >Non Parametric Statistics

2.2.4) **กรณีตัวแปร 2 ตัวแปรที่นามาหาความสัมพันธ์กันเป็นตัวแปรแจงนับ ที่ไม่เป็นอิสระต่อกัน** เช่น การหาความสัมพันธ์ของการสูบบุหรี่กับการอบรม ลด ละ เลิก การสูบบุหรี่ หรือ การเปรียบเทียบอัตราการสูบบุหรี่ก่อนกับหลังการอบรม ลด ละ เลิก การสูบบุหรี่

- ใช้สถิติ McNemar Chi-Square test -- >*Non Parametric Statistics*

2.3 การวิเคราะห์ความสัมพันธ์ กรณีตัวแปรต้นหลายตัวแปรพร้อมกัน (Multivariate Analysis)

เพื่อหาปัจจัยที่ร่วมกันมีความสัมพันธ์ หรือร่วมกันทำนายหรือร่วมกันพยากรณ์ เป็นการทดสอบว่ามี**ตัวแปรต้น**อะไรบ้าง (หลายตัวที่ร่วมกัน) ที่มี**ความสัมพันธ์** (มีผล/มีอิทธิพล/ร่วมกันทานาย/พยากรณ์)**ตัวแปรตาม**

2.3.1) **กรณีตัวแปรตามเป็นข้อมูลต่อเนื่อง**

เช่น ปัจจัยที่มีความสัมพันธ์ (ปัจจัยที่มีผล) ต่อระดับดัชนีมวลกาย

- ❑ ใช้สถิติ Multiple Linear Regression

2.3.2) **กรณีตัวแปรตามเป็นข้อมูลแจงนับ**

ปัจจัยที่มีความสัมพันธ์ (ปัจจัยที่มีผล) ต่อการรักษาหายขาดของผู้ป่วยวัณโรคที่ใช้ระบบ DOT

- ❑ ใช้สถิติ Multiple Logistic Regression

สรุปการเลือกใช้สถิติทดสอบสมมติฐาน (1)

รายการ	Parametric Statistics	Non-Parametric Statistics
1.การแจกแจงของข้อมูล	แบบปกติ (Normal Distribution)	ไม่มีข้อกาหนด (Free Distribution)
2.ระดับการวัดตัวแปรตาม	Interval Scale หรือ Ratio Scale	ใช้ได้ทุกระดับ Nominal Scale , Ordinal Scale, Interval Scale หรือ Ratio Scale
3.ตัวอย่าง	เชิงสุ่ม(ได้จากการสุ่ม)	เชิงสุ่ม(ได้จากการสุ่ม)
4.ขนาดตัวอย่าง	ใหญ่	เล็ก

อ้างอิง

- กรมวิทยาศาสตร์การแพทย์. (2564). *สถิติพื้นฐานสำหรับงานวิจัย.* สืบค้นจาก https://nih.dmsc.moph.go.th/data/data/64/activity/10_8_64/WW_v1.pdf