บทที่ 1
สถิติกับการวิจัย

สถิติเข้ามาเกี่ยวข้องกับกระบวนการวิจัยหลายขั้นตอน ตัวอย่างเช่น การคำนวณขนาดตัวอย่าง การสุ่มตัวอย่าง การสร้างเครื่องมือและหาคุณภาพของเครื่องมือในการเก็บรวบรวมข้อมูล การวิเคราะห์ข้อมูลเพื่อบรรยายลักษณะกลุ่มที่ศึกษา การวิเคราะห์ข้อมูลเพื่อตอบคำถามการวิจัยหรือสรุปผลการศึกษาตามวัตถุประสงค์ การทดสอบสมมติฐานการวิจัย การนำเสนอและสรุปผลการวิจัย ใช้เป็นเครื่องมือในการตัดสินและสรุปผลการวิจัย เพื่อให้เข้าใจสถิติมากยิ่งขึ้นและนำความรู้สถิติไปใช้ในงานวิจัยได้อย่างเหมาะสม ผู้วิจัยควรมีความรู้เบื้องต้นเกี่ยวกับสถิติดังต่อไปนี้

1. ความหมายของสถิติ (Statistics)

โดยทั่วไปสถิติมีความหมาย 2 นัย คือ ความหมายของตัวเลขข้อมูลสถิติ และสถิติที่เป็นศาสตร์ ดังนี้

1.) สถิติที่เป็นตัวเลขหรือข้อมูลสถิติ (Statistics Data)

ข้อมูลสถิติ (Statistics Data) หมายถึง ตัวเลขที่ได้จากการรวบรวมข้อมูล นำมาประมวล วิเคราะห์และสรุปเพื่อใช้บรรยายหรืออธิบายเหตุการณ์ ปรากฏการณ์หรือ ตัวแปรต่าง ๆ ที่เราสนใจ เช่น สถิติการเจ็บป่วย 10 อันดับแรกของผู้มารับบริการในแต่ละปีของสถานบริการ อัตราการเกิดอุบัติเหตุในช่วง 10 วันอันตรายในช่วงสงกรานต์ เป็นต้น

2.) สถิติที่เป็นศาสตร์ (Statistics)

หมายถึง วิชาที่ว่าด้วยการจัดกระทำต่างๆ เกี่ยวกับข้อมูลเพื่อให้สามารถบรรยายลักษณะของสิ่งที่ศึกษา ได้แก่ การเก็บรวบรวมข้อมูล การวิเคราะห์ข้อมูล การนำเสนอข้อมูล และการแปลความหมายของข้อมูล รวมถึงการนำเอาข้อมูลที่รวบรวมได้จากตัวอย่างไปใช้ในการคาดคะเนและตัดสินใจต่าง ๆ ที่เกี่ยวกับประชากร

2. องค์ประกอบของสถิติ

สถิติประกอบด้วย 2 องค์ประกอบ ดังนี้

1.) สถิติเชิงพรรณนา (Descriptive Statistics)

เป็นวิธีการทางสถิติเพื่อใช้ในการพรรณนาหรือบรรยายลักษณะของสิ่งที่ศึกษา เพื่อให้เข้าใจถึงลักษณะของข้อมูลที่เก็บรวบรวมมาได้ จะพรรณนาภายในขอบเขตของข้อมูลที่เก็บรวบรวมมาเท่านั้น ไม่สามารถจะคาดคะเนลักษณะต่างๆ ออกไปนอกเหนือจากข้อมูลที่มีอยู่ได้ หรือไม่มีการอ้างอิงหรืออนุมานไปถึงกลุ่มอื่น หากผู้วิจัยสามารถศึกษาทุกหน่วยของประชากรได้ ก็จะใช้สรุปหรือบรรยายลักษณะของประชากรที่ศึกษา หากผู้วิจัยไม่สามารถศึกษาสมาชิกทุกหน่วยของประชากรได้ สุ่มสมาชิกเพียงบางส่วนมาศึกษา(ศึกษาจากกลุ่มตัวอย่าง) ก็จะใช้สรุปหรือบรรยายเฉพาะกลุ่มตัวอย่างที่สุ่มมาศึกษาเท่านั้น

R2R สำนักการแพทย์ กรุงเทพมหานคร ปรับปรุงล่าสุด ปี 2560

2.) สถิติเชิงอ้างอิงหรือสถิติอนุมาน (Inferential Statistics)

เป็นวิธีการทางสถิติที่ใช้ในการสรุปลักษณะของประชากร จากผลการศึกษาข้อมูลในกลุ่มตัวอย่าง โดยอาศัยทฤษฎีความน่าจะเป็นในการอนุมานลักษณะประชากร โดยศึกษาจากกลุ่มตัวอย่าง แต่อ้างอิงหรืออนุมานไปถึงประชากร ประกอบด้วย 2 องค์ประกอบ คือ การประมาณค่าพารามิเตอร์ (Parametric Estimation) และการทดสอบสมมติฐาน (Testing Hypothesis)

3. ความหมายของพารามิเตอร์ (Parameter)และค่าสถิติ(Statistic)

1.) พารามิเตอร์ (Parameter) เป็นค่าตัวเลขที่คำนวณได้จากข้อมูลประชากร เพื่อใช้สรุป หรือบรรยายแสดงลักษณะของประชากร เช่น ค่าเฉลี่ยของประชากร แทนด้วยสัญลักษณ์ μ (มิว) ค่าส่วนเบี่ยงเบนมาตรฐานของประชากร แทนด้วยสัญลักษณ์ σ (ซิกม่า) เป็นต้น

2.) ค่าสถิติ (Statistic) เป็นค่าตัวเลขที่คำนวณได้จากข้อมูลของกลุ่มตัวอย่าง เพื่อใช้สรุป หรือบรรยายแสดงลักษณะของกลุ่มตัวอย่าง เช่น ค่าเฉลี่ยของกลุ่มตัวอย่าง แทนด้วยสัญลักษณ์ $\bar{x}$ (เอ็กซ์บาร์) ค่าส่วนเบี่ยงเบนมาตรฐานของของกลุ่มตัวอย่าง แทนด้วยสัญลักษณ์ S.D. (เอสดี) เป็นต้น

ตารางที่ 1.1 เปรียบเทียบสัญลักษณ์ของพารามิเตอร์และค่าสถิติ

ค่าที่ใช้บรรยายลักษณะ	พารามิเตอร์ (ประชากร)	ค่าสถิติ (ตัวอย่าง)
1. ค่าเฉลี่ยเลขคณิต (Mean)	μ (มิว)	$\bar{x}$ (เอ็กซ์บาร์)
2. ส่วนเบี่ยงเบนมาตรฐาน (Standard deviation)	σ (ซิกม่า)	s หรือ S.D.
3. ความแปรปรวน (Variance)	σ^2	s^2 หรือ $S.D.^2$
4. สัดส่วน (Proportion)	π (พาย) หรือ P	p
5. สัมประสิทธิ์สหสัมพันธ์ (Correlation Coefficient)	ρ (โร)	r
6. สัมประสิทธิ์การถดถอย (Regression Coefficient)	β (เบต้า)	b
6. ค่าอื่นๆ	θ (เซต้า)	$\hat{\theta}$

4. ระดับการวัด (Scales of Measurement or Levels of Measurement)

ระดับการวัดของตัวแปรหรือข้อมูลมีประโยชน์ในการสรุป นำเสนอและวิเคราะห์ข้อมูล ดังนั้นจึงมีความจำเป็นที่จะต้องเข้าใจถึงระดับการวัดของตัวแปรหรือข้อมูลเพื่อสรุป นำเสนอและวิเคราะห์ข้อมูลได้เหมาะสม ถูกต้อง ระดับการวัดของข้อมูลแบ่งเป็น 4 ระดับดังนี้

1.) มาตรานามบัญญัติ (Nominal Scale)

การวัดข้อมูลในระดับนี้เป็นเพียงการเรียกชื่อ หรือจัดประเภท แบ่งกลุ่ม แบ่งพวก (Classification) ชนิดของสิ่งต่าง ๆ เท่านั้นเอง จะแสดงให้เห็นเพียงความแตกต่างของสิ่งต่างๆ เท่านั้น

ตัวอย่าง- เพศ แบ่งเป็น หญิง และ ชาย
- กลุ่มเลือด แบ่งเป็น A B AB และ O

R2R สำนักงานสาธารณสุขจังหวัดขอนแก่น ปี 2560

- สถานภาพสมรส แบ่งเป็น โสด คู่ หม้าย หย่า และ แยก
- ศาสนา แบ่งเป็น พุทธ คริสต์ อิสลาม และอื่นๆ ฯลฯ
- บ้านเลขที่ เบอร์โทรศัพท์ หมายเลขหลังเสื้อนักกีฬา ฯลฯ

ลักษณะการวัดระดับนี้

- ไม่สามารถบอกปริมาณว่ามากน้อย จัดลำดับสูงต่ำระหว่างกลุ่มได้
- ไม่สามารถนำมาบวก ลบ คูณ หรือ หาร ระหว่างกลุ่มได้ แม้จะให้ค่าสังเกตเป็นตัวเลข เช่น บ้านเลขที่ เบอร์โทรศัพท์ เพราะตัวเลขดังกล่าวเป็นเพียงการกำหนด สัญลักษณ์ ไม่มีความหมายในเชิงปริมาณ
- มีข้อจำกัดในการวิเคราะห์ทางสถิติ ใช้ได้บางวิธีเท่านั้น
- ถ้ามีค่าสังเกตเพียงสองค่า เรียก ตัวแปรทวินาม (Binomial หรือ Dichotomous Variable) เช่น เพศ (ชาย และหญิง) การเจ็บป่วย (ป่วย และไม่ป่วย) ผลการรักษา (หาย และไม่หาย) เป็นต้น

2.) มาตราเรียงอันดับ (Ordinal Scale)

สามารถจัดลำดับ (Ranking) ของข้อมูลได้ว่ามาก น้อย สูงหรือต่ำอย่างไร เพิ่มจากระดับมาตรานามบัญญัติ แต่ไม่สามารถบอกความแตกต่างระหว่างลำดับได้ว่าต่างกันปริมาณเท่าใด

ตัวอย่าง - ความเจ็บปวด แบ่งเป็น ปวดมาก ปานกลาง น้อย

- ความคิดเห็น แบ่งเป็น เห็นด้วย เฉย ๆ ไม่เห็นด้วย
- ความพึงพอใจ แบ่งเป็น พอใจมาก พอใจน้อย ไม่พอใจ ฯลฯ
- ลำดับที่ของการประกวด แบ่งเป็น รางวัลที่ 1 ที่ 2 ที่ 3....
- ลำดับที่ของความสูง

ลักษณะของระดับการวัดแบบนี้

- จัดเรียงลำดับในแต่ละกลุ่ม แต่ละพวกได้
- ไม่สามารถบอกปริมาณช่วงห่างระหว่างลำดับได้ ตัวเลขจะแทนลำดับที่หรือลำดับเท่านั้น ไม่แทนปริมาณของความแตกต่าง
- ข้อมูลที่ได้ไม่สามารถนำมา บวก ลบ คูณ หรือหารกันได้

3.) มาตราอันตรภาคหรือช่วง (Interval Scale)

สเกลการวัดที่กำหนดค่าของตัวแปรเป็นตัวเลขในลักษณะต่อเนื่องแต่มีจุดศูนย์ไม่แท้ (Arbitrary Zero or Relative Zero) บอกความแตกต่างระหว่างช่วงของตัวแปรได้ โดยแต่ละหน่วยการวัดจะมีระยะห่างเท่า ๆ กัน

ตัวอย่าง อุณหภูมิ I.Q. คะแนนความรู้ คะแนนทัศนคติ คะแนนการปฏิบัติ ฯลฯ

ลักษณะของระดับการวัดแบบนี้

- วัดออกมาเป็นตัวเลขเชิงปริมาณ
- บอกความแตกต่างของแต่ละช่วงได้
- แต่ละหน่วยการวัดมีระยะห่างเท่าๆ กัน
- มีศูนย์ (0) ไม่แท้จริง ที่จุดศูนย์ยังมีค่าอยู่

คลังข้อสอบ สำนักงานสาธารณสุขจังหวัดพิษณุโลก ปี 2560

4.) มาตราอัตราส่วน (Ratio Scale)

สเกลการวัดที่กำหนดค่าของตัวแปรเป็นตัวเลขในลักษณะต่อเนื่องแต่แตกต่างจากระดับการวัด Interval Scale ที่มีจุดศูนย์แท้จริง (Absolute Zero) คือ ค่าศูนย์มีความหมายชัดเจนว่าไม่มีค่า ความแตกต่างระหว่างช่วงของตัวแปรมีขนาดเท่ากัน

ตัวอย่าง น้ำหนัก ส่วนสูง ความดันโลหิต จำนวนผู้มารับบริการ ฯลฯ

ลักษณะของระดับการวัดแบบนี้

- วัดออกมาเป็นตัวเลขเชิงปริมาณ
- บอกความแตกต่างของแต่ละช่วงได้
- แต่ละหน่วยการวัดมีระยะห่างเท่าๆ กัน
- มีศูนย์ (0) ที่แท้จริง ที่จุดศูนย์ไม่มีค่า
- ค่าที่ได้สามารถนำมา บวก ลบ คูณ หาร ถอดราก หรือยกกำลังได้ จึงสมารถนำมาวิเคราะห์โดยใช้วิธีการทางสถิติได้ทุกอย่าง

การวัดตัวแปรเดียวกันอาจมีระดับการวัดที่แตกต่างกันได้ ขึ้นอยู่กับความละเอียดและวิธีการเก็บรวบรวมข้อมูล เช่น การสูบบุหรี่ ถ้าเก็บข้อมูลเป็น สูบ กับ ไม่สูบ จะมีระดับการวัดเป็น Nominal Scale ถ้าเก็บข้อมูลเป็นจำนวนบุหรี่ที่สูบต่อวัน จะมีระดับการวัดเป็น Ratio Scale

ในการวิเคราะห์ข้อมูล ในบางกรณีอาจมีการลดระดับการวัดของตัวแปรลง เพื่อสื่อสารให้คนทั่วไปเข้าใจ เช่น เก็บข้อมูลโดยการวัดระดับความดันโลหิตออกมาเป็นตัวเลข ซึ่งมีระดับการวัดเป็น Ratio Scale เวลาสรุปผลการวิจัย อาจจัดกลุ่มข้อมูลให้เป็น Ordinal Scale คือ กลุ่มที่มีความดันโลหิตต่ำ กลุ่มที่มีความดันโลหิตปกติ และ กลุ่มที่มีความดันโลหิตสูง ในทางตรงกันข้ามกรณีที่ผู้วิจัยต้องการนำข้อมูลไปวิเคราะห์สถิติขั้นสูง เช่น ตัวแปรความคิดเห็น มีระดับการวัดเป็น Ordinal Scale ผู้วิจัยอาจกำหนดให้ค่าคะแนน เป็น Rating Scale ตามแนวคิดของ Liker's Scale เพื่อปรับตัวแปรให้มีระดับการวัดที่สูงขึ้นเป็น Interval Scale โดยกำหนดให้ เห็นด้วยอย่างยิ่ง มีค่าเท่ากับ 5 คะแนน เห็นด้วย มีค่าเท่ากับ 4 คะแนน เฉยๆ มีค่าเท่ากับ 3 คะแนน ไม่เห็นด้วย มีค่าเท่ากับ 2 คะแนน ไม่เห็นด้วยอย่างยิ่ง มีค่าเท่ากับ 1 คะแนน ก็อาจทำได้แต่ต้องกำหนดการแปลผลไว้ให้ชัดเจน ซึ่งการเพิ่มระดับการวัดจากระดับต่ำไปสูงถือว่าไม่ค่อยเหมาะสม

5. ตัวแปร (Variable)และข้อมูล(Data)

ตัวแปร(Variables) หมายถึง คุณลักษณะของสิ่งที่เราสนใจศึกษา อาจวัดเป็นตัวเลขเชิงปริมาณ เช่น น้ำหนัก ส่วนสูง อายุ เป็นต้น หรือ บอกคุณลักษณะเชิงคุณภาพ เช่น เพศ สถานภาพสมรส มีค่าที่แปรเปลี่ยนไปตามหน่วยตัวอย่างที่ศึกษา

ค่าที่เป็นตัวเลขหรือคุณลักษณะของตัวแปร เรียกว่า **ข้อมูล(Data)** ได้จากการสังเกต ชั่ง ตวง วัด นับ หรือสอบถามจากหน่วยตัวอย่างที่ศึกษา โดยที่หน่วยที่ศึกษาอาจเป็นคน สัตว์ พืช และสิ่งของ เมื่อหน่วยศึกษามีลักษณะที่แตกต่างกัน ข้อมูลที่ได้จึงแตกต่างกัน เรียกคุณลักษณะหรือค่าที่แตกต่างกันว่าตัวแปร เช่น เพศ แบ่งเป็นเพศหญิงและเพศชาย กลุ่มเลือด แบ่งเป็น กลุ่มเลือด A B AB และ O น้ำหนักที่แตกต่างกัน ส่วนสูงที่แตกต่างกัน เป็นต้น

HPP สำนักงานหลักประกันสุขภาพแห่งชาติ อบรมเชิงปฏิบัติการ ปี 2560

ประเภทของตัวแปรตามลักษณะของข้อมูลแบ่งออกได้ 2 ประเภท คือ

1.) ตัวแปรเชิงคุณภาพ (Qualitative Variable) หรือตัวแปรแจงนับ (Categorical Variable)

หมายถึง ตัวแปรที่แสดงลักษณะเชิงคุณภาพของหน่วยสังเกต โดยบ่งบอกหรือบรรยายลักษณะของตัวแปรนั้น เช่น ตัวแปรสถานภาพสมรส จะแสดงลักษณะ โสด คู่ หม้าย หย่า และแยก ตัวแปรเพศ จะแสดงลักษณะเพศชาย เพศหญิง เป็นต้น ต้องทำการแจงข้อมูลออกตามคุณลักษณะเป็นกลุ่ม หรือ พวกที่มีลักษณะเหมือนกันก่อน แล้วค่อยนับความถี่ออกมาเป็นจำนวน จึงเรียกว่าตัวแปรแจงนับ จะสรุปลักษณะของตัวแปรแบบนี้ โดยการแจกแจงความถี่ เป็นจำนวน คำนวณค่าสัดส่วน(Proportion) ค่าร้อยละ(Percentage) หรือ ค่าอัตรา(Rate) อัตราส่วน(Ratio)

2.) ตัวแปรเชิงปริมาณ (Quantitative Variable) หรือตัวแปรต่อเนื่อง (Continuous Variable)

หมายถึง ตัวแปรที่แสดงลักษณะของหน่วยสังเกตออกมาเป็นตัวเลขที่บอกขนาดหรือปริมาณได้ เช่น อายุ น้ำหนัก ส่วนสูง ระดับความดันโลหิต เป็นต้น ตัวแปรเชิงปริมาณ ได้จากการชั่ง ตวง วัด หรือนับ จะสรุปลักษณะของตัวแปรแบบนี้ โดยการคำนวณการวัดแนวโน้มเข้าสู่ส่วนกลาง และค่าการกระจายของข้อมูล อาจจัดกลุ่มตัวแปรเหล่านี้เป็นช่วง แล้วค่อยแจกแจงความถี่ คำนวณค่าสัดส่วน ค่าร้อยละ หรือ ค่าอัตรา ก็ได้

ตารางที่ 1.2 สรุป ระดับการวัดประเภทของตัวแปร

ระดับการวัด	ประเภทของตัวแปร	Descriptive Statistics	Graph /Chart
Nominal Scale Ordinal Scale	Qualitative Variable or Categorical Variable	Frequency and percentage	Pie, Bar Chart
Interval Scale Ratio Scale	Quantitative Variable or Continuous Variable	การวัดแนวโน้มเข้าสู่ส่วนกลาง และค่าการกระจาย $\bar{x}$ และ S.D. Median และ Min., Max. (Rang) Median และ Q1., Q3 (Inter Quartile Rang; IQR)	Histogram, Box plot, Dot Plot, Stem and leave
จัดกลุ่ม/ช่วง	Qualitative Variable or Categorical Variable	Frequency and percentage	Pie, Bar Chart

R2R สำนักงานสาธารณสุข มหาวิทยาลัย ปี 2560

6. การวิเคราะห์ข้อมูลในงานวิจัย

การวิเคราะห์ข้อมูลในงานวิจัย ประกอบด้วยการใช้สถิติ 2 องค์ประกอบใหญ่ ๆ ได้แก่ สถิติเชิงพรรณนา (Descriptive Statistics) โดยค่าสถิติต่างๆที่คำนวณได้จากข้อมูลตัวอย่างที่ศึกษา และ สถิติเชิงอนุมาน (Inferential Statistics) โดยนำเสนอในลักษณะของช่วงเชื่อมั่น :Confidence Interval ; CI เป็นค่าการประมาณช่วงเชื่อมั่นของค่าพารามิเตอร์จากค่าสถิติที่ได้จากข้อมูลตัวอย่างที่ศึกษา และค่า p-value ที่เป็นผลจากการทดสอบสมติฐานเพื่ออธิบายหรือสรุปลักษณะของประชากร โดยอาศัยค่าสถิติที่ได้

จากข้อมูลตัวอย่าง สรุปว่ามีนัยสำคัญทางสถิติหรือไม่มีนัยสำคัญทางสถิติ) ในงานวิจัย ควรมีครบทั้งสถิติเชิงพรรณนาและช่วงเชื่อมั่น ส่วน ค่า p-value จะมีเฉพาะงานวิจัยที่มีการทดสอบสมมติฐานเท่านั้น ซึ่งรายละเอียดของการเลือกใช้การวิเคราะห์ข้อมูลมีดังนี้

1.) การวิเคราะห์ข้อมูลโดยใช้สถิติเชิงพรรณนา (Descriptive Statistics)

เป็นสถิติที่ใช้บรรยาย หรืออธิบายลักษณะข้อมูลในกลุ่มที่เราศึกษา ซึ่งสามารถใช้ค่าสถิติต่าง ๆ อธิบายได้ขึ้นอยู่กับลักษณะของข้อมูลหรือระดับการวัดของข้อมูลหรือตัวแปรที่เราศึกษา ในการสรุปหรือบรรยายลักษณะของกลุ่มที่ศึกษา อาจใช้สถิติ ดังนี้

(1) การแจกแจงความถี่ (Frequency) คือ การนำข้อมูลที่รวบรวมได้ ซึ่งเป็นข้อมูลดิบ (Row data) มาจัดให้เป็นระเบียบ เป็นหมวดหมู่โดยเรียงจากค่ามากไปหาค่าน้อย เป็นการบรรยายข้อมูลตามความถี่เป็นจำนวนและร้อยละ อาจสร้างตารางหลายรูปแบบ ดังนี้

(1.1) การแจกแจงความถี่ตารางทางเดียว

(1.2) การแจกแจงความถี่ตารางสองทาง หรือตารางไขว้ 2 ตัวแปร

(1.3) การแจกแจงความถี่ตารางสองทาง หรือตารางไขว้ หลายตัวแปร

ตัวอย่างการแจกแจงความถี่ (ตารางทางเดียว)

ตารางที่ 1.3 จำนวน ร้อยละ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน ของกลุ่มตัวอย่างจำแนกตามข้อมูลทั่วไป (n = 120 คน)

ลักษณะทั่วไป	จำนวน	ร้อยละ
1. เพศ		
หญิง	84	70.0
ชาย	36	30.0
2. อายุ(ปี)		
น้อยกว่า 41	11	9.2
41 - 50	37	30.8
51 - 60	51	42.5
61 - 70	21	17.5
ค่าเฉลี่ย (ส่วนเบี่ยงเบนมาตรฐาน): 52.75(8.32)		

กรณีที่ 2xSD. > Mean ควรนำเสนอด้วยค่า Median แทน ค่าเฉลี่ย และ Min., Max. แทนค่า S.D

R2R สำนักงานสาธารณสุขจังหวัดอำนาจเจริญ ปี 2560

ตัวอย่างการแจกแจงความถี่ (ตารางสองทาง)

ตารางที่ 1.4 จำนวน ร้อยละ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน ของกลุ่มตัวอย่างจำแนกตามข้อมูลทั่วไป (n = 120 คน)

ลักษณะของตัวอย่าง	ความเสี่ยงเบาหวาน			
	เสี่ยงเบาหวาน (n=22 คน)		ปกติ (n=98 คน)	
	จำนวน	%	จำนวน	%
1. เพศ				
หญิง	20	(23.8)	64	(76.2)
ชาย	2	(5.6)	34	(94.4)
2. อายุ (ปี)				
น้อยกว่า 41	3	(27.3)	8	(72.7)
41 - 50	12	(32.4)	25	(67.6)
51 - 60	5	(9.8)	46	(90.2)
61 - 70	2	(9.5)	19	(90.5)

2) การคำนวณค่าสัดส่วน (Proportion) ค่าอัตรา (Rate) อัตราส่วน (Ratio)

2.1) สัดส่วน (Proportion: p) คือ การเปรียบเทียบจำนวนย่อยกับจำนวนรวมทั้งหมด โดยมากแปลงเป็นรูปร้อยละหรือ % สูตร $p = \frac{x}{n}$ หรือ $p = \frac{a}{a+b}$

x= จำนวนคนหรือเหตุการณ์ที่เราสนใจ (เป็นส่วนหนึ่งของ n)

n= จำนวนตัวอย่างที่ศึกษา

ค่าสัดส่วนมีค่าอยู่ระหว่าง 0 ถึง 1 เพื่อให้ง่ายต่อการแปลผลและอธิบาย นิยมนำค่าคงที่ 100 มาคูณ เรียกว่าค่า ร้อยละ หรือ %

2.2) อัตรา (Rate) คือ ค่าเปรียบเทียบระหว่างตัวเลขตัวเศษ(Numerator) ซึ่งเป็นส่วนหนึ่งของตัวเลขตัวส่วน(Denominator) ตัวเศษจะต้องเป็นส่วนหนึ่งของตัวส่วนเสมอ เทียบกับช่วงเวลาหนึ่งๆ แต่เพื่อให้เห็นค่าชัดเจนยิ่งขึ้น จึงนำค่าคงที่ (K) ซึ่งอาจเป็น 100, 1,000, 10,000 หรือ 100,000 ก็ได้ตามความเหมาะสมคูณกับอัตรา จะได้ค่ามากกว่า 1 ต่อค่าคงที่นั้นๆ

$$\text{อัตรา (Rate)} = \frac{\text{ตัวเศษ(Numerator)} \times K}{\text{ตัวส่วน(Denominator)}} = \frac{a \times K}{a+b}$$

2.3) อัตราส่วน (Ratio) คือ การเปรียบเทียบค่าตัวเลขของจำนวนหนึ่งกับอีกจำนวนหนึ่ง ซึ่งตัวเศษไม่ได้เป็นส่วนหนึ่งตัวส่วน เช่น จำนวนผู้ป่วยโรคเอดส์เพศชาย (A) ต่อจำนวนผู้ป่วยโรคเอดส์เพศหญิง (B) อัตราส่วนผู้ป่วยโรคเอดส์เพศชายต่อเพศหญิง

R2R โดยมหาวิทยาลัยขอนแก่น ปี 2560

อัตราส่วน (Ratio) = ตัวเศษ (Numerator) = A
ตัวส่วน (Denominator) B

นิยมนำตัวเลขที่น้อยของเศษหรือส่วนไปหารทั้งเศษและส่วน เพื่อทำให้ตัวเลขที่น้อยมีค่าเท่ากับ 1 เพื่อง่ายสำหรับการแปลผล เช่น จำนวนเจ้าหน้าที่เท่ากับ 5 คน จำนวนประชากรที่รับผิดชอบเท่ากับ 5,000 คน ดังนั้น อัตราส่วนเจ้าหน้าที่สาธารณสุขต่อประชากรที่รับผิดชอบ เท่ากับ 5 ต่อ 5,000 คน เอา 5 ไปหารทั้งเศษและส่วนได้ อัตราส่วนเจ้าหน้าที่สาธารณสุขต่อประชากรที่รับผิดชอบเท่ากับ 1 ต่อ 1,000 คน เป็นต้น

3) การวัดแนวโน้มเข้าสู่ส่วนกลาง (Measures of Central Tendency) เป็นการบรรยายด้วยค่ากลางที่เป็นตัวแทนของข้อมูลที่ศึกษา ค่ากลางมีหลายค่าการนำเสนอที่เหมาะสมขึ้นอยู่กับระดับการวัดและการแจกแจงของข้อมูล ดังนี้

3.1) ค่าเฉลี่ยเลขคณิตหรือมัชฌิมเลขคณิต (Arithmetic Mean or Mean or Average; $\bar{x}$)

เป็นค่ากลางที่ได้จากการนำเอาข้อมูลแต่ละตัวของหน่วยสังเกตมารวมกัน แล้วนำผลรวมที่ได้มาหารด้วยจำนวนข้อมูลทั้งหมด ค่าเฉลี่ยของประชากร แทนด้วย สัญลักษณ์ คือ μ(มิว) ส่วนสัญลักษณ์ของค่าเฉลี่ยของกลุ่มตัวอย่าง $\bar{x}$

สูตร ค่าเฉลี่ยของประชากร $\mu = \frac{\sum x_i}{N}$ ค่าเฉลี่ยของกลุ่มตัวอย่าง $\bar{x} = \frac{\sum x_i}{n}$

เมื่อ μ = ค่าเฉลี่ยของประชากร $\bar{x}$ = ค่าเฉลี่ยของกลุ่มตัวอย่าง

X_i = ค่าสังเกตของหน่วยสังเกตแต่ละหน่วยสังเกต(ข้อมูลแต่ละคน)

$\sum X_i$ = ผลรวมของค่าข้อมูลทั้งหมด

N = จำนวนประชากรทั้งหมด n = จำนวนตัวอย่างทั้งหมด

ตัวอย่าง เก็บข้อมูลตัวแปรน้ำหนักของนักศึกษา 10 คน ได้ข้อมูลดังนี้ 50 40 45 45 50 55 50 50 55 60 จงคำนวณหาค่าเฉลี่ยของน้ำหนักนักศึกษา 10 คน

แทนค่าในสูตร $\bar{x} = \frac{\sum x_i}{n}$

$\bar{x}$ = (50 + 40 + 45 + 45 + 50 + 55 + 50 + 50 + 55 + 60)/10

$\bar{x}$ = 500 / 10 = 50 กิโลกรัม

คุณสมบัติของค่าเฉลี่ย

1.) เป็นตัวกลางที่ทุกๆค่าสังเกตมีส่วนร่วมในการกำหนดขนาดของค่าเฉลี่ย

2.) ค่าเฉลี่ยเป็นค่ากลางที่นำข้อมูลทุกตัวมาคำนวณตามหลักคณิตศาสตร์ จึงสามารถใช้ในการวิเคราะห์ทางสถิติขั้นสูงต่อไป

3.) ถ้ามีข้อมูล(ค่าสังเกต)ตัวใดตัวหนึ่งมาก หรือ น้อยเกินไป ค่าเฉลี่ยจะถูกกระทบทำให้เป็นค่ากลางที่ไม่เหมาะสม เช่น 40 45 45 50 50 50 50 55 55 ***90***

4.) ไม่สามารถใช้วัดค่ากลางในกรณีข้อมูลมีระดับการวัดเป็น Nominal Scale และ Ordinal scale

R2R สำนักงานสาธารณสุขจังหวัดขอนแก่น ปี 2560

3.2) มัธยฐาน (Median)

มัธยฐาน คือ ค่าที่มีตำแหน่งอยู่ตรงกลางของข้อมูลที่เรียงลำดับจากค่าน้อยไปมาก หรือ จากค่ามากไปน้อย จะทำให้มีจำนวนข้อมูลครึ่งหนึ่งมีค่าสูงกว่ามัธยฐานและข้อมูลอีกครึ่งหนึ่งมีค่าต่ำกว่ามัธยฐาน

วิธีการหาค่ามัธยฐาน

1.) เรียงข้อมูลจากน้อยไปหามาก หรือ จากมากไปหาน้อย

2.) ในกรณีที่จำนวนข้อมูล(n) เป็นเลขคี่ มัธยฐาน = ค่าของข้อมูลลำดับที่ (n+1)/2

3.) ในกรณีที่จำนวนข้อมูล(n) เป็นเลขคู่

มัธยฐาน = [ค่าของข้อมูลลำดับที่(n/2) + ค่าของข้อมูลลำดับที่(n/2) +1]/2

ตัวอย่าง เก็บข้อมูลตัวแปรน้ำหนักของนักศึกษา 10 คน ได้ข้อมูลดังนี้ 50 40 45 45 50 55 50 50 55 60 จงคำนวณหาค่ามัธยฐานของน้ำหนักนักศึกษา 10 คน

วิธีการหาค่ามัธยฐาน 1.) เรียงข้อมูลจากน้อยไปหามาก หรือ จากมากไปหาน้อย

ลำดับที่	1	2	3	4	5	6	7	8	9	10
น้ำหนัก(กก.)	40	45	45	50	***50***	***50***	50	55	55	60

2.) ในกรณีที่จำนวนข้อมูล(n) เป็นเลขคู่

มัธยฐาน = [ค่าของข้อมูลลำดับที่(n/2) + ค่าของข้อมูลลำดับที่(n/2) +1]/2

ข้อมูลมีจำนวน 10 ตัวอย่าง เป็นเลข คู่ ดังนั้น

มัธยฐาน = [ค่าข้อมูลในลำดับที่ 10/2 + ค่าของข้อมูลลำดับที่(10/2) +1]/2

= [ค่าข้อมูลในลำดับที่ 5 + ค่าของข้อมูลลำดับที่ 6]/2

= [50 +50]/2 = 100/2 =50 กิโลกรัม

คุณสมบัติของค่ามัธยฐาน

1.) มัธยฐานใช้เฉพาะค่าของข้อมูลที่อยู่ตรงกลางมาคำนวณ ดังนั้นข้อมูลที่มากหรือน้อยผิดปกติ จะไม่มีผลกระทบต่อค่ามัธยฐาน

2.) ในกรณีที่การแจกแจงของข้อมูลเบ้ไปทางใดหนึ่ง(เบ้ซ้ายหรือเบ้ขวา) มัธยฐานจะเป็นค่ากลางที่เหมาะสมมากกว่าค่าเฉลี่ยและค่าฐานนิยม

3.) ถ้ามีการเปลี่ยนแปลงข้อมูลบางตัว มัธยฐานจะถูกกระทบน้อยมาก

4.) ไม่สามารถใช้วัดค่ากลางในกรณีข้อมูลมีระดับการวัดเป็น Nominal Scale

5.) ไม่สามารถให้ค่ากลางที่มีความหมายเป็นมัธยฐานได้ ถ้าข้อมูลไปรวมอยู่ที่ค่าสูงสุด หรือ ต่ำสุดมากเกินไป เช่น 45 50 55 65 **65 65** 65 65 65 65

3.3) ฐานนิยม (Mode)

ฐานนิยม คือ ค่าข้อมูลที่มีความถี่มากที่สุด หรือ ค่าซ้ำกันมากที่สุด

วิธีการหาค่าฐานนิยม

1) นำข้อมูลมาเรียงจากน้อยไปมาก หรือมากไปน้อย

2) แล้วดูว่าข้อมูลใดมีค่าซ้ำกันมากที่สุด ค่าข้อมูลนั้น คือ ค่าฐานนิยม

R2R สำนักงานปลัดกระทรวงสาธารณสุข ปีงบประมาณ พ.ศ. 2560

ตัวอย่าง เก็บข้อมูลตัวแปรน้ำหนักของนักศึกษา 10 คน ได้ข้อมูลดังนี้ 50 40 45 45 50 55 50 50 55 60จงคำนวณหาค่ามัธยฐานของน้ำหนักนักศึกษา 10 คน

วิธีการหาค่าฐานนิยม

1)นำข้อมูลมาเรียงจากน้อยไปมาก หรือมากไปน้อย

ลำดับที่	1	2	3	4	5	6	7	8	9	10
น้ำหนัก(กก.)	40	45	45	**50**	**50**	**50**	**50**	55	55	60

2) แล้วดูว่าข้อมูลใดมีค่าซ้ำกันมากที่สุด ดังนั้น ค่าฐานนิยม ตามตัวอย่าง คือ 50 กิโลกรัม

คุณสมบัติของค่าฐานนิยม

1.) หาค่าง่าย รวดเร็ว

2.) ใช้หาค่ากลางในกรณีที่ข้อมูลมีระดับการวัดเป็น Nominal Scale

3.) ถ้ามีการเปลี่ยนแปลงข้อมูลบางตัว หรือ ข้อมูลที่มีค่ามากหรือน้อยผิดปกติจะไม่กระทบต่อฐานนิยม เช่น 40 45 45 50 50 50 50 55 55 ***90***

4.) ในกรณีที่มีฐานนิยมหลายตัว ทำให้ไม่สามารถสรุปได้ดี

3.4) ค่าเฉลี่ยถ่วงน้ำหนัก(Weighted Mean : $\overline{x}_w$) ในบางกรณีการหาค่าเฉลี่ยที่มีการให้ความสำคัญหรือน้ำหนักของข้อมูลแต่ละตัวไม่เท่ากัน เช่น การคำนวณเกรดเฉลี่ย ซึ่งน้ำหนักของแต่ละวิชา(จำนวนหน่วยกิต)ไม่เท่ากัน

สูตร Weighted Mean $\overline{x}_w = \dfrac{\sum w_i x_i}{\sum w_i}$

เมื่อ $\overline{x}_w$ = ค่าเฉลี่ยน้ำหนัก

X_i = ค่าสังเกตของหน่วยสังเกตตัวที่ i

W_i = ค่าน้ำหนักหน่วยสังเกตตัวที่ i

i = หน่วยสังเกตแต่ละตัว 1, 2, 3 ,....,n

4) การวัดการกระจาย (Measures of Dispersion) เป็นการบรรยายค่าความแตกต่างของข้อมูล ว่าข้อมูลมีการกระจายมากน้อยเพียงใด

การใช้ค่าการวัดแนวโน้มเข้าสู่ส่วนกลางเพื่อบรรยายลักษณะของข้อมูลอย่างเดียวยังไม่เพียงพอ เพราะในบางครั้งข้อมูลที่มีค่ากลางเท่ากัน (ค่าเฉลี่ยเท่ากัน) แต่ค่าของข้อมูลแต่ละตัวภายในกลุ่มอาจแตกต่างกัน เช่น ข้อมูล 3 ชุด ต่อไปนี้

ชุดของข้อมูล	ข้อมูล	ค่าเฉลี่ย	มัธยฐาน	ฐานนิยม
ข้อมูลชุดที่ 1	45 50 55 55 55 60 65	55	55	55
ข้อมูลชุดที่ 2	50 55 55 55 55 55 60	55	55	55
ข้อมูลชุดที่ 3	55 55 55 55 55 55 55	55	55	55

จากข้อมูลดังกล่าว พบว่า ค่าเฉลี่ย ค่ามัธยฐาน และค่าฐานนิยม ของข้อมูล 3 ชุด เท่ากัน คือ 55 แต่ถ้าดูรายละเอียดของข้อมูลแต่ละตัวในทั้ง 3 ชุด พบว่า ข้อมูลชุดที่ 1 มีความแตกต่าง

R2R สำนักงานสาธารณสุขจังหวัดขอนแก่น ปี 2560

กันมากกว่าข้อมูลชุดที่ 2 ซึ่งข้อมูลใกล้เคียงกัน และ ชุดที่ 3 ซึ่งข้อมูลเหมือนกัน ดังนั้นการบรรยายให้เห็นลักษณะของข้อมูลให้ชัดเจนยิ่งขึ้น นอกจากจะนำเสนอและบรรยายค่าการวัดแนวโน้มเข้าสู่ส่วนกลางแล้ว ยังต้องนำเสนอและบรรยายถึงลักษณะการกระจายของข้อมูลในกลุ่มด้วย ซึ่งการกระจายมีหลายวิธี ในที่นี้จะกล่าวถึงวิธีที่นิยมใช้ ดังนี้

4.1) ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation; S.D. หรือ S.) เป็นค่าเฉลี่ยของความแตกต่างของข้อมูลแต่ละตัวกับค่าเฉลี่ย(Mean) หรือ หมายถึง โดยเฉลี่ยๆ แล้ว ข้อมูลแต่ละตัวมีความแตกต่าง(ห่าง)จากค่าเฉลี่ย มากน้อยเพียงใด ถ้าห่างมากแสดงว่าข้อมูลมีการกระจายมาก ถ้าห่างน้อยแสดงว่าข้อมูลมีการกระจายน้อย ส่วนเบี่ยงเบนมาตรฐานของประชากร แทนด้วย สัญลักษณ์ คือ σ (ซิกม่า) ส่วนสัญลักษณ์ส่วนเบี่ยงเบนมาตรฐานของกลุ่มตัวอย่าง S.D. หรือ s.

สูตร ส่วนเบี่ยงเบนมาตรฐานของประชากร $\sigma = \sqrt{\frac{\sum(x_i - \mu)^2}{N}}$

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

เมื่อ μ = ค่าเฉลี่ยของประชากร $\bar{x}$ = ค่าเฉลี่ยของกลุ่มตัวอย่าง

X_i = ค่าสังเกตของหน่วยสังเกตแต่ละหน่วยสังเกต(ข้อมูลแต่ละคน)

N = จำนวนประชากรทั้งหมด n = จำนวนตัวอย่างทั้งหมด

ตัวอย่างที่ เก็บข้อมูลตัวแปรน้ำหนักของนักศึกษา 10 คน ได้ข้อมูลดังนี้ 40 45 45 50 50 50 50 55 55 60 จงคำนวณหาส่วนเบี่ยงเบนมาตรฐานของน้ำหนักนักศึกษา 10 คน

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

แทนค่า จากตัวอย่าง คำนวณค่าเฉลี่ย = 50 กิโลกรัม

คนที่	1	2	3	4	5	6	7	8	9	10	ผลรวม(∑)
x_i	40	45	45	50	50	50	50	55	55	60	500
$(x_i - \bar{x})$	-10	-5	-5	0	0	0	0	5	5	10	0
$(x_i - \bar{x})^2$	100	25	25	0	0	0	0	25	25	100	300

แทนค่า $S.D. = \sqrt{\frac{\sum(x_i - \bar{x})^2}{(n-1)}}$ $S.D. = \sqrt{\frac{300}{9}} = \sqrt{33.33} = 5.77$

4.2) ความแปรปรวน (Variance; $S.D.^2$) เป็นค่าเฉลี่ยของความแตกต่างกำลังสองของข้อมูลแต่ละตัวกับค่าเฉลี่ย (Mean Square Deviation: ส่วนเบี่ยงเบนกำลังสองเฉลี่ย) คือ ส่วนเบี่ยงเบนมาตรฐานยกกำลังสองนั่นเอง หมายถึง โดยเฉลี่ย ๆ แล้ว ข้อมูลแต่ละตัวมีความแตกต่างกำลังสอง (ห่าง) จากค่าเฉลี่ย (ค่ากลาง) มากน้อยเพียงใด ถ้าห่างมากแสดงว่าข้อมูลมีการกระจายมาก ถ้าห่างน้อยแสดงว่าข้อมูลมีการกระจายน้อย ความแปรปรวนของประชากร แทนด้วย สัญลักษณ์ คือ σ^2 (ซิกม่ายกกำลังสอง) ส่วนสัญลักษณ์ความแปรปรวนของกลุ่มตัวอย่าง แทนด้วย $S.D.^2$ หรือ $S.^2$

R2R สำนักงานสาธารณสุขจังหวัดกาฬสินธุ์ ปี 2560

สูตร ความแปรปรวนของประชากร $\sigma^2 = \frac{\sum(x_i - \mu)^2}{N}$

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

ตัวอย่าง เก็บข้อมูลตัวแปรน้ำหนักของนักศึกษา 10 คน ได้ข้อมูลดังนี้ 40 45 45 50 50 50 50 55 55 60 จงคำนวณหาความแปรปรวนของน้ำหนักนักศึกษา 10 คน

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

แทนค่า จากตัวอย่าง คำนวณค่าเฉลี่ย = 50 กิโลกรัม

คนที่	1	2	3	4	5	6	7	8	9	10	ผลรวม(∑)
x_i	40	45	45	50	50	50	50	55	55	60	500
$(x_i - \bar{x})$	-10	-5	-5	0	0	0	0	5	5	10	0
$(x_i - \bar{x})^2$	100	25	25	0	0	0	0	25	25	100	300

แทนค่า $S.D.^2 = \frac{\sum(x_i - \bar{x})^2}{(n-1)}$ $S.D.^2 = \frac{300}{9} = 33.33$

4.3) สัมประสิทธิ์ของความแปรปรวน (Coefficient of Variation : CV)

เป็นค่าที่ใช้ในการเปรียบเทียบการกระจายของข้อมูล ในข้อมูลที่หน่วยการวัดที่แตกต่างกันหรือมีค่าเฉลี่ยแตกต่างกัน ซึ่งเป็นการเปรียบเทียบว่าข้อมูลแต่ละชุดมีการกระจายเป็นกี่เท่าของค่าเฉลี่ย เป็นขนาดของส่วนเบี่ยงเบนมาตรฐานที่เทียบเป็นเปอร์เซ็นต์ของค่าเฉลี่ย

สูตร $CV. = \frac{S.D.}{\bar{x}}$ หรือ $CV. = \frac{S.D.x100}{\bar{x}}$ %

เมื่อ CV. = สัมประสิทธิ์ของความแปรปรวน

S.D. = ส่วนเบี่ยงเบนมาตรฐาน $\bar{x}$ = ค่าเฉลี่ย

ตัวอย่าง เก็บข้อมูลตัวแปรน้ำหนักของนักศึกษา 10 คน ได้ข้อมูลดังนี้ 40 45 45 50 50 50 50 55 55 60 จงคำนวณหาสัมประสิทธิ์ของความแปรปรวนของน้ำหนักนักศึกษา 10 คน

สูตร $CV. = \frac{S.D.x100}{\bar{x}}$

แทนค่า จากตัวอย่าง ค่าเฉลี่ย= 50 กิโลกรัม ส่วนเบี่ยงเบนมาตรฐาน = 5.77

$CV. = \frac{5.77}{50} = .1154$ หรือ $CV. = \frac{5.77x100}{50} = 11.54$ %

4.4) พิสัย (Rage) คือ ผลต่างระหว่างข้อมูลที่มีค่าสูงสุดและค่าต่ำสุด ถ้าพิสัยมีค่ามากแสดงว่าข้อมูลมีการกระจายมาก ถ้ามีค่าน้อยแสดงว่าข้อมูลมีการกระจายน้อย แต่พิสัยเป็นการวัดการกระจายอย่างหยาบๆ เท่านั้น เพราะคำนวณจากค่าสูงสุดและค่าต่ำสุดเท่านั้น หากข้อมูลมีค่าที่มากหรือน้อยผิดปกติ จะทำให้ค่าพิสัยมีค่ามาก

สูตร พิสัย (Range)= Maximun – Minimun

ตัวอย่าง เก็บข้อมูลตัวแปรน้ำหนักของนักศึกษา 10 คน ได้ข้อมูลดังนี้ 50 40 45 45 50 55 50 50 55 60 จงคำนวณหาค่าพิสัยของน้ำหนักนักศึกษา 10 คน

R2R สำนักงานปลัดกระทรวงสาธารณสุข ปีงบประมาณ 2560

สูตร พิสัย (Range) = Maximun – Minimun

แทนค่า พิสัย (Range) = 60 – 40 = 20

4.5) พิสัยควอไทล์ (Interquartile Rang; IQR) = Q_3 - Q_1

ตารางที่ 1.5 สรุปการเลือกใช้ค่าการวัดแนวโน้มเข้าสู่ส่วนกลางและการวัดการกระจายตามระดับของข้อมูล

Level of Measurement	Measurement of Center (Central tendency)	Measurement of spread (Dispersion)
Nominal	Mode	-
Ordinal	Median	IQR =Q3-Q1 or Range= Maximun-Minimum
Interval and Ratio		
Normal Distribution (Symmetric data)	Mean	S.D. or Variance
Skewed data	Median	IQR =Q3-Q1 or Range= Maximun-Minimum

5) การหาค่าที่แสดงตำแหน่งของข้อมูล (Measure of Position Standing) การเรียงข้อมูลจากน้อยไปมาก แล้วหาค่าของข้อมูลที่จะแบ่งข้อมูลออกเป็นส่วนๆ ดังนี้

5.1) ควอไทล์ (Quartile) แบ่งข้อมูลออกเป็น 4 ส่วนเท่าๆ กัน

5.2) เดไซด์ (Decide) แบ่งข้อมูลออกเป็น 10 ส่วนเท่าๆ กัน

5.3) เปอร์เซนไทล์ (Percentile) แบ่งข้อมูลออกเป็น 100 ส่วนเท่าๆ กัน

6) การแจกแจงข้อมูล (Distribution) เป็นการบรรยายว่าข้อมูลที่ศึกษามีการแจกแจงเป็นลักษณะใด โดยดูจาก

6.1) การแจกแจงที่เป็นโค้งปกติ (Normal Distribution)

6.2) ลักษณะความเบ้ (Skewness)

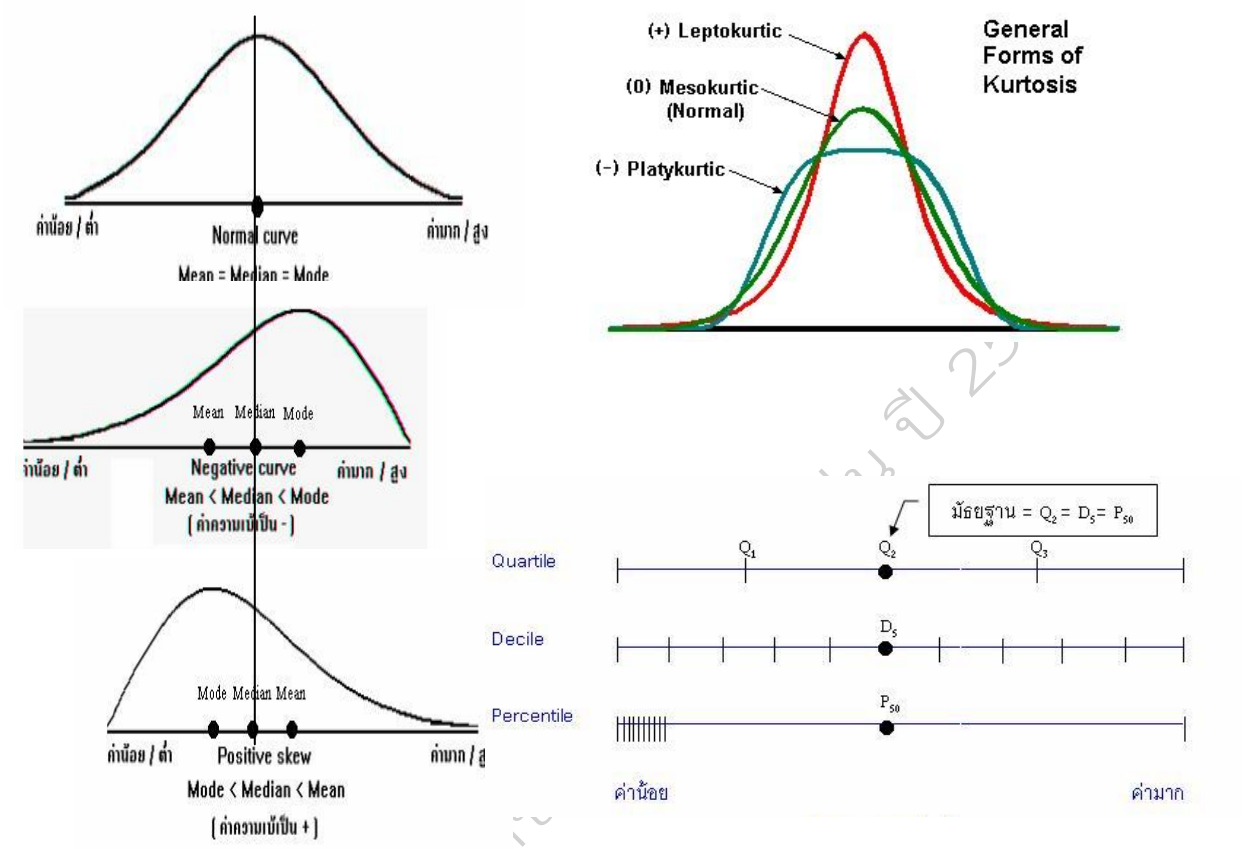
6.3) ลักษณะความโด่ง (Kurtosis)

การแจกแจงแบบปกติมีค่าความเบ้(Skewness) =0 (เบ้ขวา Skewness >0, เบ้ซ้าย Skewness <0

การแจกแจงแบบปกติมีค่าความโด่ง(Kurtosis) =0 (โด่งมาก Kurtosis >0, โด่งน้อย Kurtosis <0

7) การนำเสนอข้อมูล (Presentation of Data) การนำเสนอข้อมูลอาจนำเสนอในลักษณะบรรยายในรูปบทความ ตาราง(Table) กราฟเส้น (Line Graph) แผนภูมิแท่ง (Bar Chart) แผนภูมิวงกลม (Pie Chart) ฮิสโตแกรม (Histogram) แผนภาพกระจาย (Scatter Diagram) Stem and Leaf Plot และ Box Plot ประกอบการบรรยาย เป็นต้น

R2R สำนักงานสาธารณสุขจังหวัดขอนแก่น ปี 2560



2.) สถิติเชิงอนุมานหรือสถิติอ้างอิง (Inferential Statistics)

เป็นการใช้ระเบียบวิธีการทางสถิติในการอนุมานค่าพารามิเตอร์จากค่าสถิติ (Statistic) ที่คำนวณได้จากข้อมูลของกลุ่มตัวอย่างที่เราสุ่มมาศึกษา โดยอาศัยทฤษฎีความน่าประกอบด้วย 2 องค์ประกอบ คือ การประมาณค่าพารามิเตอร์ (Parameter Estimation) นำเสนอในลักษณะของช่วงเชื่อมั่นและการทดสอบสมมติฐาน (Testing Hypothesis) นำเสนอค่า p value เพื่อสรุปผลการทดสอบสมมติฐานว่ามีนัยสำคัญทางสถิติหรือไม่มีนัยสำคัญทางสถิติ (Significance หรือ Non-Significance) การวิเคราะห์ข้อมูลโดยใช้สถิติเชิงอนุมาน ประกอบด้วย

2.1) การประมาณค่าพารามิเตอร์(Parametric Estimation)

2.1.1) การประมาณค่าแบบจุด(Point Estimation)

เป็นการประมาณค่าพารามิเตอร์ (Parameter) เพื่อสรุปหรือบรรยายลักษณะของประชากรด้วยสถิติ (Statistic) ที่ได้จากคำนวณข้อมูลจากตัวอย่างที่ศึกษาด้วยค่าตัวเลขเพียงตัวเดียว ซึ่งในทางปฏิบัติไม่นิยมกระทำ เพราะไม่สามารถบอกโอกาสความผิดพลาดในการสรุปได้ เช่น การประมาณค่า μ จากค่า $\bar{x}$ ซึ่งในทางปฏิบัติจะเห็นได้ว่าค่า $\bar{x}$ ที่คำนวณได้จากกลุ่มตัวอย่าง อาจมีค่าไม่เท่ากันพอดีกับค่า μ หากสามารถศึกษาทั้งประชากร ดังนั้นจึงทำการหาขอบเขต (Interval) ของค่า μ ผลทำให้เกิดการประมาณค่าพารามิเตอร์แบบช่วง (Interval Estimation)

2.1.2) 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%

2.2) การทดสอบสมมติฐาน(Testing Hypothesis)

การทดสอบสมมติฐานเป็นวิธีการของสถิติเชิงอนุมานที่เกี่ยวข้องกับการตัดสินใจสรุปลักษณะของประชากรที่สนใจ โดยใช้ข้อมูลจากตัวอย่าง มีขั้นตอนดังนี้

1.) การตั้งสมมติฐานทางสถิติ

2.) การกำหนดระดับนัยสำคัญทางสถิติ (α)

3.) การเลือกสถิติในการทดสอบ

4.) การหาค่าวิกฤติ/หาค่า p value

5.) ตัดสินใจและสรุปผลการทดสอบ

3.) ขั้นตอนการทดสอบสมมติฐาน

3.1) การตั้งสมมติฐานในการทดสอบ

สมมติฐาน (Hypothesis) คือ ข้อสมมติเกี่ยวกับค่าพารามิเตอร์ที่คาดว่าจะเกิดขึ้น ขั้นตอนการทดสอบสมมติฐานทางสถิติจะเริ่มต้นโดยการนำสมมติฐานของการวิจัย (Research Hypothesis) มากำหนดเป็นสมมติฐานทางสถิติ (Statistical Hypothesis)

สมมติฐานแบ่งเป็น 2 ประเภท คือ

1.) สมมติฐานของการวิจัย (Research Hypothesis)

สมมติฐานของการวิจัย คือ คำตอบที่ผู้วิจัยสงสัยและคาดคะเนไว้ล่วงหน้าอย่างมีเหตุผล โดยเขียนเป็นประโยคข้อความที่แสดงถึงความสัมพันธ์ของตัวแปร เป็นข้อความที่สามารถทดสอบได้

ตัวอย่าง เช่น ผู้วิจัยมีความสนใจเปรียบเทียบค่าเฉลี่ยดัชนีมวลกายของคนที่ออกกำลังกายอย่างสม่ำเสมอกับคนที่ไม่ออกกำลังกาย ว่าความแตกต่างกันหรือไม่

อาจตั้งสมมติฐานของการวิจัยว่า “ค่าเฉลี่ยดัชนีมวลกายในกลุ่มที่ออกกำลังกายอย่างสม่ำเสมอมีค่า**น้อยกว่า**กลุ่มที่ไม่ออกกำลังกาย” เป็นการตั้งสมมติฐานแบบ**มีทิศทาง**

หรือ “ค่าเฉลี่ยดัชนีมวลกายในกลุ่มที่ออกกำลังกายอย่างสม่ำเสมอมีค่า**แตกต่าง**จากกลุ่มที่ไม่ออกกำลังกาย” เป็นการตั้งสมมติฐานแบบ**ไม่มีทิศทาง**

R2R สำนักงานสาธารณสุขจังหวัดอ่างทอง ปี 2560

2.) สมมติฐานทางสถิติ (Statistical Hypothesis)

เป็นสมมติฐานที่กล่าวถึงลักษณะที่สนใจในประชากร อาจเขียนในรูปแบบของข้อความหรือในรูปแบบของสัญลักษณ์ของค่าพารามิเตอร์

สมมติฐานทางสถิติแบ่งเป็น 2 ประเภท คือ

2.1) สมมติฐานของการทดสอบ (Null Hypothesis : H_0)

สมมติฐานของการทดสอบหรือสมมติฐานกลางหรือสมมติฐานว่าง หรือสมมติฐานหลัก เป็นสมมติฐานที่ตั้งว่า **"ไม่มีความแตกต่าง (No Difference)"** หรือ **"ไม่มีความสัมพันธ์ (No Association)"** ของลักษณะที่สนใจในประชากร เป็นสมมติฐานของความไม่แตกต่าง หรือ สมมติฐานของความเท่ากับ ตั้งให้เป็นกลางไว้ก่อน

จากตัวอย่าง สมมติฐานของการวิจัยตั้งไว้ว่า "ค่าเฉลี่ยดัชนีมวลกายในกลุ่มที่ออกกำลังกายอย่างสม่ำเสมอมีค่าน้อยกว่ากลุ่มที่ไม่ออกกำลังกาย"

สมมติฐานการทดสอบกำหนดได้ดังนี้

H_0 : ค่าเฉลี่ยดัชนีมวลกายในกลุ่มที่ออกกำลังกายอย่างสม่ำเสมอมีค่า**<u>เท่ากับ</u>**ค่าเฉลี่ยดัชนีมวลกายในกลุ่มที่ไม่ออกกำลังกาย (เขียนในรูปของข้อความ)

$H_0 : \mu_1 = \mu_2$ (เขียนในรูปของสัญลักษณ์)

เมื่อ μ_1 คือ ค่าเฉลี่ยดัชนีมวลกายในกลุ่มที่ออกกำลังกายอย่างสม่ำเสมอ

μ_2 คือ ค่าเฉลี่ยดัชนีมวลกายในกลุ่มที่ไม่ออกกำลังกาย

2.2) สมมติฐานทางเลือก (Alternative Hypothesis : H_A)

สมมติฐานทางเลือกหรือสมมติฐานรอง เป็นสมมติฐานที่กำหนดขึ้นเพื่อใช้สรุปเมื่อผลการทดสอบพบว่าปฏิเสธสมมติฐาน H_0 ซึ่งจะตั้งให้สอดคล้องกับสมมติฐานของการวิจัย ซึ่งตั้งได้ 2 แบบ คือ

2.2.1) สมมติฐานทางเลือกแบบทางเดียว (One Tail Alternative Hypothesis) เป็นสมมติฐานที่ตั้งในลักษณะความมากกว่าหรือน้อยกว่าในทิศทางใดทิศทางหนึ่งเท่านั้น

ตัวอย่าง สมมติฐานของการวิจัยตั้งไว้ว่า "ค่าเฉลี่ยดัชนีมวลกายในกลุ่มที่ออกกำลังกายอย่างสม่ำเสมอมีค่า**<u>น้อยกว่า</u>**กลุ่มที่ไม่ออกกำลังกาย"

H_A : "ค่าเฉลี่ยดัชนีมวลกายในกลุ่มที่ออกกำลังกายอย่างสม่ำเสมอมีค่า***น้อยกว่า***กลุ่มที่ไม่ออกกำลังกาย" (เขียนในรูปของข้อความ)

$H_A : \mu_1 < \mu_2$ (เขียนในรูปของสัญลักษณ์)

เมื่อ μ_1 คือ ค่าเฉลี่ยดัชนีมวลกายในกลุ่มที่ออกกำลังกายอย่างสม่ำเสมอ

μ_2 คือ ค่าเฉลี่ยดัชนีมวลกายในกลุ่มที่ไม่ออกกำลังกาย

2.2.2) สมมติฐานทางเลือกแบบสองทาง (Two - Tail Alternative Hypothesis) เป็นสมมติฐานที่ตั้งในลักษณะความแตกต่างกัน โดยไม่คำนึงว่าจะต้องอยู่ในทิศทางใด

จากตัวอย่าง สมมติฐานของการวิจัยตั้งไว้ว่า "ค่าเฉลี่ยดัชนีมวลกายในกลุ่มที่ออกกำลังกายอย่างสม่ำเสมอมีค่า**<u>แตกต่าง</u>**จากกลุ่มที่ไม่ออกกำลังกาย"

R2R สำนักงานสาธารณสุขจังหวัดยโสธร ปี 2560

สมมติฐานทางเลือกกำหนดได้ดังนี้

H_A : “ค่าเฉลี่ยดัชนีมวลกายในกลุ่มที่ออกกำลังกายอย่างสม่ำเสมอมีค่า**<u>แตกต่าง</u>**จากกลุ่มที่ไม่ออกกำลังกาย” (เขียนในรูปของข้อความ)

$H_A : \mu_1 \neq \mu_2$ (เขียนในรูปของสัญลักษณ์)

เมื่อ μ_1 คือ ค่าเฉลี่ยดัชนีมวลกายในกลุ่มที่ออกกำลังกายอย่างสม่ำเสมอ

μ_2 คือ ค่าเฉลี่ยดัชนีมวลกายในกลุ่มที่ไม่ออกกำลังกาย

การกำหนดสมมติฐาน H_A ว่าจะเป็นแบบใดต้องพิจารณาจากสมมติฐานของการวิจัย หรือคำตอบที่ผู้วิจัยคาดหวังไว้

2.) การกำหนดระดับนัยสำคัญทางสถิติ (ความผิดพลาดของการทดสอบ Alpha :α) ซึ่งเป็นเกณฑ์เพื่อใช้ประกอบการตัดสินใจ

เกณฑ์ที่ใช้ตัดสินใจว่าจะยอมรับหรือปฏิเสธสมมติฐาน H_0 คือ ระดับนัยสำคัญทางสถิติ (Significance Level) นิยมเขียนอยู่ในรูปสัญลักษณ์ α (Alpha) คือ ความน่าจะเป็นสูงสุดของการปฏิเสธสมมติฐาน H_0 เมื่อ H_0 เป็นจริงซึ่งผู้วิจัยยอมให้เกิดขึ้นได้ การกำหนดค่า α ไม่มีเกณฑ์มาตรฐานใด ๆ ระบุว่าจะต้องมีค่าเท่าใด ผู้วิจัยเป็นผู้กำหนดเองว่าต้องการให้มีโอกาสหรือความน่าจะเป็นในการตัดสินใจหรือสรุปผิดพลาดมากน้อยแค่ไหน โดยทั่ว ๆ ไปนิยมกำหนดไว้ให้เท่ากับ 0.05 หรือ 0.01 ถ้ากำหนดให้ α = 0.05 จะมีความหมายว่า โอกาสที่จะปฏิเสธ H_0 เมื่อ H_0 เป็นจริงมีได้ไม่เกิน 5 ครั้ง ใน 100 ครั้ง

ความผิดพลาดในการตัดสินใจในการทดสอบสมมติฐาน

ผลลัพธ์ที่ได้จากการตัดสินใจ ไม่ว่าจะพบว่ามีนัยสำคัญทางสถิติ (ปฏิเสธสมมติฐาน H_0) หรือไม่พบนัยสำคัญทางสถิติ (ยอมรับสมมติฐาน H_0) ย่อมมีโอกาสเกิดความผิดพลาดขึ้นได้เสมอเนื่องจาก สมมติฐาน H_0 ที่กำหนดขึ้นอาจเป็นจริงหรือไม่จริง

ความสัมพันธ์ของผลลัพธ์ที่ได้จากการทดสอบสมมติฐานและสถานการณ์จริงของสมมติฐาน H_0 ที่กำหนดขึ้นสามารถแสดงดังตาราง ซึ่งในการตัดสินใจไม่ว่าจะยอมรับที่ปฏิเสธสมมติฐาน H_0 ก็จะเกิดความผิดพลาดขึ้น ซึ่งความผิดพลาดจะมี 2 ชนิด คือ Type I Error (α Error) และ Type II Error (β Error)

ตารางที่ 1.6 ความผิดพลาดในการตัดสินใจ

ผลการตัดสิน	**สถานการณ์จริง**	
	H_0 ถูกต้อง	H_0 ผิด (H_A ถูกต้อง)
ปฏิเสธ H_0	สรุปผิด ✘ -- >เกิด α Error	สรุปถูกต้อง ✓ = (1-β) อำนาจของการทดสอบ(Power of test)
ยอมรับ H_0	สรุปถูกต้อง ✓ = (1-α) ความเชื่อมั่น (Confidence Interval)	สรุปผิด ✘-- >เกิด β Error

R2R สำนักงานสาธารณสุขจังหวัดนครราชสีมา ปี 2560

ความผิดพลาดในการตัดสินใจ ประกอบด้วย 2 แบบ

1.) Type I error : α-error : ความผิดพลาดในการปฏิเสธ H_0 ที่เป็นจริง โดยตัดสินใจสรุปว่า แตกต่าง ในสถานการณ์จริงไม่แตกต่าง , ส่วนกลับ 1-α = ความเชื่อมั่น→ใช้ในการประมาณค่าช่วงเชื่อมั่น

2.) Type II error : β-error : ความผิดพลาดในการยอมรับ H_0 ที่เป็นเท็จ โดยตัดสินใจสรุปว่า ไม่แตกต่าง ในสถานการณ์จริงแตกต่าง, ส่วนกลับ1-β = อำนาจการทดสอบ(Power of test) →ใช้ในการคำนวณขนาดตัวอย่าง

3.) การเลือกสถิติในการทดสอบ

การเลือกสถิติทดสอบเป็นอีกขั้นตอนหนึ่งที่สำคัญของการทดสอบสมมติฐาน ซึ่งการที่จะเลือกสถิติใดเป็นสถิติทดสอบให้เหมาะสมกับข้อมูลนั้นต้องพิจารณาหลายอย่าง เช่น วัตถุประสงค์ของงานวิจัย สมมติฐานของการวิจัย ระดับการวัดของข้อมูลหรือตัวแปรที่ศึกษา จำนวนกลุ่มที่ศึกษา

การเลือกสถิติเพื่อทดสอบสมมติฐานที่อาศัยข้อกำหนดหรือข้อตกลงเบื้องต้น (Assumption)เกี่ยวกับการแจกแจง เรียกว่า Parametric Statistics แต่ถ้าข้อมูลที่ศึกษามีการแจกแจงไม่เป็นไปตามข้อตกลงเบื้องต้น การใช้ Parametric Statistics จะมีความผิดพลาด จึงต้องใช้สถิติที่ไม่มีข้อกำหนดหรือข้อตกลงเบื้องต้นดังกล่าว ซึ่งเรียกว่า Non-Parametric Statistics

ซึ่งตัวอย่างของการเลือกใช้สถิติในการทดสอบดังตารางที่ 1.7

ตารางที่ 1.7 ตัวอย่างการเลือกใช้สถิติเพื่อทดสอบสมมติฐาน

ข้อกำหนด	Parametric Statistics	Non-Parametric Statistics
1.การแจกแจงของข้อมูล	แบบปกติ (Normal Distribution)	ไม่มีข้อกำหนด (Free Distribution)
2.ระดับการวัดตัวแปรตาม	Interval Scale หรือ Ratio Scale	ใช้ได้ทุกระดับ Nominal , Ordinal , Interval หรือ Ratio Scale
3.ตัวอย่าง	เชิงสุ่ม(ได้จากการสุ่ม)	เชิงสุ่ม(ได้จากการสุ่ม)
4.ขนาดตัวอย่าง	ใหญ่	เล็ก
5.1 ข้อได้เปรียบ	-ประสิทธิภาพในการทดสอบสูง -เป็นที่นิยม และเป็นรู้จัก	-ใช้ได้ทั้งตัวแปรต่อเนื่องและไม่ต่อเนื่อง -สามารถใช้ได้ในกรณีตัวอย่างขนาดเล็ก -ไม่อิงกับการแจกแจงของประชากร
5.2 ข้อเสียเปรียบ	-ข้อมูลต้องเป็นไปตาม Assumption เช่น ข้อมูลต้องแจกแจงแบบปกติ (Normal Distribution) - ข้อมูลต้องเป็น Interval Scale หรือ Ratio Scale -ขนาดตัวอย่างต้องใหญ่พอ	-อาจจะทำให้เกิดการสูญเสียสารสนเทศที่จำเป็นและสำคัญ -ประสิทธิภาพต่ำกว่าหากข้อมูลเป็นไปตาม Assumption ของ Parametric Statistics -หากคำนวณด้วยมือ กรณีขนาดตัวอย่างใหญ่ต้องใช้เวลามาก -ตารางใช้ยาก

R2R สำนักงานสาธารณสุขจังหวัดนครปฐม ปี 2560

ตารางที่ 1.7 ตัวอย่างการเลือกใช้สถิติเพื่อทดสอบสมมติฐาน (ต่อ)

วัตถุประสงค์	Parametric Statistics	*Non-Parametric Statistics*
การเปรียบเทียบค่ากลาง 1 กลุ่ม	t test แบบ กลุ่มเดียว	-
การเปรียบเทียบค่ากลาง 2 กลุ่ม ที่เป็นอิสระต่อกัน	t test แบบ สองกลุ่มที่เป็นอิสระต่อกัน(Independent t test)	*Mann-Whitney U Test หรือ Wilcoxon Rank Sum Test*
การเปรียบเทียบค่ากลาง 2 กลุ่ม ที่ไม่เป็นอิสระต่อกัน	Paired t test แบบ สองกลุ่มที่ไม่เป็นอิสระต่อกัน	*Wilcoxon (Matched Pairs) Signed Rank Test*
การเปรียบเทียบค่ากลาง มากกว่า 2 กลุ่ม ที่เป็นอิสระต่อกัน	One Way ANOVA	*Kruskal Wallis Test*
การเปรียบเทียบค่ากลาง มากกว่า 2 กลุ่ม ที่ไม่เป็นอิสระต่อกัน	Repeated Measurement	*Friedman Test*
การวิเคราะห์ความสัมพันธ์ของตัวแปร 2 ตัว(เชิงปริมาณ)	Pearson's Product Moment Correlation Coefficient	*Spearman's Rank Correlation Coefficient*
การวิเคราะห์ความสัมพันธ์ของตัวแปร 2 ตัว(แจงนับ)		*Chi Square test หรือ Fisher's Exact Test , Crude OR,RR 95% CI*
การเปรียบเทียบความแตกต่างของค่าสัดส่วน ที่เป็นอิสระต่อกัน	Z Test	*Chi Square testหรือFisher's Exact Test, Crude OR,RR 95% CI*
การเปรียบเทียบความแตกต่างของค่าสัดส่วน ที่ไม่เป็นอิสระต่อกัน(ตัวแปรเชิงคุณภาพ วัดก่อน-หลัง)		*McNemar*
หาสมการในการทำนาย(ปัจจัยที่มีผลต่อ.....)		
1. ตัวแปรตามเป็นข้อมูลเชิงปริมาณ (ตัวแปรต่อเนื่อง) ตัวอย่าง ปัจจัยที่มีความสัมพันธ์กับ(มีผลต่อ/ทำนาย)ระดับโคเลสเตอรอลของผู้ป่วยความดันโลหิตสูง	Multiple Linear Regression	
2. ตัวแปรตามเป็นข้อมูลเชิงคุณภาพ (ตัวแปรแจงนับ) ตัวอย่าง ปัจจัยที่มีความสัมพันธ์กับ(มีผลต่อ/ทำนาย)การมีภาวะแทรกซ้อนของผู้ป่วยเบาหวาน		*Multiple Logistic Regression Adjusted OR , RR 95% CI*

R2R สำนักการแพทย์ กรุงเทพมหานคร สถาบันการแพทย์ฉุกเฉินแห่งชาติ ปี 2560

4.) การคำนวณค่าสถิติทดสอบ หรือการใช้คำสั่งในการวิเคราะห์สถิติทดสอบโดยใช้โปรแกรมสำเร็จรูป

การคำนวณค่าสถิติทดสอบหรือการวิเคราะห์สถิติทดสอบโดยใช้คำสั่งในโปรแกรมสำเร็จรูป เป็นการหาค่าวิกฤต (Critical Value) หรือ p value บนการแจกแจงทางทฤษฎีของค่าสถิติจากข้อมูลตัวอย่างที่ศึกษาภายใต้ข้อสมมติว่า H_0 ที่กำหนดเป็นจริง เพื่อนำค่านี้ไปเปรียบเทียบกับระดับนัยสำคัญที่ตั้งเป็นเกณฑ์ไว้จะยอมรับหรือปฏิเสธสมมติฐานทดสอบ

การคำนวณค่าสถิติทดสอบอาจใช้การคำนวณด้วยมือจากสูตรของสถิติทดสอบ หรืออาจใช้การคำนวณโดยใช้โปรแกรมสำเร็จรูป ซึ่งในปัจจุบันนิยมใช้โปรแกรมสำเร็จรูปในการวิเคราะห์เพราะมีความสะดวก รวดเร็วและถูกต้อง หากคำนวณโดยใช้โปรแกรมสำเร็จรูป โปรแกรมจะคำนวณค่า p value (Sig.) มาพร้อม แต่ถ้าใช้วิธีการคำนวณด้วยมือจากสูตร ผู้วิจัยต้องนำค่าสถิติที่ได้จากการคำนวณไปเปิดตารางสถิติเพื่อหาค่าวิกฤติหรือค่า p value ไม่สะดวก ยุ่งยาก อาจเกิดความผิดพลาดได้

5.) ตัดสินใจและสรุปผลการทดสอบ

เป็นขั้นตอนของการเปรียบเปรียบเทียบค่า p value ที่ได้จากการคำนวณจากข้อมูลที่เราศึกษาแล้วนำมาเปรียบเทียบกับระดับนัยสำคัญทางสถิติ (α) ที่เราตั้งไว้เป็นเกณฑ์เพื่อจะสรุปว่าจะยอมรับหรือปฏิเสธสมมติฐานในการทดสอบ (H_0) โดยคำนึงถึงสมมติฐาน H_A ด้วยว่าเป็นแบบทางเดียวหรือสองทาง ซึ่งเกณฑ์การตัดสินใจ มีหลักเกณฑ์ดังนี้

กรณี สมมติฐานเลือก (H_A) เป็นแบบทางเดียว (One Tail Hypothesis)

จะปฏิเสธสมมติฐานในการทดสอบ (H_0) เมื่อค่า p - value $<\alpha$ สรุปว่า มีความแตกต่าง (มีความสัมพันธ์) กันอย่างมีนัยสำคัญทางสถิติ(Significance)

จะยอมรับสมมติฐานในการทดสอบ (H_0) เมื่อค่า p - value $\geq\alpha$ สรุปว่า มีความแตกต่าง (มีความสัมพันธ์) กันอย่าง**ไม่**มีมีนัยสำคัญทางสถิติ (Non-Significance)

กรณี สมมติฐานเลือก (H_A) เป็นแบบสองทาง (Two Tail Hypothesis)

จะปฏิเสธสมมติฐานในการทดสอบ (H_0) เมื่อค่า p - value $<\alpha/2$ สรุปว่า มีความแตกต่าง (มีความสัมพันธ์) กันอย่างมีนัยสำคัญทางสถิติ (Significance)

จะยอมรับสมมติฐานในการทดสอบ (H_0) เมื่อค่า p - value $\geq\alpha/2$ สรุปว่า มีความแตกต่าง (มีความสัมพันธ์) กันอย่าง**ไม่**มีมีนัยสำคัญทางสถิติ (Non-Significance)

R2R สำนักงานสาธารณสุขจังหวัดหนองบัวลำภู ปี 2560

7. สรุป

การวิจัย (Research) เป็นกระบวนการแสวงหา “องค์ความรู้(knowledge)” การได้มาซึ่งองค์ความรู้โดยการใช้กระบวนการทางการวิจัยนั้นมีหลายขั้นตอนที่เกี่ยวข้องกับสถิติ ดังนั้นสถิติจึงเป็นเครื่องมือที่สำคัญในกระบวนการแสวงหาความรู้ องค์ความรู้ที่เราต้องการ คือ องค์ความรู้ที่มุ่งอธิบายลักษณะ ปรากฏการณ์ หรือเรื่องศึกษาในประชากร (Population) ซึ่งค่าที่ใช้อธิบายหรือสรุปลักษณะของประชากร จะเรียกว่า ค่าพารามิเตอร์ (Parameter) แต่ในการศึกษาส่วนใหญ่เราไม่สามารถศึกษาสมาชิกทุกหน่วยของประชากรได้ เนื่องจาก ประชากรมีขนาดใหญ่เกินไป ศึกษาได้ยาก สิ้นเปลืองทรัพยากร ทั้งเงิน คน และ เวลา หรือในบางกรณีเราไม่ทราบขนาดหรือขอบเขตที่แท้จริงของประชากร แต่เราสามารถได้องค์ความรู้เกี่ยวกับประชากรโดยใช้สถิติเข้ามาช่วยในการสรุปและตัดสินใจ โดยอาศัยข้อมูลหรือการศึกษาจากสมาชิกบางส่วนที่เราสุ่มมาจากประชากร เรียกว่า ตัวอย่าง (Sample) การอธิบายหรือสรุปลักษณะของกลุ่มตัวอย่างจากข้อมูลกลุ่มตัวอย่างที่สุ่มมา เรียกว่า สถิติเชิงพรรณนา (Descriptive Statistics) ค่าที่ใช้สรุปลักษณะของกลุ่มตัวอย่าง เรียกว่า ค่าสถิติ (Statistic) แต่เป้าหมายหลักของการศึกษาคือการได้องค์ความรู้เพื่อใช้อธิบายหรือสรุปลักษณะของประชากร การอธิบายหรือสรุปลักษณะของประชากรโดยใช้ข้อมูลที่ได้จากการศึกษาจากกลุ่มตัวอย่างนี้ เรียกว่า การอนุมานทางสถิติ (Inferential Statistics) หรือการสรุปค่าพารามิเตอร์จากค่าสถิตินั่นเอง การอนุมานทางสถิติที่อาศัยข้อกำหนดหรือข้อตกลงเบื้องต้น(Assumption)เกี่ยวกับการแจกแจง เรียกว่า Parametric Statistics แต่ถ้าข้อมูลที่ศึกษามีการแจกแจงไม่เป็นไปตามข้อตกลงเบื้องต้น การใช้ Parametric Statistics จะมีความผิดพลาด จึงต้องใช้สถิติที่ไม่มีข้อกำหนดหรือข้อตกลงเบื้องต้นดังกล่าว ซึ่งเรียกว่า Non-Parametric Statistics

R2R สำนักงานสาธารณสุขจังหวัดพิษณุโลก ปี 2560