

## บทนำ ความหมายความเป็นมาของสถิติศาสตร์และความน่าจะเป็น

วิชาสถิติหรือสถิติศาสตร์วางแนวสำคัญในการวางแผนการเก็บตัวเลข ของเรื่องราวความจริงในธรรมชาติหรือความจริงที่มนุษย์สร้างสรรค์ขึ้นมาการวิเคราะห์หรือพรรณาส่งเหล่านั้นทั้งหมดในองค์รวมตามขอบเขตที่กำหนด หรือใช้ตัวเลขบางส่วนที่สามารถจัดหาได้อย่างสมเหตุสมผลภายใต้แนวทางความน่าจะเป็นหรือเหตุผลอื่นๆ แล้วดำเนินการประมาณค่าหรือดำเนินการทดสอบสมมุติฐานไปยังองค์รวม ตลอดจนการสรุปผลและการตัดสินใจอย่างมีเหตุผล ตามความเป็นจริงในธรรมชาติหรือ ความเป็นจริงภายใต้หลักเกณฑ์ที่เหมาะสม (Validity) ที่สำคัญอย่างยิ่ง สถิติถูกใช้สำหรับการวางแผนงาน การกำหนดทางเลือก การประเมินทางเลือก และการตัดสินใจทางเลือกเพื่อแก้ปัญหาการบริหารงานธุรกิจ คุณภาพการตัดสินใจทางธุรกิจและการบริหาร ขึ้นอยู่กับข้อมูลและข้อสนเทศมีประโยชน์สำหรับช่วยในการตัดสินใจที่มีประสิทธิภาพสูงยิ่งขึ้น ข้อมูลและข้อสนเทศ

### สถิติศาสตร์

สถิติ ตรงกับภาษาอังกฤษว่า Statistics พจนานุกรมศัพท์คณิตศาสตร์ฉบับ ราชบัณฑิตยสถานใช้คำว่าสถิติศาสตร์ เชื่อว่าคำ Statistics มาจากคำว่า Status และจากภาษาลาตินสมัยโรมันว่า State หมายถึงรัฐ ในความหมายเดิมสถิติจึงหมายถึง เรื่องการจัดเก็บตัวเลข การบันทึกข่าวสาร ข้อสนเทศเกี่ยวกับคนสัตว์ สิ่งของหรือ ปรัชญาการต่าง ๆ อันเกี่ยวกับประชาชนของรัฐ เพื่อเป็นประโยชน์ต่อการบริหารรัฐ จึงนับว่าต้นกำเนิดของวิชาสถิติมาจากนักปกครองนั่นเอง

ในต้นคริสต์ศตวรรษที่ 17 มีนักคณิตศาสตร์ชาวฝรั่งเศสชื่อ อะบราฮัมเดอ มัวฟ (Abraham de Moivre ค.ศ. 1667-1754) ได้ค้นพบสมการการแจกแจงปกติเป็นกราฟ เส้นโค้งรูประฆัง (Normal curve) แล้วคาร์ลเฟรดริค เกาส์ (Carl Friedrich Gauss ค.ศ. 1777-1855) ชาวเยอรมันได้นำสมการกราฟเส้นโค้งนี้ไปใช้อธิบายความผิดพลาด ของการวัดระยะทางดาราศาสตร์ (Astronomy) จนถือได้ว่าเป็นต้นตำหรับของทฤษฎีว่า ด้วยค่าผิดพลาด (Theory of error) คนทั่วไปจึงรู้จักกราฟเส้นโค้งปกติในนาม Normal Function of error หรือ Gaussian distribution.

คริสต์ศตวรรษที่ 18 นักสถิติชาวเบลเยียมชื่อ อะดอร์ฟ เควเตเลต (Adolph Quetelet ค.ศ. 1796-1874) เป็นคนแรกที่ประยุกต์วิชาสถิติไปใช้ในแขนงวิชาการ ศึกษาและสังคมวิทยา โดยนำมาใช้ในการรวบรวมข้อมูลและการแปลความหมายข้อมูล อย่างได้ผลดีและเป็นที่ยอมรับ จนท่านได้รับการยกย่องเป็น Father of the quantitative method in sociology ต่อมาเซอร์ฟรานซิสกาลตัน (Sir Francis Galton ค.ศ. 1822-1911) ได้นำไปใช้ศึกษาทางชีววิทยาในแขนงพันธุกรรมและที่มาของ เชื้อชาติ คาร์ล เพียร์สัน (Karl Pearson ค.ศ.1857-1936) ได้วางหลักการทดสอบสมมุติฐาน

(Pearson's test:การทดสอบของเพียร์สัน) การวิเคราะห์ความแปรปรวน การวางแผนการทดลอง และอื่น ๆ อันเกี่ยวข้องไปถึงหลักการชักสิ่งตัวอย่าง

ในปี ค.ศ. 1908 นักคณิตศาสตร์ชาวไอริชชื่อ วิลเลียม เอส. กอสเสต (William S. Gosset ค.ศ.1876-1937) ใช้นามปากกาว่า Student นำเสนอการใช้สิ่งตัวอย่างขนาดเล็ก (a small sample) ทำให้การแจกแจงสิ่งตัวอย่างขนาดเล็กเรียกว่า Student's t-distribution หรือ t-distribution จนนำมาใช้ในสถิติอนุมานอย่างกว้างขวาง ผู้ที่ริเริ่มบุกเบิกงานทดลองอย่างจริงจัง คิดค้นวิธีการวิเคราะห์ความแปรปรวน (Analysis of Variance:ANOVA)คือ โรนัลด์ ไอเมอร์ ฟิชเชอร์ (Ronald Aylmer Fisher ค.ศ.1809-1962) ทำให้การแจกแจงของฟิชเชอร์ (Fisher's distribution หรือ F-distribution) เป็นสิ่งสำคัญในการวิเคราะห์ความแปรปรวน โดยเฉพาะความแปรปรวนในงานทดลอง

จากที่กล่าวมายังมีนักคิด นักค้นคว้าอีกหลายท่านที่ไม่ได้กล่าวถึงแสดงให้เห็นวิวัฒนาการที่ยาวนานของวิชาสถิติและจะยังคงก้าวหน้าต่อไปตราบใจที่มนุษย์ยังอยากรู้ อยากรู้ เห็น อยากรู้ พัฒนา สักคมอยู่ เราอาจให้ความหมายของวิชาสถิติศาสตร์พอเป็นสังเขปว่าเป็นศาสตร์ที่กล่าวถึง กระบวนการสำรวจ การวางแผนการทดลองการเก็บตัวเลขและประมวลผลข้อมูล การทดสอบสมมุติฐาน การประมาณค่าเกี่ยวกับประชากร ตลอดจนการแปลความหมายและการทำข้อสรุป (Interpretation and drawing conclusion) หรือวิชาสถิติหมายถึงวิชาที่ว่าด้วย “ระเบียบวิธีทางสถิติ” (Statistical method) อันประกอบด้วย

- (1) การเก็บรวบรวมข้อมูล
- (2) การนำเสนอข้อมูล
- (3) การวิเคราะห์ข้อมูล
- (4) การตีความหมายและการหาข้อสรุป

## สาขาของวิชาสถิติ

จากขอบเขตอันกว้างไกลของวิชาสถิติ เราอาจแบ่งแยกได้เป็น 3 สาขาคือสถิติพรรณนา สถิติอนุมานและสถิติปฏิบัติ

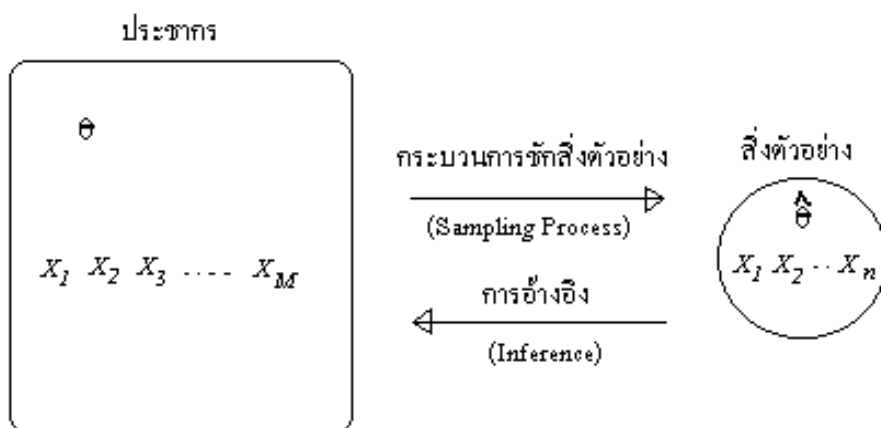
(ก) สถิติพรรณนา (Descriptive statistics) เป็นสถิติที่ว่าด้วยการอธิบาย พรรณนาหรือกล่าวถึงเหตุการณ์ที่เกิดขึ้นจริง เช่น สถิติรายชื่อร้านค้าปลีก พ.ศ.2532 สถิติการตายเนื่องจากอุบัติเหตุทางรถยนต์ในปี พ.ศ.2530 เป็นต้น รวมถึง การวิเคราะห์ข้อมูลในรูปการบรรยายเป็นบทความ การแสดงด้วยตาราง แผนภูมิ กราฟ รูปภาพ ตลอดจนการคำนวณค่าเฉลี่ย ค่าเบี่ยงเบนมาตรฐานหรืออื่น ๆ อันเกี่ยวกับสิ่งที่ต้องการกล่าวถึง

(ข) สถิติอนุมาน (Inferential statistics หรือ Inductive statistics)เป็นสถิติที่เกี่ยวกับการศึกษาคุณลักษณะของประชากร (Parameter) หรือสถานการณ์ที่เป็นจริงในธรรมชาติ โดยอาศัยข้อมูล ข่าวสารหรือข้อสนเทศเพียงบางส่วนที่เรียกว่า “ สิ่งตัวอย่าง หรือ ตัวอย่าง ” (sample)โดยเทคนิคการเลือกสิ่งตัวอย่าง (Sampling Technique)หรือกระบวนการชักสิ่งตัวอย่าง (Sampling

Process)ที่เหมาะสม เช่น ต้องการศึกษผลผลิตถั่วเหลืองพันธุ์ สจ.10 ที่ปลูกในเขตภาคกลาง อาจเก็บข้อมูลจากตำบลต่าง ๆ ของภาคกลางแบบสุ่ม แล้วสรุปผลว่าผลผลิตถั่วเหลืองพันธุ์นี้เป็น 90 กก./ไร่ เป็นต้น

ในสถิติอนุมานเมื่อกล่าวถึงประชากร (Population) ขนาด  $M$  มักจะอ้างคุณลักษณะของประชากร เช่น ค่าเฉลี่ย( $\mu$ ), ค่าเบี่ยงเบนมาตรฐาน ( $\sigma$ ) หรืออื่นๆ คุณลักษณะเหล่านี้เราเรียกว่า “พารามิเตอร์”(parameter)ของประชากรที่เรากำลังสนใจอยู่ สัญลักษณ์แทนพารามิเตอร์ที่ใช้กันเสมอคือ  $\Theta$  หรือ  $\theta$  (theta : ที่ต้า,อักษรกรีก)

สำหรับสิ่งตัวอย่างขนาด  $n$  เป็นส่วนหนึ่งของประชากรและมีเนื้อเดียวกัน ต่างกันที่ขนาดเล็กกว่า คุณลักษณะของสิ่งตัวอย่าง เช่น ค่าเฉลี่ย ( $\bar{X}$ ),ค่าเบี่ยงเบนมาตรฐาน(S)หรืออื่นๆจะถูกกำหนดสัญลักษณ์ให้แตกต่างจากที่ใช้กับประชากร คุณลักษณะเหล่านี้เราเรียกว่า “สถิติ” “ตัวสถิติ” หรือ “ค่าสถิติ ” (Statistic) สัญลักษณ์ทั่วไปที่ใช้กันเสมอคือ  $\hat{\Theta}$  หรือ  $\hat{\theta}$  แล้วนำผลได้ไปอ้างอิงหรือสรุปเกี่ยวกับ พารามิเตอร์



รูปแสดงประชากรและสิ่งตัวอย่าง ในสาขาสถิติอนุมาน

(ค) สถิติปฏิบัติ เป็นการกล่าวถึง ระเบียบวิธีและเทคนิคในทางปฏิบัตินับตั้งแต่การวางแผนการรวบรวมข้อมูล การเก็บข้อมูลโดยการสำรวจ การชักสิ่งตัวอย่าง การเก็บข้อมูลทั้งด้วยแรงคนและด้วยคอมพิวเตอร์

### ความน่าจะเป็น

โอกาสพบเรื่องราวหรือเหตุการณ์ที่อยู่ในความสนใจภายในขอบเขตที่กำหนดไว้หรือความน่าจะเป็น (probability) ของเรื่องราวนั้นด้วยการแสดงค่าเป็นตัวเลขขณะที่เรื่องราวเหล่านั้นยังไม่เกิดด้วยแสดงการเปรียบเทียบจำนวนหนทางที่เกิดเหตุการณ์ที่เรากำลังสนใจอยู่กับจำนวนหนทางที่จะเกิดเหตุการณ์ที่พึงเป็นไปได้ทั้งหมด

ถ้ากำหนด  $N$  : เป็นจำนวนหนทางที่จะเกิดเหตุการณ์ที่พึงเป็นไปได้ทั้งหมด  
 $a$  : เป็นจำนวนหนทางที่เกิดเหตุการณ์ที่เรากำลังสนใจอยู่  
 $A$  : เป็นเหตุการณ์ที่เรากำลังสนใจอยู่  
และ  $P(A)$  : เป็นความน่าจะเป็นเหตุการณ์  $A$  จะได้สมการ (0-1)

$$P(A) = \frac{a}{N} \quad \dots(0-1)$$

จำนวนหนทางที่จะเกิดเหตุการณ์ที่พึงเป็นไปได้ทั้งหมด ( $N$ ) นิยมแสดงไว้ในรูปของเซต (set) เรียกว่า เซตผลลัพธ์ที่พึงเป็นไปได้ทั้งหมด (all possible event set or all possible outcomes set or sample space) เขียนแทนเซตนี้ด้วย  $S$  เช่น ผลลัพธ์ที่พึงเป็นไปได้ทั้งหมดของสินค้าขึ้นที่วางขายเป็น  $S$  คือ {ขายได้, ขายไม่ได้} ผลลัพธ์ที่พึงเป็นไปได้ทั้งหมดของผลการเรียนนักศึกษาผู้หนึ่งเป็น  $S$  คือ {A, B, C, D, E} ผลลัพธ์ที่พึงเป็นไปได้ทั้งหมดของการออกรางวัลเลขท้ายสองท้ายสองตัวสำหรับหนึ่งปี  $S$  คือ {00, 01, 02, 03, ..., 99} เป็นต้น

เราอาจกล่าวได้ว่าความน่าจะเป็นหรือโอกาสคือความถี่สัมพัทธ์ของเหตุการณ์ที่กำลังสนใจอยู่ในจำนวนเหตุการณ์ที่พึงเป็นไปได้ทั้งหมด ทำให้ความน่าจะเป็นมีลักษณะเฉพาะ ดังนี้คือ

- (1) ความน่าจะเป็นของเหตุการณ์ใด ๆ มีค่าอยู่ระหว่าง 0 ถึง 1
- (2) ผลรวมความน่าจะเป็นของทุกเหตุการณ์ในจำนวนเหตุการณ์ที่พึงเป็นไปได้ทั้งหมด เท่ากับ 1
- (3) ความน่าจะเป็นของเหตุการณ์ใด ๆ จะมีความหมายเป็น ความน่าจะเป็น หรือ โอกาสก็ต่อเมื่อ เหตุการณ์นั้นยังไม่เกิด หรือเหตุการณ์นั้นเกิดขึ้นแล้วแต่ไม่ทราบผลแน่ชัด

-----

**คำถามท้ายบทนำ (งานมอบหมายที่ W00)**

- (1) รากศัพท์คำว่า STATISTICS มาจากคำใดในภาษาอังกฤษ และหมายถึงอะไร
- (2) คำว่า STATISTICS และ STATISTIC มีความหมายแตกต่างกันอย่างไร จงอธิบาย
- (3) ท่านเข้าใจว่า “ สถิติเป็นวิทยาศาสตร์และศิลปะ ” อย่างไร อธิบาย
- (4) มีผู้กล่าวว่า “นักสถิติกับหมอดูไม่แตกต่างกัน” ท่านเห็นด้วยหรือไม่ อย่างไร
- (5) จงอธิบายความแตกต่างระหว่าง “ สถิติพรรณนา ” กับ “ สถิติอนุมาน ”
- (6) จงอธิบายถึงกระบวนการชักสิ่งตัวอย่างในสาขาสถิติอนุมาน บอกความแตกต่างและแสดงความสัมพันธ์กันระหว่าง ประชากร (Population) และตัวอย่าง (Sample)
- (7) จงอธิบายวิธีการใช้ความน่าจะเป็นในชีวิตประจำวัน