

บทนำ

สถิติคณิตศาสตร์ นับวันที่จะเพิ่มความสำคัญหรือบทบาทในสาขาวิชาต่าง ๆ อย่างเช่น ในด้านผลิตภัณฑ์อุตสาหกรรมจำนวนมาก ๆ การแพทย์และชีววิทยา เศรษฐศาสตร์ รัฐศาสตร์ จิตวิทยา สังคมศาสตร์ เกษตรศาสตร์ ศึกษาการจราจร และวิทยาศาสตร์

สถิติคณิตศาสตร์เกี่ยวกับทฤษฎีและการประยุกต์ของวิธีการสำหรับการรวบรวมข้อมูล ทางสถิติ การวิเคราะห์ข้อมูล การทำการอนุมาน ข้อมูลสถิติอาจประกอบด้วยหมายเลขจัดเรียง ในรูปตารางหรือรูปภาพ ข้อมูลเหล่านี้อาจได้มาจากปัญหาการจราจร การทดลองทางการแพทย์ เพื่อกำหนดผลของยาชนิดใหม่ การควบคุมคุณภาพของผลิตภัณฑ์อุตสาหกรรมในโรงงาน การวัดประชามติและคนที่ชอบ ของผู้ออกเสียงเลือกตั้ง การสำรวจค่าครองชีพ

ด้วยเหตุนี้เราเห็นว่า ข้อมูลทางสถิติอาจรวบรวมได้ในรูปต่าง ๆ กัน แต่กระนั้น ข้อมูลก็ยังมีลักษณะเหมือนกัน ที่ได้มาจากเหตุการณ์โดยบังเอิญ นั่นคือ เราไม่สามารถทำนายผลที่ได้ เพราะว่าตัวประกอบบางอย่างที่ไม่สามารถควบคุมได้ อย่างเช่น ถ้าหากว่าเราทอดลูกเต๋าหนึ่งลูก ทั้งที่เรารอบสรวงหน้าว่า จะต้องปรากฏหน้าใดหน้าหนึ่งของหมายเลข 1, 2, 3, 4, 5, 6, เกิดขึ้น แต่เราก็ไม่สามารถทำนายผลลัพธ์ในการดำเนินของการทดลองเชิงสุ่มนี้ นั่นคือ เราไม่สามารถทำนายว่าหมายเลขอันไหนจะปรากฏเกิดขึ้นในการทอด อย่างไรก็ตาม ถ้าหากว่าลูกเต๋าสมดุแล้ว ในการทดลองระยะยาว เราสังเกตเห็นว่าหมายเลขทั้งหกเหล่านั้น จะปรากฏเกิดขึ้นจำนวนครั้งเท่ากันโดยประมาณ อย่างเช่น ในการทดลองทอดลูกเต๋า 600 ครั้ง แต่ละหมายเลขจะปรากฏเกิดขึ้นประมาณ 100 ครั้ง

ในทำนองเดียวกัน การกำเนิดเด็กก็อาจเป็นการทดลองเชิงสุ่มที่มีสองผลลัพธ์ที่อาจเป็นไปได้ “ชาย” หรือ “หญิง” เราไม่สามารถทำนายในกรณีใดกรณีหนึ่งโดยเฉพาะ แต่ประสบการณ์แสดงว่า ในประชากรใหญ่ ๆ อย่างเช่น ในนครใหญ่ ๆ ทดสอบผลลัพธ์ (ชายหรือหญิง) ปรากฏว่าเท่ากันโดยประมาณ

ถ้าหากว่าเรามีโรงงานผลิตสกรู เราไม่สามารถทำนายว่า สกรูตัวต่อไปที่ผลิต จะดีหรือจะเสีย แต่เราอาจสังเกตทุก ๆ วัน จำนวนเปอร์เซ็นต์ของสกรูเสียเกือบเท่ากัน ถ้าหากว่ากระบวนการผลิตยังคงไม่เปลี่ยน

ในทำนองเดียวกัน เราไม่สามารถทำนายว่า ที่อายุเท่าไรที่บุคคลหนึ่งจะตาย แต่เราก็สามารถหาตารางของการมีชีวิตอยู่ที่คาดหวังได้โดยพิจารณาประชากรที่มีขนาดใหญ่ และเราอาจใช้ตารางเหล่านี้เพื่อคำนวณทุนประกันชีวิต เทียบกับชนิดอื่น ๆ ของการประกันภัย (ไฟ, ขโมย,

รายนต์) สภาวะก็คล้ายคลึงกัน

ในตัวอย่างของเรา มีอยู่กรณีหนึ่งแสดงการสุ่มอย่างไม่มีการเลือกสรร นั่นทำให้การทำนายไม่อาจเป็นไปได้ แต่กรณีมีจำนวนมาก ๆ ก็สามารถแสดงว่ามีกฎเกณฑ์ทางสถิติได้ ความสำคัญที่จะหารายละเอียดตามกฎทางสถิติ เพื่อที่เราอาจปรับตัวแบบทางคณิตศาสตร์ ตัวแบบนั้นช่วยในสภาวะที่เข้าใจเกี่ยวกับผลของโอกาส และอาจใช้เป็นฐานสำหรับการทำนายหรือมีอิทธิพลต่อพฤติกรรมในอนาคต

ความมุ่งหมายของการเลือกตัวอย่างทางสถิติ

ถ้าหากว่า เราประสงค์เพื่อที่จะศึกษาการแพร่เชื้อโรคโดยนกพิราบ ในกรุงเทพฯ เราไม่อาจที่จะจับนกพิราบทั้งหมดในกรุงเทพฯ และตรวจสอบโดยทางแพทย์ เราสามารถจับจำนวน n ตัว แทนที่จะเป็นทั้งหมด (อย่างเช่น $n = 50$, หรือ 100 หรือ 500 ตัว) และตรวจสอบมันเพื่อค้นคว้าว่านกพิราบเหล่านี้เป็นตัวนำเชื้อไวรัสซึ่งเป็นสาเหตุของเชื้อโรค (เป็น 1) หรือไม่ เป็นเชื้อโรค (เป็น 0) เราสามารถบันทึกผลได้ด้วยการเขียน

1 1 0 1 0 0 0 ... 0 1 0,

ในที่นี้หมายความว่า พบเชื้อไวรัสในสองตัวแรก และไม่พบในตัวที่สาม ฯลฯ จำนวนตัวเลขข้างต้นนี้เรียกว่าตัวอย่างสุ่มจากประชากรจำนวนทั้งหมดที่ต้องการหาเพื่อตรวจสอบในกรุงเทพฯ หมายเลขในตัวอย่างเรียกว่า ค่าสังเกตหรือค่าของตัวอย่าง จำนวน n เรียกว่าขนาดของตัวอย่าง ถ้าหากว่า 15% ของค่าสังเกตเป็น 1 เราก็สรุปได้ว่า ประมาณ 15% ของนกพิราบทั้งหมดในกรุงเทพฯ นำเชื้อไวรัสเหล่านั้น ในข้อสรุปนี้เราต้องสมมติว่า วิธีการของเราในการจับนกพิราบนั้น นกพิราบทั้งหมดในกรุงเทพฯ มีโอกาสถูกจับได้เท่า ๆ กัน ด้วยเหตุนี้ นกพิราบ n ตัว ที่ถูกจับจริง ๆ ใช้แทนตัวอย่างสุ่มของนกพิราบทั้งหมดในกรุงเทพฯ

ข้อบ่งชี้ของการสุ่มดังจะได้พิจารณาต่อไปนี้

ในบางกรณีไม่มีความจำเป็นเลยที่จะไปเกี่ยวข้องกับประชากรทั้งหมด เพราะเป็นการทำลาย อย่างเช่นการทดลองอายุของหลอดไฟ การเลือกตั้งในชุมชน ถึงแม้ว่าเราสามารถสำรวจความเห็นของทุก ๆ คนในชุมชน ผลลัพธ์ก็ไม่แน่ว่าจะมีคุณค่ามากกว่าตัวอย่างที่มีขนาดใหญ่พอ เพราะว่าบุคคลเหล่านี้บางคนอาจเปลี่ยนใจของเขาในความตั้งใจครั้งแรกที่ออกเสียงก็ได้ ฯลฯ

มีหลาย ๆ กรณีที่ประชากรมีขนาดใหญ่เกินไปที่จะศึกษาได้สมบูรณ์ เมื่อเทียบกับเวลา

และต้นทุนที่ต้องการวิธีการสุ่มตัวอย่าง ตัวอย่างเช่น ผลิตภัณฑ์ที่ผลิตจำนวนมาก (วิทยุ, สกรู, ยางลบ, ดินสอ ฯลฯ) เราไม่สามารถตรวจสอบแต่ละชิ้นได้ เพราะจะสิ้นค่าใช้จ่ายและเวลามาก จึงไม่ต้องสงสัยว่า การเรียนวิธีการทางสถิติจะช่วยได้มาก วิธีการนี้เป็นฐานของสถิติคณิตศาสตร์สมัยใหม่ทั้งหมด เพื่อได้ตัวอย่างเหล่านี้มา และจากตัวอย่างเหล่านี้ก็นำมาสรุปเกี่ยวกับประชากรที่ได้ตัวอย่างมา

ขั้นตอนของการสำรวจทางสถิติ

ชนิดของปัญหาที่สามารถนำสถิติคณิตศาสตร์มาใช้ค่อนข้างจะเหมือนกัน ขั้นตอนของการสำรวจทางสถิติเกือบทุกกรณีเหมือนกัน

ขั้นตอนที่ 1 การกำหนดกฎเกณฑ์ของปัญหา ลำดับสำรวจปัญหา เริ่มแรกเราต้องสร้างแนวทางศึกษาที่แน่นอน กำหนดคำถามชัดเจน และตั้งขอบเขตให้เหมาะสมกับปัญหา เวลาที่ใช้ จำนวนเงิน และความชำนาญของผู้สำรวจ แนวทางศึกษา อย่างเช่น สิ่งของที่เสีย โรคจิต และการบริการลูกค้า อาจแปรจากกรณีหนึ่งไปยังอีกกรณีหนึ่ง และในแต่ละกรณีเฉพาะ เราต้องทำตามคำจำกัดความที่เหมาะสมของเทอมที่เกี่ยวข้อง

ขั้นตอนที่ 2 ออกแบบการทดลอง ความต้องการของเรา คือ ต้องการหารายละเอียดให้มากที่สุด โดยใช้เวลาและทุนน้อยที่สุด นั่นคือ เราต้องกำหนดขนาดของตัวอย่างหรือจำนวน และชนิดของข้อมูลที่จะแก้ปัญหาให้มีประสิทธิภาพมากที่สุด และเราต้องเลือกวิธีการนี้กับวิธีการสุ่มตัวอย่าง เมื่อเทียบกับวิธีการหลัง มันก็ไม่ง่ายนักที่ได้วิธีการเลือกที่เป็นแบบสุ่มได้อย่างสมบูรณ์ มีอันตรายอยู่เสมอที่การเลือก อาจเป็นวิธีการบางอย่างที่มีความเอียงเอนได้ ข้อเสนอแนะการเอาชนะความยากลำบากนี้ได้ก็ต้องศึกษาวิธีการเฉพาะในหนังสือการสำรวจตัวอย่าง

ขั้นตอนที่ 3 การรวบรวมข้อมูล โดยทั่ว ๆ ไป การรวบรวมข้อมูลเป็นขั้นตอนที่ใช้เวลามากที่สุดของการสำรวจทั้งหมด ควรยึดกฎอย่างเหนียวแน่น ผลลัพธ์ก็จะดีขึ้น

ขั้นตอนที่ 4 การบรรยายผลลัพธ์ที่ได้ ในขั้นนี้ใส่ข้อมูลที่ได้จากการทดลองลงในแบบที่สามารถอ่านได้ และแสดงได้ด้วยกราฟ (แผนภาพ แผนภาพแท่ง ฯลฯ) นอกจากนั้น คำนวณค่าเฉลี่ยการกระจายจากค่าสังเกตของตัวอย่าง

ขั้นตอนที่ 5 การอนุมานทางสถิติและการกำหนดกฎเกณฑ์ของคำตอบ ใช้วิธีการทางสถิติที่เลือกในขั้นที่ 2 เราเลือกสรุปจากตัวอย่างเกี่ยวกับประชากรที่เลือกตัวอย่างมา และกำหนดคำตอบกับปัญหาของเรา