

สถิติ

1. ความหมายและบทบาทของสถิติ

สถิติ (Statistics) เป็นศาสตร์สาขาหนึ่งที่เกี่ยวข้องกับการจัดจำแนกและแจกแจงสิ่งของ หรือปรากฏการณ์ต่าง ๆ เพื่อค้นหาข้อเท็จจริงและข้อสรุปในสิ่งของหรือปรากฏการณ์ต่าง ๆ เหล่านั้น ในระยะเริ่มแรกนั้นใช้สถิติเกี่ยวกับการนับจำนวนชายฉกรรจ์เพื่อเตรียมไว้ทำศึกสงคราม นับจำนวนที่นาที่ราษฎรจะต้องเสียภาษีให้แก่รัฐ จัดเก็บข้อมูลชนิดต่าง ๆ เพื่อใช้ในการปกครองประเทศ ความหมายของสถิติในสมัยดั้งเดิมจึงเป็นเรื่องการจัดเก็บตัวเลข การบันทึกข้อมูลเกี่ยวกับคน สัตว์และสิ่งของต่าง ๆ เพื่อประโยชน์ในการบริหารประเทศเป็นส่วนใหญ่ ดังนั้นสถิติหรือ statistics จึงมาจากคำว่า state ที่แปลว่ารัฐ ซึ่งสอดคล้องกับคำภาษาละตินว่า status ต่อมาได้มีผู้นำสถิติไปใช้ในการเล่นเกมการพนัน และนำไปใช้ในการศึกษาวิชาต่าง ๆ จนเป็นเหตุให้ศาสตร์ในสาขาวิชาต่าง ๆ เหล่านั้นเจริญก้าวหน้าไปอย่างรวดเร็ว ดังที่ประจักษ์อยู่ทั่วไปในขณะนี้

สถิติเป็นระเบียบวิธีที่ประกอบด้วย การเก็บรวบรวมข้อมูล การนำเสนอข้อมูลเหล่านั้นในรูปของตารางหรือแผนภูมิ การคำนวณค่าเฉลี่ย การหาเปอร์เซ็นต์ การทดสอบสมมติฐาน และยังหมายรวมถึงการออกแบบทดลอง (experimental design) การสุ่มตัวอย่าง ตลอดจนการประมวลผลข้อมูล ปัจจุบันถือว่าสถิติเป็นเทคโนโลยีของวิธีการทางวิทยาศาสตร์อย่างหนึ่ง เพราะเราใช้สถิติเป็นเครื่องมือในการตัดสินใจแก้ปัญหาต่าง ๆ เมื่อมีภาวะการณ์ที่ไม่แน่ใจเกิดขึ้นในทางวิทยาศาสตร์และสังคมศาสตร์

เราใช้สถิติในการอธิบายค่าเฉลี่ยหรือตำแหน่งที่ของชุดข้อมูลที่จะศึกษา ใช้สถิติเพื่อวัดการกระจายจากค่าเฉลี่ยของข้อมูล ใช้วัดความสัมพันธ์ระหว่างข้อมูลสองชุดหรือระหว่างตัวแปรสองตัวซึ่งเราเรียกค่าที่วัดความสัมพันธ์นั้นว่าสัมประสิทธิ์สหสัมพันธ์ (correlation coefficient)

และยังใช้ในการพยากรณ์ การทดสอบสมมติฐาน ตลอดจนใช้ประกอบพิจารณาการตัดสินใจ ปัจจุบันโรงงานอุตสาหกรรมต่าง ๆ ช่างนักสถิติไว้เพื่อควบคุมคุณภาพ (quality control) เกี่ยวกับการผลิต การจำหน่าย ตลอดจนการโฆษณาสินค้าต่าง ๆ ของโรงงาน นักสถิติทางด้านธุรกิจใช้สถิติในการวิเคราะห์อนุกรมเวลา (analysis of time series) เกี่ยวกับการผลิตและจำหน่ายสินค้า จะเห็นว่าสถิติเป็นเครื่องมือที่มีบทบาทสำคัญมากในการพัฒนากิจการด้านต่าง ๆ

2. ความเป็นมาของสถิติ

สถิติเชิงพรรณนามีประวัติอันยาวนานพอ ๆ กับการกำเนิดของโลกมนุษย์ สมัยโบราณเขาใช้สถิติในการจัดหาข้อมูลเกี่ยวกับภาษี การสงคราม ผลผลิตทางเกษตรกรรม และการกีฬา สถิติเชิงอนุมาน ซึ่งส่วนใหญ่ใช้ทฤษฎีของความน่าจะเป็น ได้เจริญก้าวหน้าอย่างรวดเร็วตั้งแต่ศตวรรษที่ 16 เป็นต้นมา โดยเฉพาะอย่างยิ่งความเจริญก้าวหน้าของสถิติเป็นผลอันเกิดจากการวิจัยของบุคคลในสาขาต่าง ๆ ในช่วงระยะเวลา 400 ปีที่ผ่านมา

นอกจากนี้ความเจริญก้าวหน้าของสถิติยังเป็นผลสืบเนื่องมาจากการเล่นเกมการพนันของมนุษย์ คือผู้เล่นแต่ละฝ่ายพยายามที่จะเอาชนะเกมการพนัน นักการพนันทั้งหลายจึงได้พยายามกระตุ้นให้นักคณิตศาสตร์คิดค้นหาวิธีการที่จะเอาชนะเกมการพนันให้ได้มากที่สุด จึงได้นำไปสู่การพัฒนาทฤษฎีของความน่าจะเป็น ในยุคนั้นมีนักคณิตศาสตร์หลายท่านที่ได้นำทฤษฎีทางคณิตศาสตร์มาพัฒนาวิธีการทางสถิติ เช่น ปาสกาล (Pascal) ไลบ์นิทซ์ (Leibniz) แฟร์มาต์ (Fermat) และ เจมส์ เบอรรินนูลี (James Bernoulli)

ในปี ค.ศ. 1733 เดอ มัวท์ (De Moivre) ได้ค้นพบสมการของการแจกแจงแบบปกติ (normal distribution) ซึ่งนำไปใช้กันมากในทางสถิติอนุมาน (inferential statistics) บางครั้งเรามักจะเรียกการแจกแจงแบบปกติซึ่งมีลักษณะคล้ายระฆัง (bell-shaped distribution) นี้ว่าการแจกแจง

แบบเกาส์เซียน (Gaussian distribution) เพื่อเป็นเกียรติแก่ เกาส์ (Gauss, 1777 - 1855) ซึ่งเป็นผู้ค้นพบเป็นคนแรก นอกจากนี้ เกาส์ ยังได้สร้างสถารเพื่อศึกษาความคลาดเคลื่อนอันเกิดจากการวัดสิ่งใดสิ่งหนึ่งหลาย ๆ ครั้ง อนึ่ง ลาปลาซ (Laplace) ยังได้นำเอาสถิติมาใช้ในการตรวจสอบทฤษฎีบทเป็นครั้งแรกซึ่งนับว่าเป็นผลงานที่โด่งดังมากในยุคนั้น

ระหว่างคริสต์ศตวรรษที่ 19 มีนักสถิติชาวบราซิลชื่อ อะดอล์ฟ เทเชเตอ์ (Adolph Quetelet, 1796 - 1874) ได้นำวิธีการทางสถิติมาประยุกต์ใช้ทางด้านการศึกษาและสังคมวิทยา เทเชเตอ์ เป็นนักสถิติคนแรกที่แสดงให้เห็นว่าเทคนิคทางสถิติที่ได้มาจากการวิจัยสาขาหนึ่งสามารถนำไปประยุกต์ใช้ในสาขาอื่น ๆ ได้

เซอร์ ฟรานซิส กัลตัน (Sir Francis Galton, 1822 - 1911) เป็นอีกผู้หนึ่งที่มีบทบาทในการนำสถิติมาประยุกต์ใช้ในทางสังคมศาสตร์ โดยเฉพาะการนำสถิติมาแก้ปัญหาที่เกี่ยวข้องพันธุกรรมและการปรับปรุงชาติพันธุ์ของมนุษย์ นอกจากนี้ กัลตัน ยังได้ค้นพบเปอร์เซนไทล์ และ คาร์ล เพียร์สัน (Karl Pearson : 1857 - 1936) ซึ่งทำงานร่วมกับกัลตันในการพัฒนาทฤษฎีของการถดถอยและสหสัมพันธ์ ได้คิดค้นวิธีทางสถิติอีกมากมาย และเป็นผู้ค้นพบทฤษฎีการสุ่มตัวอย่าง (theory of sampling)

ต้นคริสต์ศตวรรษที่ 20 วิลเลียม เอส กอสเสต (William S. Gosset) ได้พัฒนาวิธีการตัดสินใจโดยอาศัยข้อมูลจำนวนน้อย เขาได้ทำการศึกษาในโรงเรียนแห่งหนึ่งของชาวไอริช แต่ไม่ได้พิมพ์ผลงานออกเผยแพร่เพราะเกรงจะเป็นประโยชน์ต่อคู่แข่ง กอสเสตได้พิมพ์เฉพาะผลของการศึกษาของเขาโดยใช้นามปากกาว่า 'Student' นอกจากนี้เซอร์ โรนัลด์ ฟิชเชอร์ (Sir Ronald Pa. Fisher, 1890 - 1962) ได้พัฒนาทฤษฎีเกี่ยวกับกลุ่มตัวอย่างขนาดเล็ก และแบบแผนการทดลอง ซึ่งเป็นผลงานทางสถิติที่สำคัญมากอีกคือ

คริสต์ศตวรรษที่ 20 มีนักสถิติที่มีชื่อเสียงโด่งดังมากมาย ซึ่งล้วนแล้วแต่มีผลงานในการพัฒนาทฤษฎีสถิติใหม่ ๆ ตลอดจนการนำสถิติไปประยุกต์ใช้ในการแก้ปัญหาต่าง ๆ

มากมาย จนกระทั่งในปี ค.ศ. 1950 มหาวิทยาลัยต่าง ๆ ในสหรัฐอเมริกาได้จัดตั้งภาควิชาและคณะวิชาที่เปิดสอนวิชาทางสถิติขึ้นมาโดยเฉพาะ และได้มีการให้ปริญญาโทบัณฑิตทางสถิติศาสตร์เป็นครั้งแรกในปี ค.ศ. 1955 ระยะต่อมาคอมพิวเตอร์ได้เข้ามามีบทบาทอย่างมากในการพัฒนาทฤษฎีและวิธีการต่าง ๆ ทางสถิติจนเข้าสู่ยุคสถิติสมัยใหม่ ในอนาคตคงจะมีทฤษฎีและหลักการทางสถิติใหม่ ๆ เกิดขึ้นมากมาย ปัจจุบันสถิติเป็นเครื่องมืออันสำคัญที่สุดของการทำวิจัย

3. ชนิดของสถิติ

สถิติแบ่งตามวิธีได้ 2 ชนิด คือ สถิติเชิงพรรณนา (descriptive statistics) และสถิติเชิงอนุมาน (inferential statistics)

สถิติเชิงพรรณนา เกี่ยวข้องกับการนำเสนอข้อมูลด้วยตาราง แผนภูมิ แผนภาพ หรือ คำอธิบายการจัดเก็บข้อมูลต่าง ๆ ข้อมูลเหล่านี้ อาจเป็นจำนวนเชิงปริมาณ (quantitative data) เช่น ความสูง น้ำหนัก เป็นต้น หรือเป็นจำนวนเชิงคุณภาพ (qualitative data) เช่น ที่ศคนคิดต่อระบบการปกครอง การนับถือศาสนา อาชีพผู้ปกครอง เป็นต้น การนำข้อมูลต่าง ๆ เหล่านี้มาจัดจำแนกตามประเภท ลักษณะและตามจุดมุ่งหมายที่เราต้องการจะทราบ ก่อนที่จะตีความหมาย เพื่อให้เข้าใจความหมายในธรรมชาติของข้อมูลเหล่านั้น สถิติเชิงพรรณนาเป็นแต่เพียงวิธีการหาข้อสรุปหรือบรรยายจากข้อมูลเท่านั้น มิได้มีเทคนิคพิเศษอย่างใดที่จะนำมาช่วยในการตีความหมาย

สถิติเชิงอนุมาน เป็นเทคนิคการแก้ปัญหาอีกระดับหนึ่งซึ่งยากกว่า วิธีนี้เป็นวิธีของการหาข้อสรุป (infer) เกี่ยวกับประชากร (population) จากการพิจารณาและศึกษาข้อมูลในกลุ่มตัวอย่าง (sample) ที่สุ่มมาจากข้อมูลในประชากรนั้น ตัวอย่างเช่น พยาบาลของโรงเรียนในท้องถิ่นแห่งหนึ่ง อยากทราบว่านักเรียนชั้นประถมศึกษาปีที่ 4 ที่ไม่สอบปลุกผีเลย มีสัดส่วนเท่าใดของนักเรียนชั้นประถมศึกษาปีที่ 4 ทั้ง

หมด ก็ไม่จำเป็นต้องไปตามนักเรียนแต่ละคนทั้งหมด เพียงแต่สุ่มตัวอย่างนักเรียนชั้นประถมศึกษาปีที่ 4 มา 100 คน แล้วถามว่าใครไม่เคยปลูกผีบ้าง ก็สามารถประมาณสัดส่วนของจำนวนนักเรียนชั้นประถมศึกษาปีที่ 4 ที่ไม่เคยปลูกผี ต่อจำนวนนักเรียนชั้นประถมศึกษาปีที่ 4 ทั้งหมดได้ เป็นต้น จุดมุ่งหมายของสถิติอนุมานก็เพื่อเสาะแสวงหาคุณสมบัติเฉพาะของประชากรจากความรู้ที่ได้ศึกษาจากตัวอย่างที่สุ่มมา อย่างไรก็ตาม สถิติย่อมเกี่ยวข้องกับวิธีการเก็บรวบรวมข้อมูลเพื่อนำมาใช้ในการตัดสินใจอย่างมีเหตุผลและกำจัดอิทธิพลอันเนื่องมาจากความคลาดเคลื่อนต่าง ๆ ให้มีน้อยที่สุด

4. ค่าสถิติ (statistic)

ค่าสถิติเป็นตัวเลขที่คำนวณได้จากตัวอย่าง (sample)

เช่น ค่าเฉลี่ยของตัวอย่าง \bar{x} (sample mean) ความแปรปรวนของข้อมูลในตัวอย่าง s^2 (sample variance) เป็นต้น ค่า \bar{x} และ s^2 เรียกว่าค่าสถิติ (statistic) แต่ถ้าคำนวณจากประชากร (population) เช่น ค่าเฉลี่ยของข้อมูลในประชากร (population mean) ความแปรปรวนของข้อมูลในประชากร σ^2 (population variance) ค่า μ และ σ^2 เรียกว่า พารามิเตอร์ (parameter) โดยทั่วไปเราจะประมาณค่า พารามิเตอร์ของประชากรด้วยค่าสถิติ ที่คำนวณจากข้อมูลในกลุ่มตัวอย่างที่สุ่มมาจากประชากรนั้น บางครั้งเรียกค่าสถิตินี้ว่าเป็นตัวประมาณ (estimator) การประมาณค่าพารามิเตอร์นั้นอาจจะระบุช่วงความเชื่อมั่น (confidence interval) ไว้ด้วย

อนันต์ ศรีโสภาก

บรรณานุกรม

อนันต์ ศรีโสภาก สถิติเบื้องต้น บริษัทสำนักพิมพ์ไทยวัฒนาพานิชจำกัด 891 ถนนพระราม 1 กรุงเทพมหานคร พ.ศ. 2525

Mood, Alexander and Graybill, Franklin A. *Introduction to the Theory of Statistics.* 2d ed. New York : McGraw-Hill, 1963.

Walpole, R.E. *Introduction to Statistics.* New York : Macmillan Company, 1968.