บทที่ 2

ทฤษฎีและผลงานที่เกี่ยวข้องกับการวิจัย

คำว่า สถิติ ตรงกับคำในภาษาอังกฤษว่า statistics เป็นคำที่แปลมาจากคำว่า statistik ในภาษาเยอรมัน ซึ่ง Gottfried Ackenwell นักปรัชญาชาวเยอรมัน เป็นผู้บัญญัติขึ้น ในปี ค.ศ. 1749 โดยมีความหมายถึง ข้อมูลหรือข่าวสารที่เกี่ยวข้องและจะเป็นประโยชน์ต่อการบริหารของรัฐ

ในปัจจุบัน ความหมายคำว่า สถิติ ได้ถูกกำหนดให้มีความหมายกว้างขวางออกไปจากความหมายดั้งเดิมมาก โดยจำแนกออกได้เป็น 2 ความหมาย คือ

ประการแรก ถ้ามองในแง่ของศาสตร์ สถิติ หมายถึงศาสตร์ที่ว่าด้วยการเก็บข้อมูลซึ่งแสดงข้อเท็จจริงต่าง ๆ และการประมวลผล หรือดำเนินการกับข้อมูลเหล่านั้น โดยครอบคลุมไปถึง การนำเสนอข้อมูล การวิเคราะห์เพื่อหาข้อสรุปต่าง ๆ ที่เกี่ยวกับลักษณะหรือธรรมชาติของข้อมูล การนำข้อมูลไปใช้ในการพยากรณ์ หรือนำไปประกอบการตัดสินใจเกี่ยวกับอนาคต

เมื่อพิจารณากระบวนการปฏิบัติงานในขั้นตอนต่าง ๆ ของสถิติศาสตร์ จะเห็นได้ว่าสถิติศาสตร์ เป็นทั้งศิลป (Art) และ วิทยาศาสตร์ (Science) ทั้งนี้เนื่องจากวิธีการต่าง ๆ ทางสถิติ เป็นวิธีการทางวิทยาศาสตร์ที่อาศัยการหาข้อสรุปอย่างมีเหตุผล ภายในขอบเขตของข้อสมมติที่ตั้งขึ้น

อย่างไรก็ตาม นักสถิติยังต้องใช้ประสบการณ์และวิจารณญาน ในการเลือกรูปแบบ และวิธีการเก็บรวบรวมข้อมูล ซึ่งในส่วนหลังนี้จัดได้ว่า เป็นศิลปแขนงหนึ่งเหมือนกัน

ประการที่สอง พิจารณาในแง่ของตัวเลข สถิติ หมายถึง ตัวเลขที่ได้มาด้วยวิธีการใด ๆ จากข้อมูลจำนวนมาก ๆ ที่เก็บรวบรวมมาได้

โดยทั่วไปมักจะเรียกสถิติในความหมายนี้ว่า ข้อมูลสถิติ (statistical data)

สถิติศาสตร์ เริ่มมีบทบาทในการประยุกต์ใช้กับศาสตร์อื่น ๆ ในปลายศตวรรษที่ 19 โดยการคิดค้นของนักปรัชญา และนักทดลองชาวอังกฤษ 2 ท่าน คือ เซอร์ ฟรานซีส กัลตัน (Sir Francis Galton : ค.ศ. 1822-1911) และ คาร์ล เพียร์สัน (Karl Pearson: ค.ศ. 1857-1936)

ในปี ค.ศ. 1893 ท่านทั้งสองได้นำเอาทฤษฎีความน่าจะเป็น และทฤษฎีของความคลาดเคลื่อน (Theory of Error) ไปใช้ในการวิเคราะห์ข้อมูล ที่ได้จากการทดลองทางการเกษตร พันธุ์ศาสตร์ และชีววิทยา

ในราว ค.ศ. 1895-1930 ถือได้ว่าเป็นช่วงเวลาที่วิวัฒนาการทางสถิติสมัยใหม่ได้รุดหน้าไปอย่างรวดเร็ว แนวคิดที่สำคัญ ๆ ในทางทฤษฎี และ เทคนิคทางการวิเคราะห์เชิงสถิติ ได้เกิดขึ้นในคาบเวลานี้ทั้งสิ้น นักสถิติผู้มีชื่อเสียงและได้รับการยกย่องว่ามีส่วนสำคัญในการวางรากฐานให้แก่ทฤษฎีสถิติสมัยใหม่ ได้แก่ เซอร์ โรนัลด์ ไอล์เมอร์ ฟิเชอร์ (Sir Ronald Aylmer Fisher: ค.ศ. 1890-1962) เจอร์ซี เนย์แมน (Jerzey Neyman: ค.ศ. 1894-1984) และ อีกอน เพียร์สัน (Egon Pearson: ค.ศ. 1895-1980)

ในบทนี้จะกล่าวถึงรายละเอียดของสถิติศาสตร์ ในส่วนที่เป็นตัวทฤษฎี ซึ่งใช้อ้างอิงในการวิจัย ดังรายละเอียดต่อไปนี้

2.1 แนวคิดของความน่าจะเป็น

ทฤษฎีความน่าจะเป็น (Probability Theory) นับเป็นรากฐานที่สำคัญที่สุดในการสร้างทฤษฎีการแจกแจง (Distribution Theory) ตัวทฤษฎีการแจกแจงนั้นเปรียบได้กับกระดูกสันหลังของ ทฤษฎีการอนุมานเชิงสถิติ (Theory of Statistical Inference) ซึ่งมีการประยุกต์ใช้กับศาสตร์อื่นๆอย่างกว้างขวาง แต่เป็นที่น่าเสียดายที่ในปัจจุบัน แนวคิดเกี่ยวกับความน่าจะเป็นของทั้งนักสถิติและนักคณิตศาสตร์ ได้แตกแยกออกเป็น 2 ฝ่าย โดยฝ่ายหนึ่งเห็นว่า ความน่าจะเป็น คือ คุณสมบัติภายใน (intrinsic property)* ส่วนอีกฝ่ายเห็นว่า ความน่าจะเป็น คือ มาตรวัดระดับความเชื่อ (measure of belief or degree of belief)**

ในปัจจุบัน ไม่เพียงแต่จะมีแนวความคิดที่แตกกันออกเป็น 2 ทางเท่านั้น แม้แต่การให้นิยามของความน่าจะเป็น ก็ไม่สามารถให้ความหมายที่เป็นอันหนึ่งอันเดียวกัน ซึ่งทุกฝ่ายให้การยอมรับ ทำให้มีการใช้นิยามของความน่าจะเป็นในความหมายที่แตกต่างกันซึ่งพอจะกล่าวได้เป็น 3 นัย หรือแนวทางดังนี้

1. นิยามตามแนวความน่าจะเป็นแบบคลาสสิค
2. นิยามตามแนวความน่าจะเป็นในรูปความถี่
3. นิยามตามแนวความน่าจะเป็นอัตตวิสัย

---

* เรียกแนวความคิดในลักษณะนี้ว่า Objective Probability

** แนวความคิดในลักษณะนี้เรียกว่า Subjective Probability

2.1.1 <u>ความน่าจะเป็นแบบคลาสสิค</u> (Classical Probability)

นิยามของความน่าจะเป็นตามแนวคลาสสิคนี้ เข้าใจกันว่าเกิดขึ้นก่อนนิยามตามแนวอื่น ๆ โดยมีหลักฐานปรากฏในคริสต์ศตวรรษที่ 17 เป็นคำอธิบายของ แบลส ปาสกาล (Blaise Pascal: ค.ศ. 1623-1662) เกี่ยวกับเกมการพนัน (Games of chance) ให้แก่ เชอวาริเย เดอ เมเร (Chevalier de Mere) นักการพนันที่มีชื่อในยุคนั้น คำถามของ เดอ เมเร มีใจความว่า

เนื่องจากเขาได้ชนะการพนันในเกมทอดลูกเต๋า 1 ลูก 4 ครั้ง ซึ่งเขาพนันว่าจะขึ้นแต้ม 6 อย่างน้อย 1 ครั้ง จึงทำให้เขาพนันต่อไปว่า ถ้าทอดลูกเต๋า 2 ลูก 24 ครั้ง จะขึ้นแต้ม 6 ทั้งคู่ อย่างน้อย 1 ครั้ง แต่ปรากฏว่าครั้งนี้เขาแพ้พนัน จึงได้มาปรึกษา เพื่อขอคำอธิบายจากปาสกาล

คำอธิบายของปาสกาล ในเรื่องการทอดลูกเต๋ามีดังนี้ ปาสกาลอธิบายว่า ในการทอดลูกเต๋า 1 ลูก 1 ครั้ง ความน่าจะเป็นที่ลูกเต๋าจะขึ้นแต้ม 6 คือ 1/6 ความน่าจะเป็นที่ลูกเต๋าจะไม่ขึ้นแต้ม 6 ในการทอด 1 ครั้ง คือ 5/6

ในการทอดลูกเต๋า 1 ลูก 4 ครั้ง ความน่าจะเป็นที่ลูกเต๋าจะไม่ขึ้นแต้ม 6 คือ $(5/6)^4$ เท่ากับ 0.4823 และความน่าจะเป็นที่ลูกเต๋าจะขึ้นแต้ม 6 อย่างน้อย 1 ครั้ง คือ $1 - (5/6)^4$ เท่ากับ 0.5177 ซึ่งเกินครึ่ง แสดงว่าโอกาสที่ เดอ เมเร จะชนะมีมาก

แต่เมื่อทอดลูกเต๋า 2 ลูก พร้อมกัน ความน่าจะเป็นที่จะไม่ได้แต้ม 6 ทั้งคู่ คือ 35/36 เมื่อทอดลูกเต๋า 2 ลูก 24 ครั้ง ความน่าจะเป็นที่จะได้แต้ม 6 ทั้ง 2 ลูก อย่างน้อย 1 ครั้ง คือ $1 - (35/36)^4$ เท่ากับ 0.4914 ซึ่งน้อยกว่าครึ่ง เดอ เมเร จึงมีโอกาสที่จะแพ้พนัน

คำอธิบายของ ปาสกาล นับเป็นตัวอย่างของการคำนวณความน่าจะเป็นตามแนวคลาสสิคที่เก่าแก่มากชิ้นหนึ่ง ก่อนที่จะกล่าวถึงนิยามของความน่าจะเป็นแบบคลาสสิค จะขอทบทวนพื้นฐานบางประการเสียก่อน

การทดลองเชิงสุ่ม (Random experiment) เป็นการกระทำใด ๆ ที่ผู้กระทำไม่สามารถทำนายผลได้ล่วงหน้า และ จะเสร็จสิ้นลงด้วยผล (Outcomes) อย่างใดอย่างหนึ่งเท่านั้น เซตของผลทั้งหมดที่เป็นไปได้จากการทดลอง เรียกว่า แซมเปิลสเปซ (Sample space) สมาชิกแต่ละตัวของแซมเปิลสเปซ เรียกว่า แซมเปิลพอยท์ (Sample point)

เหตุการณ์ (Event) หมายถึง สับเซตของแซมเปิลสเปซ สมาชิกแต่ละตัวของเหตุการณ์ คือ แซมเปิลพอยท์ ที่อยู่ในแซมเปิลสเปซ เหตุการณ์ที่มีสมาชิกเพียง 1 ตัว จะเรียกว่าเหตุการณ์อย่างง่าย (Simple event) ส่วนเหตุการณ์ที่มีสมาชิกมากกว่า 1 ตัว เรียกว่าเหตุการณ์เชิงประกอบ (Compound event)

ผลลัพธ์ที่มีโอกาสเกิดขึ้นเท่า ๆ กัน (Equally likely outcome) หมายถึง แซมเปิลพอยท์ทุกตัวที่อยู่ในแซมเปิลสเปซ จะมีโอกาสในการเกิดขึ้นเท่ากันทุกตัว

ในการทดลอง เราสนใจผลลัพธ์ของการทดลองที่จะเกิดขึ้น การหาความน่าจะเป็นของการเกิดเหตุการณ์ใดเหตุการณ์หนึ่ง ใช้วิธีการคำนวณจากผลลัพธ์ของการทดลองที่เป็นไปได้ทั้งหมดที่เรียกว่า แซมเปิลสเปซ โดยหาอัตราส่วนระหว่างจำนวนแซมเปิลพอยท์ที่เป็นสมาชิกในเหตุการณ์ กับจำนวนแซมเปิลพอยท์ทั้งหมดในแซมเปิลสเปซ โดยมีเงื่อนไขว่า แซมเปิลพอยท์แต่ละตัวที่เป็นสมาชิกในแซมเปิลสเปซ และในเหตุการณ์ จะต้องมีโอกาสเกิดขึ้นได้เท่า ๆ กัน เรานิยามความน่าจะเป็นของเหตุการณ์ใด ๆ ได้ดังนี้

<u>นิยาม 2.1.1</u> ความน่าจะเป็นของเหตุการณ์

ถ้าแซมเปิลสเปซของการทดลองเชิงสุ่ม มีสมาชิก N ตัว แต่ละตัวมีโอกาสเกิดขึ้นเท่า ๆ กัน และเหตุการณ์ E ซึ่งเกิดจากการทดลองนี้ประกอบ

ด้วยสมาชิก m ตัว ความน่าจะเป็นที่เหตุการณ์ E จะเกิดขึ้นมีค่าเท่ากับ m/n เขียนแทนด้วย

$$P(E) = \frac{m}{n}$$

2.1.2 ความน่าจะเป็นในรูปความถี่[1] (Frequency Probability)

การนิยามความน่าจะเป็นตามแนวความน่าจะเป็นในรูปความถี่ มีกำเนิดมาจากการสังเกตผลของการทดลองเชิงสุ่ม ที่ทำซ้ำ ๆ กันภายใต้สภาวการณ์เดียวกัน แล้วบันทึกค่าความถี่ของผลลัพธ์ที่เกิดขึ้น ตัวอย่างของการทดลองในลักษณะนี้ได้บันทึกเป็นหลักฐานไว้โดย J.E. Kerrick (1946) ในผลงานชื่อ " An Experimental Introduction to the Theory of Probability." ผลงานดังกล่าว Kerrick ได้เป็นผู้กระทำการทดลองเองทั้งหมด ในระหว่างที่เขาถูกกักกันอยู่ในคุกชื่อ Hald ที่ประเทศเดนมาร์ก ในช่วงสงครามโลกครั้งที่ 2 เขาได้ทำการทดลองหลายการทดลอง โดยหนึ่งในนั้นเป็นการทดลองโยนเหรียญซ้ำ ๆ กัน เป็นจำนวน 10,000 ครั้ง แล้วนับจำนวนที่ออกหัว ผลปรากฏว่า นับจำนวนครั้งที่ออกหัวได้ 5,067 ครั้ง จากนั้นเขาได้คำนวณหาความถี่สัมพัทธ์ของจำนวนหัวที่เกิดขึ้น ได้เท่ากับ 5,067/10,000 ซึ่งมีค่าเท่ากับ 0.5067

---

[1]ปรัชญาเกี่ยวกับความน่าจะเป็นในรูปความถี่ ศึกษาได้จาก Ian Hacking, Logic of statistical inference (Cambridge: Cambridge University Press, 1965) , pp. 1-26.

การนิยามความน่าจะเป็นในรูปความถี่มีแนวคิดดังนี้ ถ้าให้ E เป็นเหตุการณ์ที่เราสนใจในการทดลองเชิงสุ่ม สมมุติว่าในการทำการทดลองเชิงสุ่มซ้ำกัน n ครั้ง เหตุการณ์ E เกิดขึ้นทั้งหมด m ครั้ง ค่าอัตราส่วน m/n เรียกว่า ความถี่สัมพัทธ์ (relative frequency) สำหรับค่าความถี่สัมพัทธ์ ที่เกิดจากการทดลองซึ่งมีจำนวนครั้งที่มากพอจนทำให้ค่าความถี่สัมพัทธ์เข้าใกล้ค่าคงที่ค่าใดค่าหนึ่ง จะกำหนดให้ค่าคงที่ดังกล่าว คือ ค่าความน่าจะเป็น เรานิยามความน่าจะเป็นของเหตุการณ์ใด ๆ ได้ดังนี้

<u>นิยาม 2.1.2</u> ความน่าจะเป็นของเหตุการณ์จากการทดลอง*

ถ้าเหตุการณ์ E เกิดขึ้นได้ m ครั้ง จากการทดลองเชิงสุ่มที่ทำซ้ำกัน n ครั้ง ความน่าจะเป็นที่เหตุการณ์ E จะเกิดขึ้น มีค่าเท่ากับ

$$P(E) = \lim_{n \to \infty} \frac{m}{n}$$

2.1.3 <u>ความน่าจะเป็นอัตตวิสัย</u>** (Personal Probability)

นิยามความน่าจะเป็นตามแนวความน่าจะเป็นอัตตวิสัย คือ การใช้ความรู้สึกของมนุษย์ เป็นเครื่องมือในการวัดค่าความน่าจะเป็นของเหตุการณ์ใดเหตุการณ์หนึ่งว่าควรจะเป็นเท่าใด ค่าความน่าจะเป็นดังกล่าว จะขึ้นอยู่กับตัวบุคคลที่เป็นผู้กำหนด

---

*หนังสือบางเล่มเรียกว่า long run relative frequency หรือ long run frequency

**หนังสือส่วนใหญ่จะใช้คำว่า Subjective Probability

ค่าความน่าจะเป็นนั้น ซึ่งการที่ค่าความน่าจะเป็นจะมีความถูกต้องใกล้เคียงกับความจริงมากน้อยเพียงใดนั้น ย่อมขึ้นอยู่กับ ประสบการณ์ ความเชี่ยวชาญ และข้อมูลต่าง ๆ ที่บุคคลนั้นมีอยู่ทำให้ค่าความน่าจะเป็นของเหตุการณ์เดียวกัน ที่ถูกกำหนดโดย บุคคล 2 คน สามารถมีค่าที่แตกต่างกันได้[2]

แนวความคิดแบบอัตตวิสัยนี้ ถูกนำเสนอขึ้นเป็นครั้งแรก ในผลงาน ชื่อ " A Treatise on Probability." โดย J.M. Keynes ในปี ค.ศ. 1921 ติดตามด้วยผลงานของ F.P. Ramsey ที่ชื่อ " Truth and Probability." ในปี ค.ศ. 1931 และต่อมาได้ถูกนำเสนอในหนังสือ " Theory of Probability." ของ H. Jeffreys ซึ่งตีพิมพ์เป็นครั้งแรกในปี ค.ศ. 1939 (พิมพ์ครั้งที่สอง ปี ค.ศ. 1940 และครั้งที่สาม ในปี ค.ศ. 1961)

---

[2]ศึกษาเพิ่มเติมใน L.J. Savage, The Foundation of statistics, 2 nd rev. ed.(New York: Dover Publication Inc.,1972), pp. 27-68.

2.2 ทฤษฎีการแจกแจง

ทฤษฎีการแจกแจง (Distribution Theory) เป็นเรื่องราวเกี่ยวกับการแจกแจงความน่าจะเป็นของตัวแปรสุ่ม เราทราบมาแล้วว่า ตัวแปรสุ่ม คือ ฟังก์ชันค่าจริงที่เปลี่ยนแซมเปิลพอยท์ทุกแซมเปิลพอยท์ในแซมเปิลสเปซ ให้เป็นตัวเลขจำนวนจริง การแบ่งประเภทของตัวแปรสุ่ม ทำได้โดยใช้ค่าของตัวแปรสุ่มเป็นเกณฑ์ในการจำแนก

ถ้าค่าทั้งหมดที่เป็นไปได้ของตัวแปรสุ่ม มีจำนวนจำกัดสามารถนับได้ถ้วน หรือมีค่าเป็นจำนวนอนันต์แต่สามารถนับได้ กล่าวคือ สามารถนำค่าแต่ละค่าไปจับคู่แบบหนึ่งต่อหนึ่งกับเลขจำนวนนับได้ เราจะจัดให้ตัวแปรสุ่มที่มีลักษณะดังกล่าว อยู่ในกลุ่มของตัวแปรสุ่มชนิดไม่ต่อเนื่อง

ถ้าค่าที่เป็นไปได้ของตัวแปรสุ่มมีจำนวนอนันต์ และไม่สามารถนำค่าแต่ละค่าไปจับคู่แบบหนึ่งต่อหนึ่งกับเลขจำนวนนับ กล่าวคือ ค่าของตัวแปรสุ่มมีค่าต่อเนื่องกันได้หลายค่าจนนับไม่ถ้วน เราจะจัดให้ตัวแปรสุ่มที่มีลักษณะดังกล่าว อยู่ในกลุ่มของตัวแปรสุ่มชนิดต่อเนื่อง

ในกรณีของตัวแปรสุ่มชนิดไม่ต่อเนื่อง (Discrete random variable) เราสามารถหาค่าความน่าจะเป็นของการเกิดค่าของตัวแปรสุ่มค่าหนึ่ง ๆ ได้จาก ฟังก์ชันความน่าจะเป็นแบบไม่ต่อเนื่อง* (Discrete probability function)

ส่วนกรณีของตัวแปรสุ่มชนิดต่อเนื่อง (Continuous random variable) ค่าความน่าจะเป็นของการเกิดค่าของตัวแปรสุ่ม ที่ค่าใด ๆ มีค่าเท่ากับศูนย์ การพิจารณาเกี่ยวกับความน่าจะเป็น จะเป็นการหาค่าความน่าจะเป็นของค่าตัวแปรสุ่ม ที่เกิดขึ้นในช่วง

---

*หนังสือบางเล่มใช้คำว่า probability mass function (p.m.f) เมื่อกล่าวถึง ฟังก์ชันความน่าจะเป็นแบบไม่ต่อเนื่อง

หนึ่ง ๆ แทน โดยหาได้จากการอินทิเกรต ฟังก์ชันความน่าจะเป็นแบบต่อเนื่อง* (Continuous probability function)

ตัวแปรสุ่ม สามารถมีค่าได้หลายค่าด้วยความน่าจะเป็นต่าง ๆ กัน ทั้งนี้ขึ้นอยู่กับว่า ตัวแปรสุ่มนั้นมีฟังก์ชันการแจกแจงความน่าจะเป็นในรูปแบบใด เมื่อนำเอาฟังก์ชันความน่าจะเป็นมาพลอตลงกระดาษกราฟ จะทำให้เราทราบถึงรูปร่างของฟังก์ชันการแจกแจงความน่าจะเป็นของตัวแปรสุ่มนั้น

ค่าเฉลี่ย (mean) และ ค่าความแปรปรวน (variance) ของตัวแปรสุ่ม เป็นตัวเลขที่บอกถึงลักษณะทางสถิติโดยรวม ๆ ของการแจกแจงของตัวแปรสุ่มนั้น ตัวเลขดังกล่าวจัดเป็นข้อมูลพื้นฐานของฟังก์ชันการแจกแจง โดยค่าเฉลี่ยหรือค่าคาดคะเนเชิงคณิตศาสตร์ (Mathematical expectation) คือตัวกลางเลขคณิตของตัวแปรสุ่ม เป็นค่าที่ถือว่าเป็นตัวแทนที่ดีของการแจกแจง จุดที่เป็นค่าคาดคะเนเชิงคณิตศาสตร์ เปรียบได้กับจุดศูนย์กลางของการแจกแจง (Center of Distribution) ส่วนค่าความแปรปรวนของตัวแปรสุ่มเป็นค่าที่บอกให้ทราบว่า ค่าต่าง ๆ ของตัวแปรนั้น มีการกระจายออกจากค่าเฉลี่ยของตัวแปรสุ่มโดยเฉลี่ยเป็นเท่าใด และสามารถใช้บอกลักษณะความกว้างบริเวณฐานของโค้งการแจกแจงได้ กล่าวคือ โค้งที่มีความแปรปรวนน้อย ฐานโค้งจะแคบและยอดโค้งจะสูง ส่วนโค้งที่มีค่าความแปรปรวนมากฐานโค้งจะกว้าง และยอดโค้งจะต่ำ

รายละเอียดของการแจกแจงความน่าจะเป็นแบบต่าง ๆ ที่กล่าวถึงในการวิจัยครั้งนี้ สำหรับการแจกแจงแบบไม่ต่อเนื่อง และการแจกแจงแบบต่อเนื่อง แสดงดังตารางที่ 2.1

---

*หนังสือหลายเล่มใช้คำว่า probability density function (p.d.f.) เมื่อกล่าวถึง ฟังก์ชันความน่าจะเป็นแบบต่อเนื่อง

ตารางที่ 2.1 การแจกแจงแบบต่อเนื่องและแบบไม่ต่อเนื่องบางชนิด

Name of parametric family of distributions	Cumulative distribution function $F(\cdot)$ or probability density function $f(\cdot)$	Parameter space	Mean $\mu = \mathscr{E}[X]$
Bernoulli	$f(x) = p^x q^{1-x} I_{\{0,\,1\}}(x)$	$0 \le p \le 1$ $(q = 1 - p)$	p
Binomial	$f(x) = \binom{n}{x} p^x q^{n-x} I_{\{0,\,1,\,\ldots,\,n\}}(x)$	$0 \le p \le 1$ $n = 1, 2, 3, \ldots$ $(q = 1 - p)$	np
Normal	$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp[-(x-\mu)^2/2\sigma^2]$	$-\infty < \mu < \infty$ $\sigma > 0$	μ
Beta	$f(x) = \frac{1}{B(a,\,b)} x^{a-1}(1-x)^{b-1} I_{(0,1)}(x)$	$a > 0$ $b > 0$	$\frac{a}{a+b}$
F disuibution	$f(x) = \frac{\Gamma[(m+n)/2]}{\Gamma(m/2)\Gamma(n/2)} \left(\frac{m}{n}\right)^{m/2} \times \frac{x^{(m-2)/2}}{[1 + (m/n)x]^{(m+n)/2}} I_{(0,\infty)}(x)$	$m, n = 1, 2, \ldots$	$\frac{n}{n-2}$ for $n > 2$
Chi-square distribution	$f(x) = \frac{1}{\Gamma(k/2)} \left(\frac{1}{2}\right)^{k/2} x^{k/2-1} e^{-(1/2)x} I_{(0,\infty)}(x)$	$k = 1, 2, \ldots$	k

ที่มา: A.M. Mood, F.A. Graybill and D.C. Boes, <u>Introduction to the Theory of Statistics</u>, 3 rd ed. (New York: McGraw-Hill Book Co., 1974), pp. 538-543.

ตารางที่ 2.1 การแจกแจงแบบต่อเนื่องและแบบไม่ต่อเนื่องบางชนิด (ต่อ)

Variance $\sigma^2 = \mathscr{E}[(X-\mu)^2]$	Moments $\mu_r' = \mathscr{E}[X^r]$ or $\mu_r = \mathscr{E}[(X-\mu)^r]$ and/or cumulants κ_r	Moment generating function $\mathscr{E}[e^{tX}]$
pq	$\mu_r' = p$ for all r	$q + pe^t$
npq	$\mu_3 = npq(q-p)$ $\mu_4 = 3n^2p^2q^2 + npq(1-6pq)$	$(q + pe^t)^n$
σ^2	$\mu_r = 0$, r odd; $\mu_r = \frac{r!}{(r/2)!}\frac{\sigma^r}{2^{r/2}}$, r even; $\kappa_r = 0$, $r > 2$	$\exp[\mu t + \frac{1}{2}\sigma^2 t^2]$
$\frac{ab}{(a+b+1)(a+b)^2}$	$\mu_r' = \frac{B(r+a,b)}{B(a,b)}$	not useful
$\frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}$ for $n > 4$	$\mu_r' = \left(\frac{n}{m}\right)^r \frac{\Gamma(m/2+r)\Gamma(n/2-r)}{\Gamma(m/2)\Gamma(n/2)}$ for $r < \frac{n}{2}$	does not exist
$2k$	$\mu_j' = \frac{2^j\Gamma(k/2+j)}{\Gamma(k/2)}$	$\left(\frac{1}{1-2t}\right)^{k/2}$ for $t < 1/2$

ทฤษฎีที่จะนำเสนอต่อไป เป็นความสัมพันธ์ระหว่างการแจกแจงบางการแจกแจงที่อ้างถึงในการวิจัย

ทฤษฎีที่ 2.2.1 ความสัมพันธ์ระหว่างการแจกแจงทวินาม กับการแจกแจงเบอร์นูลี*

ถ้า $X_1, X_2, \ldots, X_n$ เป็นตัวแปรสุ่มที่มีการแจกแจงเบอร์นูลี ที่มีพารามิเตอร์ p เหมือนกันและเป็นอิสระจากกัน จะได้ว่าตัวแปรสุ่ม Y ที่กำหนดโดย

$$Y = X_1 + X_2 + \ldots + X_n$$

จะมีการแจกแจงทวินามที่มีพารามิเตอร์ (n,p)

ทฤษฎีที่ 2.2.2 ความสัมพันธ์ระหว่างการแจกแจงยูนิฟอร์ม กับการแจกแจงเบต้า

ถ้า X เป็นตัวแปรสุ่มที่มีการแจกแจงยูนิฟอร์มในช่วง $(0,1)$ จะได้ว่า ตัวแปรสุ่ม X ดังกล่าวคือ ตัวแปรสุ่มเบต้าที่มีพารามิเตอร์ $(1,1)$

*ตัวแปรสุ่มเบอร์นูลี ที่มีพารามิเตอร์เท่ากับ p คือตัวแปรสุ่มทวินามที่มีพารามิเตอร์ $(1,p)$

<u>ทฤษฎีที่ 2.2.3</u> ความสัมพันธ์ระหว่างการแจกแจงปกติมาตรฐาน กับการแจกแจงไคสแควร์

ถ้า $X_1, X_2, \ldots, X_n$ เป็นตัวแปรสุ่มที่มีการแจกแจงแบบปกติมาตรฐานเหมือนกัน และเป็นอิสระจากกัน จะได้ว่าตัวแปรสุ่ม Q_k ที่กำหนดโดย

$$Q_k = X_1^2 + X_2^2 + \ldots + X_k^2$$

จะมีการแจกแจงไคสแควร์ที่มีค่าองศาอิสระเท่ากับ k

<u>ทฤษฎีที่ 2.2.4</u> ความสัมพันธ์ระหว่างการแจกแจงไคสแควร์ กับการแจกแจงเอฟ

กำหนดให้ X_1 เป็นตัวแปรสุ่มไคสแควร์ ที่มีค่าองศาอิสระเท่ากับ m

X_2 เป็นตัวแปรสุ่มไคสแควร์ ที่มีค่าองศาอิสระเท่ากับ n

และ X_1 กับ X_2 เป็นอิสระจากกัน จะได้ว่าตัวแปรสุ่ม W ที่กำหนดโดย

$$W = \frac{X_1 / m}{X_2 / n}$$

จะมีการแจกแจงเอฟที่มีพารามิเตอร์ (m,n)

<u>ทฤษฎีที่ 2.2.5</u> ความสัมพันธ์ระหว่างการแจกแจงเอฟ กับการแจกแจงเบต้า*

ถ้า X เป็นตัวแปรสุ่มเอฟที่มีพารามิเตอร์ (m,n) จะได้ว่าตัวแปรสุ่ม U และ V ที่กำหนดโดย

$$U = \frac{n}{n + mX}$$ จะมีการแจกแจงเบต้าที่มีพารามิเตอร์ (n/2,m/2)

$$V = \frac{mX}{n + mX}$$ จะมีการแจกแจงเบต้าที่มีพารามิเตอร์ (m/2,n/2)

<u>ทฤษฎีที่ 2.2.6</u> ความสัมพันธ์ระหว่างการแจกแจงไคสแควร์ กับการแจกแจงเบต้า

กำหนดให้ X_1 เป็นตัวแปรสุ่มไคสแควร์ที่มีองศาอิสระเท่ากับ 2a
X_2 เป็นตัวแปรสุ่มไคสแควร์ที่มีองศาอิสระเท่ากับ 2b
และ X_1 กับ X_2 เป็นอิสระจากกัน จะได้ว่าตัวแปรสุ่ม Z ที่กำหนดโดย

$$Z = \frac{X_1}{X_1 + X_2}$$

จะมีการแจกแจงเบต้าที่มีพารามิเตอร์ (a,b)

---

* ถ้า X เป็นตัวแปรสุ่มเบต้าที่มีพารามิเตอร์ (a,b) จะได้ว่า ตัวแปรสุ่ม Y ที่กำหนด โดย $Y = 1 - X$ จะมีการแจกแจงเบต้าที่มีพารามิเตอร์ (b,a)

2.3 ทฤษฎีค่าจำกัดสู่ส่วนกลาง

ในการสุ่มตัวอย่างขนาด n หน่วย จากประชากรที่มีการแจกแจงปกติ ซึ่งมีค่าเฉลี่ย μ และค่าความแปรปรวนเท่ากับ σ^2 เราทราบว่าค่าเฉลี่ยที่คำนวณจากตัวอย่างขนาด n จะมีการแจกแจงปกติที่มีค่าเฉลี่ยเท่ากับ μ และค่าความแปรปรวนเท่ากับ σ^2/n

ในกรณีที่สุ่มตัวอย่างขนาด n หน่วย จากประชากรที่มีการแจกแจงแบบอื่น ค่าเฉลี่ยที่คำนวณจากตัวอย่างขนาด n ดังกล่าว จะมีการแจกแจงเป็นรูปอื่นที่ไม่ใช่การแจกแจงแจงปกติ แต่ถ้าขนาดตัวอย่าง n มีค่าใหญ่พอ การแจกแจงของค่าเฉลี่ยตัวอย่างจะมีการแจกแจงใกล้เคียงกับการแจกแจงปกติโดยทฤษฎีค่าจำกัดสู่ส่วนกลาง กล่าวคือ ถ้าสุ่มตัวอย่างขนาด n หน่วยจากประชากรที่มีการแจกแจงอื่น ๆ ที่ไม่ใช่การแจกแจงปกติ โดยที่ประชากรดังกล่าวมีค่าเฉลี่ยเท่ากับ EX และ ค่าความแปรปรวนเท่ากับ VarX ในกรณีที่ค่าขนาดตัวอย่าง n มีค่าใหญ่พอ การแจกแจงของค่าเฉลี่ยตัวอย่างจะใกล้เคียงกับการแจกแจงปกติ ที่มีค่าเฉลี่ยเท่ากับ EX และค่าความแปรปรวนเท่ากับ (VarX)/n

ทฤษฎีค่าจำกัดสู่ส่วนกลาง นับเป็นหัวใจของทฤษฎีการอนุมานเชิงสถิติ การแจกแจงของผลบวกของตัวแปรสุ่มใด ๆ ที่เป็นแบบต่อเนื่องและไม่ต่อเนื่อง เมื่อกล่าวโดยลิมิตจะมีการแจกแจงลู่เข้าสู่การแจกแจงปกติ ภายใต้เงื่อนไขของทฤษฎีค่าจำกัดสู่ส่วนกลาง ใจความของทฤษฎีค่าจำกัดสู่ส่วนกลางมีดังนี้

<u>ทฤษฎีที่ 2.3.1</u> ทฤษฎีค่าจำกัดสู่ส่วนกลาง (The Central Limit Theorem)

กำหนดให้ $X_1, X_2, \ldots, X_n$ เป็นตัวแปรสุ่มที่มีการแจกแจงเหมือนกันและเป็นอิสระจากกัน โดยการแจกแจงดังกล่าว มีค่าเฉลี่ยเท่ากับ $E(X)$ และความแปรปรวนเท่ากับ $Var(X)$ ซึ่งทั้งคู่มีค่าจำกัด
ให้

$$\bar{X}_n = (X_1 + X_2 + \ldots + X_n) / n$$

จะได้ว่าในกรณีที่ n มีค่าใหญ่* ตัวสถิติ Y_n ซึ่งกำหนดโดย

$$Y_n = \frac{\sqrt{n}\,(\bar{X}_n - E(X))}{\sqrt{Var(X)}}$$

จะมีลิมิตของการแจกแจงลู่เข้าสู่การแจกแจงปกติมาตรฐาน

---

*เงื่อนไขของทฤษฎีค่าจำกัดสู่ส่วนกลางข้อหนึ่ง คือขนาดตัวอย่าง n ต้องใหญ่พอ ซึ่งในปัจจุบันยังไม่มีผู้ใดบอกได้แน่นอนว่า ค่า n ที่ใหญ่พอนั้นมีค่าเท่าใด

2.4 ทฤษฎีการประมาณค่าแบบช่วง[3]

ในการประมาณค่าพารามิเตอร์โดยใช้วิธีการประมาณแบบจุด ซึ่งเป็นวิธีการที่ใช้ค่าประมาณเพียงค่าเดียวที่คำนวณได้จากตัวอย่าง เป็นค่าประมาณสำหรับพารามิเตอร์ของประชากร เราไม่สามารถบอกได้ว่าค่าประมาณแบบจุดนั้น มีค่าเท่ากับค่าพารามิเตอร์หรือแตกต่างจากค่าพารามิเตอร์อย่างไร จึงได้มีแนวคิดที่จะนำเอาความคลาดเคลื่อนของการประมาณค่ามาเป็นเกณฑ์ในการวัดความเชื่อถือของค่าประมาณที่ได้ การทราบลักษณะบางประการของตัวประมาณแบบจุด เช่น ลักษณะการแจกแจง ค่าเฉลี่ย และค่าความแปรปรวน จะช่วยให้ทราบถึงความถูกต้องและเชื่อถือได้ของค่าประมาณเหล่านั้น

การประมาณค่าแบบช่วง เป็นแนวคิดที่นำเอาความแปรปรวนของตัวประมาณแบบจุด การแจกแจงตัวอย่าง (Sampling distribution) ของตัวประมาณ และค่าประมาณแบบจุดมาประกอบเข้าด้วยกัน เพื่อสร้างช่วงซึ่งเป็นสับเซตของค่าที่เป็นไปได้ของค่าพารามิเตอร์ขึ้นมา และอาศัยการแจกแจงของตัวประมาณเป็นตัวกำหนดความมั่นใจ ว่าช่วงที่สร้างขึ้นจะครอบคลุมค่าที่แท้จริงของพารามิเตอร์ด้วยความน่าจะเป็นเท่าใด

การประมาณค่าแบบช่วงของพารามิเตอร์ ส่วนมากจะอยู่ในรูปช่วง (a,b) ที่สมมติให้ค่าที่แท้จริงของพารามิเตอร์อยู่ภายในช่วงนี้ โดยที่ a และ b จะเป็นฟังก์ชันของค่าสังเกตที่ได้จากตัวอย่าง และจะขึ้นอยู่กับการแจกแจงของตัวประมาณแบบจุดของค่าพารามิเตอร์ กับค่าประมาณแบบจุดที่คำนวณได้จากตัวอย่าง ดังนั้นเมื่อสุ่มตัวอย่างมาหนึ่งชุด จะสามารถสร้างช่วง (a,b) ได้หนึ่งช่วง และคำนวณค่าประมาณแบบจุดได้หนึ่งค่า ถ้าสุ่มตัวอย่างมาหลาย ๆ ชุด จากประชากรเดียวกัน ก็จะสามารถคำนวณค่าประมาณแบบจุดได้หลายค่า ซึ่งจะมีค่าแตกต่างกันหรือไม่ก็ได้ ขึ้นกับชุดตัวอย่างที่สุ่มมา ทำให้ได้ช่วง (a,b) ที่อาจแตกต่างกันหรือไม่ก็ได้ หลายช่วง บางช่วงอาจจะครอบคลุมค่าพารามิเตอร์ และ

---

[3]ศึกษาเพิ่มเติมใน ประชุม สุวัตถี, ทฤษฎีการอนุมานเชิงสถิติ (กรุงเทพมหานคร: สำนักพิมพ์โอเดียนสโตร์, 2527) หน้า 318-369.

บางช่วงอาจจะไม่ครอบคลุมค่าพารามิเตอร์ จึงต้องมีการกำหนดระดับความเชื่อมั่นว่าช่วง (a,b) ที่คำนวณได้ มีความน่าจะเป็นร้อยละเท่าใด ที่จะครอบคลุมค่าที่แท้จริงของพารามิเตอร์ การกำหนดระดับความเชื่อมั่น จะกำหนดในรูปเปอร์เซนต์ของความเชื่อมั่น $(1 - \alpha)100$

ความหมายของช่วงความเชื่อมั่น $(1 - \alpha)100$ เปอร์เซนต์ อาจตีความได้ดังนี้ ถ้าเราทำการหาช่วง (a,b) ด้วยวิธีการอย่างเดียวกัน ซ้ำ ๆ กัน 100 ครั้ง เราจะได้ช่วง 100 ชุด ซึ่งอาจจะมีค่าแตกต่างกันหรือไม่ก็ได้ ขึ้นอยู่กับชุดตัวอย่าง และเราคาดหมายว่า จะมีช่วงจำนวน $(1 - \alpha)100$ ชุด ที่จะครอบคลุมค่าที่แท้จริงของพารามิเตอร์ หรือกล่าวอีกนัยหนึ่งได้ว่า เมื่อเราสร้างช่วง (a,b) ที่คำนวณได้จากตัวอย่างหนึ่งชุด เราจะมีความเชื่อมั่น $(1 - \alpha)100$ เปอร์เซนต์ ว่าช่วงนั้นจะคลุมค่าที่แท้จริงของพารามิเตอร์

ปกติระดับความเชื่อมั่นที่นิยมใช้คือ 0.90, 0.95 และ 0.99 ในการประมาณค่าแบบช่วงนั้น ช่วงความเชื่อมั่นที่มีระดับความเชื่อมั่นสูงกว่า จะมีความยาวช่วงมากกว่าช่วงความเชื่อมั่นที่มีระดับความเชื่อมั่นต่ำกว่า ในการใช้ประโยชน์ ค่าประมาณช่วงที่กว้างจะใช้ประโยชน์ได้น้อยกว่าค่าประมาณช่วงที่แคบกว่า สิ่งที่เราต้องการ คือ ช่วงความเชื่อมั่นที่มีความยาวช่วงสั้น ๆ แต่มีระดับความเชื่อมั่นสูง

การประมาณค่าแบบช่วง มีรากฐานทางคณิตศาสตร์เป็นผลงานริเริ่มของปิแยร์ ซิมง เดอ ลาปลาซ (Pierre Simon de Laplace : ค.ศ. 1749-1827) นักคณิตศาสตร์ชาวฝรั่งเศส ตั้งแต่ปี ค.ศ. 1812 ในผลงานชื่อ " Theorie Analytique des Probabilities." หลังจากนั้นได้มีการพัฒนาแนวความคิดนี้เรื่อยมา จนในปี ค.ศ. 1937 เจอร์ซี เนย์แมน (Jerzey Neyman: ค.ศ. 1894-1984) ได้เสนอผลงานชื่อ " Outline of a Theory of Statistical Estimation based on the Classical Theory of Probability." ผลงานดังกล่าว เป็นต้นกำเนิดของช่วงความเชื่อมั่น (Confidence Interval) ที่ใช้อยู่ในปัจจุบัน แนวคิดของเนย์แมน มีดังนี้

ในการประมาณค่าพารามิเตอร์ θ เราอาจจะใช้ฟังก์ชันของตัวอย่างสุ่ม $t(x_1,x_2,...x_n)$ โดยพิจารณารูปแบบการแจกแจงของตัวสถิติ t ดังกล่าว ซึ่งจะเป็นฟังก์ชันของพารามิเตอร์ θ และ จะพยายามหาสับเซตของค่าที่เป็นไปได้ทั้งหมดของพารามิเตอร์ θ ที่ทำให้เราเชื่อมั่นได้ว่า ค่าที่แท้จริงของ θ จะตกอยู่ในสับเซตดังกล่าว และจะเรียกสับเซต ซึ่งส่วนใหญ่เป็นช่วง (Interval) บนเส้นจำนวนจริงนี้ ว่า ช่วงความเชื่อมั่น

<u>นิยาม 2.4.1</u>[4]

ให้ $X_1, X_2, ..., X_n$ เป็นตัวอย่างสุ่มจากประชากรที่มีฟังก์ชันความน่าจะเป็น $f(x:\theta)$ เมื่อ θ เป็นพารามิเตอร์ที่ไม่ทราบค่า ตัวสถิติ $L(X_1,X_2,...,X_n)$ และ $U(X_1,X_2,...,X_n)$ เป็นตัวสถิติที่ให้ค่า L น้อยกว่า U และ ให้สอดคล้อง

$$Pr(L < \theta < U) = 1 - \alpha$$

โดยที่ α ไม่ขึ้นกับ θ และ $0 < \alpha < 1$

เราจะเรียก (L, U) ว่าช่วงสุ่ม (random interval)

---

[4]นิยามในลักษณะอื่น ๆ ที่เกี่ยวกับช่วงความเชื่อมั่น สามารถศึกษาเพิ่มเติมได้จาก Shelemyahu Zacks, <u>The Theory of Statistical Inference</u> (New York: John Wiley and Sons ,1971) ,pp. 499-582.

เมื่อกำหนดค่า $1 - \alpha$ ใด ๆ ขึ้นมา และจะเรียกช่วงสุ่ม (L, U) ที่สอดคล้อง

$$Pr(L < \theta < U) > 1 - \alpha \qquad (2.4.1)$$

สำหรับทุกค่าของ θ ว่าช่วงความเชื่อมั่น $(1 - \alpha)100$ เปอร์เซนต์ ของพารามิเตอร์ θ

ตัวเลข $1 - \alpha$ ซึ่งเป็นค่า อินฟิมัม (infimum) ของเทอมทางซ้ายมือของอสมการ 2.4.1 เรียกว่า ค่าสัมประสิทธิ์ความเชื่อมั่น (Confidence Coefficient)*

ค่าสถิติ L เรียกว่า ลิมิตล่างของช่วงความเชื่อมั่น (Lower Confidence Limit)

ค่าสถิติ U เรียกว่า ลิมิตบนของช่วงความเชื่อมั่น (Upper Confidence Limit)

---

*ค่า $Pr(L < \theta < U)$ ซึ่งเป็นค่าทางซ้ายมือของอสมการ 2.4.1 โดยปกติจะเรียกว่า ค่าระดับความเชื่อมั่น (Confidence Level) เฉพาะค่า infimum ซึ่งเท่ากับ $\min\{ Pr(L < \theta < U):$ สำหรับทุกค่าที่เป็นไปได้ของพารามิเตอร์ $\theta \}$ จึงจะเรียกว่า ค่าสัมประสิทธิ์ความเชื่อมั่น

2.5 <u>การทดสอบภาวะสารูปดีโดยใช้ไคสแควร์</u>

เนื้อหาส่วนสำคัญที่จัดเป็นหัวข้อใหญ่ในเรื่องการอนุมานเชิงสถิติอีกหัวข้อหนึ่ง นอกจากการประมาณค่าพารามิเตอร์ ก็คือ การทดสอบสมมติฐานเกี่ยวกับพารามิเตอร์ของประชากร ซึ่งเป็นวิธีการเปรียบเทียบความเชื่อ ที่นักวิจัยมีต่อค่าพารามิเตอร์ของประชากร กับข้อมูลที่เก็บรวบรวมมาได้จากการสุ่มเลือกตัวอย่าง โดยผ่านขั้นตอนการวิเคราะห์เชิงสถิติ แล้วหาข้อสรุปจากการเปรียบเทียบ

ข้อความเกี่ยวกับการทดสอบสมมติฐานจะถูกต้องหรือไม่นั้น ขึ้นอยู่กับข้อมูลที่ใช้ในการตัดสินใจเป็นสำคัญ กล่าวคือ เราบอกได้แต่เพียงว่า จากข้อมูลที่มีอยู่ ข้อสรุปเกี่ยวกับพารามิเตอร์ของประชากร ควรเป็นเช่นนี้เท่านั้น

กระบวนการของการทดสอบสมมติฐานแบ่งออกเป็น 4 ขั้นตอนใหญ่ ๆ คือ

1. การกำหนดข้อความที่เป็นสมมติฐาน

ประกอบไปด้วย สมมติฐานว่าง (null hypothesis) และ สมมติฐานแย้ง (alternative hypothesis)

2. การกำหนดระดับนัยสำคัญของการทดสอบสมมติฐานทางสถิติ (statistical significance level)

ค่าดังกล่าว เป็นค่าที่มากที่สุดของ ค่าความน่าจะเป็นที่ในการทดสอบสมมติฐาน จะเกิดความผิดพลาดชนิดที่ 1 (type I error probability) ซึ่งเป็นโอกาสที่ผู้ทำการตัดสินใจคนหนึ่ง ๆ จะทำการตัดสินใจผิดพลาด โดยการปฏิเสธสมมติฐานว่างที่ตั้งขึ้น เมื่อในความเป็นจริงสมมติฐานว่างนั้นถูกต้อง

3. กำหนดตัวสถิติที่ใช้ในการทดสอบ (test statistics) และ กฎการตัดสินใจ (decision rules) ของการทดสอบ

ในการกำหนดตัวสถิติที่ใช้ในการทดสอบ มีความจำเป็นที่จะต้องทราบการแจกแจงของตัวสถิตินั้น เพื่อที่จะได้สามารถคำนวณค่าสถิติ จากฟังก์ชันการแจกแจงของตัวสถิติ

ณ. ระดับนัยสำคัญที่กำหนด แล้วจึงใช้ค่าที่คำนวณได้ เป็นค่าวิกฤต* (critical value) ซึ่งจะใช้ในการกำหนด กฎการตัดสินใจของการทดสอบสมมติฐาน

ค่าวิกฤตดังกล่าว สามารถมีค่าเป็นค่าเดียวหรือสองค่าก็ได้ ขึ้นอยู่กับลักษณะของการทดสอบว่า เป็นการทดสอบแบบข้างเดียว (one-sided test) หรือ เป็นการทดสอบแบบสองข้าง (two-sided test)

กฎของการตัดสินใจ คือ การกำหนดบริเวณวิกฤต (critical region) ซึ่งเป็นบริเวณที่เราจะตัดสินใจปฏิเสธสมมติฐานว่าง (rejected: H_o) ถ้าค่าสถิติที่คำนวณได้จากตัวอย่างตกอยู่ในบริเวณนี้ ในกรณีที่ค่าสถิติที่คำนวณได้จากตัวอย่าง ไม่ได้มีค่าอยู่ในบริเวณวิกฤต เราจะตัดสินใจไม่ปฏิเสธสมมติฐานว่าง (do not reject: H_o) ในความหมายว่า ข้อมูลที่รวบรวมมาได้ ไม่ได้ปรากฏหลักฐานที่ชวนให้ให้ปฏิเสธสมมติฐานว่าง

4. การเก็บรวบรวมข้อมูล

ขั้นตอนการเก็บรวบรวมข้อมูลนับเป็นขั้นตอนที่สำคัญที่สุด ข้อมูลที่นำมาใช้ทดสอบสมมติฐานจะมีค่าถูกต้องเชื่อถือได้มากน้อยเพียงใด ขึ้นอยู่กับว่าตัวอย่างที่เลือกมานั้นมีความเป็นตัวแทนที่ดีของประชากรที่ต้องการทดสอบสมมติฐานหรือไม่ ภายหลังจากที่ได้เก็บรวบรวมข้อมูลเสร็จสิ้นแล้ว ก็จะสามารถคำนวณค่าสถิติได้ โดยการแทนค่าสังเกตที่ได้จากข้อมูล ซึ่งเป็นค่าของตัวแปรสุ่ม ลงในสูตรของตัวสถิติ แล้วทำการตัดสินใจเกี่ยวกับข้อความในสมมติฐานว่าง ตามกฎการตัดสินใจที่กำหนดไว้ในข้อ 3

จากขั้นตอนของกระบวนการทดสอบสมมติฐาน จะพบว่าในการหาการแจกแจงของตัวสถิติ เราจะต้องตั้งข้อกำหนดเกี่ยวกับรูปแบบการแจกแจงของข้อมูลเสียก่อน หลังจากนั้น จึงจะสามารถหาการแจกแจงของตัวสถิติได้ โดยกระทำภายใต้ข้อกำหนดที่ตั้งขึ้น

---

*สำหรับตัวสถิติที่ใช้กันบ่อย ๆ มักจะมีผู้สร้างตารางการแจกแจงของตัวสถิตินั้นเอาไว้แล้ว จึงสามารถหาค่าวิกฤตได้จากตาราง โดยไม่ต้องเสียเวลาคำนวณเอง

ภายหลังจากรวบรวมค่าสังเกต $x_1, x_2, \ldots, x_n$ มาได้แล้ว เราจะต้องทำการตรวจสอบว่า ข้อมูลที่ได้มามีความสอดคล้องกับข้อตกลงเบื้องต้นของการแจกแจงที่กำหนดหรือไม่ โดยการตั้งสมมติฐานว่างว่าชุดตัวอย่างสุ่ม $x_1, x_2, \ldots, x_n$ มีรูปแบบการแจกแจงเป็นไปตามข้อกำหนดที่ตั้งขึ้น (ต้องระบุค่าพารามิเตอร์ของการแจกแจงด้วย)

ปัญหาการทดสอบสมมติฐานในลักษณะดังกล่าวเรียกว่า การทดสอบภาวะสารูปดี (goodness-of-fits test) ซึ่งเป็นการทดสอบหรือตรวจสอบว่า ลักษณะการแจกแจงที่กำหนดไว้ในสมมติฐานว่าง จะสามารถเข้ากันได้ (fits) กับการแจกแจงของข้อมูล $x_1, x_2, \ldots, x_n$ ซึ่งเป็นค่าสังเกต ที่เราสังเกตมาได้ดีเพียงใด

ในทางปฏิบัติ เราจะไม่พยายามทดลองไปเรื่อย ๆ เพื่อหาว่า ข้อมูลของเรามีความสอดคล้องกับการแจกแจงใดมากที่สุด กล่าวคือ เราจะไม่ทำการทดสอบข้อมูลกับการแจกแจงทุกการแจกแจง แต่จะเลือกทดสอบกับการแจกแจงที่เรามีเหตุผล หรือมีความเชื่อที่ได้มาจากประสบการณ์มากพอที่จะกล่าวว่า ค่าสังเกต $x_1, x_2, \ldots, x_n$ น่าจะถูกสุ่มเลือกมาจากประชากรที่มีการแจกแจงนั้น ๆ

ดังนั้นอาจกล่าวได้ว่า การทดสอบภาวะสารูปดี คือ การหาคำตอบของคำถามซึ่งมีใจความว่า "เป็นไปได้หรือไม่ ที่จะกล่าวว่าค่าสังเกต $x_1, x_2, \ldots, x_n$ ได้ถูกสุ่มเลือกมาจากประชากร ที่มีรูปแบบการแจกแจงตามที่ระบุไว้ในสมมติฐานว่าง".

ข้อได้เปรียบ ของการนำเอารูปแบบการแจกแจงทางทฤษฎี มาใช้สำหรับการทดสอบภาวะสารูปดี ที่เห็นได้ชัดเจน คือ ในกรณีที่สมมุติฐานว่างไม่ถูกปฏิเสธ เราจะสามารถดึงเอาคุณสมบัติต่าง ๆ ของการแจกแจงที่ระบุไว้ในสมมุติฐานว่าง มาใช้เป็นข้อสรุปเกี่ยวกับประชากร ทำให้ได้ข้อสรุปที่กว้างกว่าการใช้ข้อสรุปที่ได้รับจากข้อมูลตัวอย่าง

ถ้าไม่มีเหตุผลเพียงพอที่จะกล่าวว่า การแจกแจงของข้อมูล ซึ่งได้จากตัวอย่างมีความสอดคล้องหรือเข้ากันได้กับการแจกแจงในทางทฤษฎี เราจะปฏิเสธสมมุติฐานว่างในกรณีเช่นนี้ เราจะสามารถหาข้อสรุปที่ดีที่สุดเกี่ยวกับประชากร ได้จากข้อมูลตัวอย่างเท่านั้น

อย่างไรก็ตาม ในความเป็นจริง สมมุติฐานว่าง (H_o) อาจจะไม่มีโอกาสเกิดขึ้นจริง ๆ เนื่องจาก เราอาจจะไม่สามารถหารูปแบบการแจกแจงที่แท้จริงของตัวแปรสุ่ม $x_1, x_2, \ldots, x_n$ ดังนั้น การล้มเหลวที่จะปฏิเสธสมมุติฐานว่าง (failure to reject: H_o) จึงไม่ได้หมายความว่า เรายอมรับว่าสมมุติฐานว่าง (H_o) เป็นจริง

หลักการของการทดสอบภาวะสารูปดี เป็นการเปรียบเทียบความเข้ากันได้ระหว่างการแจกแจงที่ระบุในสมมุติฐานว่าง กับการแจกแจงของข้อมูลตัวอย่าง เครื่องมือที่ใช้เป็นแนวทาง สำหรับบ่งชี้ความเข้ากันได้ คือการเปรียบเทียบกันระหว่างโค้งของฟังก์ชันความน่าจะเป็น กับฮิสโตแกรมของข้อมูลในกรณีข้อมูลเป็นค่าต่อเนื่อง หรือเปรียบเทียบกันระหว่างกราฟเส้นของฟังก์ชันความน่าจะเป็นแบบไม่ต่อเนื่อง กับกราฟเส้นของข้อมูลในกรณีที่ข้อมูลเป็นค่าไม่ต่อเนื่อง

ถ้าผลของการเปรียบเทียบพบว่า มีความใกล้เคียงกันระหว่างเส้นโค้งหรือกราฟเส้น จะถือว่า การแจกแจงที่ระบุในสมมุติฐานว่างเข้ากันได้กับการแจกแจงของข้อมูลตัวอย่าง

จากหลักการดังกล่าว ทำให้การทดสอบภาวะสารูปดี ถูกกล่าวถึงในฐานะที่เป็นวิธีที่มีระบบแบบแผน ซึ่งใช้ในการคัดเลือก หรือจำแนกความเข้ากันได้ระหว่างการแจกแจงของข้อมูลตัวอย่าง กับการแจกแจงทางทฤษฎีเท่านั้น

การทดสอบไคสแควร์ เป็นการทดสอบภาวะสารูปดีที่เก่าแก่ที่สุด การทดสอบนี้ถูกเสนอขึ้นในปี ค.ศ. 1900 โดย คาร์ล เพียร์สัน (Karl Pearson: ค.ศ. 1857-1936) ในผลงานชื่อ " On a Criterien That a Given System of Deviations from the Probable in the Case of a Correlated System of Variables Is Such That It Can Be Reasonably Supposed to Have Arisen in Random Sampling."

วิธีการทดสอบภาวะสารูปดีโดยใช้ไคสแควร์ (The Chi-Square goodness-of-fits test) สามารถใช้ได้ดีกับการแจกแจงแบบต่อเนื่อง และการแจกแจงแบบไม่ต่อเนื่อง โดยมีหลักการ ดังนี้

1. ทำการแบ่งช่วงของค่าที่เป็นไปได้ทั้งหมดของตัวแปรสุ่ม ที่มีการแจกแจงตามที่ระบุในสมมติฐานว่าง ออกเป็น k ช่วง โดยที่แต่ละช่วงจะต้องไม่มีสมาชิกซ้ำกัน (อินเตอร์เซกชันของทุกช่วงจะต้องเป็นเซตว่าง)

สมมติว่าได้เป็น $(a_0, a_1), [a_1, a_2), \ldots, [a_{k-1}, a_k)$ โดยที่ a_0 สามารถมีค่าเป็น $-\infty$ และ a_k สามารถมีค่าเป็น $+\infty$ ได้

กำหนดให้

$$N_j = \text{จำนวนค่าสังเกต } x \text{ ที่ตกอยู่ในช่วงที่ } j : [a_{j-1}, a_j)$$

สำหรับทุกค่า $j = 1,2,\ldots,k$ และจะต้องได้ว่า $N_1 + N_2 + \ldots + N_k$ เท่ากับ n เมื่อ n คือจำนวนทั้งหมดของค่าสังเกต x

2. คำนวณค่าสัดส่วนที่คาดหมายของจำนวนค่าสังเกต x ที่ตกอยู่ในช่วงที่ j ในการหาค่าสัดส่วนที่คาดหมาย จะต้องกระทำภายใต้เงื่อนไขว่า ค่าสังเกต $x_1, x_2, \ldots, x_n$ ได้ถูกสุ่มเลือกมาจากประชากร ที่มีรูปแบบการแจกแจงเป็นไปตามที่ระบุในสมมติฐานว่าง

กำหนดให้

p_j เป็นค่าสัดส่วนที่คาดหมายของจำนวนค่าสังเกต x ที่ตกอยู่ในช่วงที่ j

np_j เป็นจำนวนที่คาดหมายของค่าสังเกต x ที่จะตกอยู่ในช่วงที่ j

3. คำนวณค่าสถิติ Q_{k-m-1} จากสูตร

$$Q_{k-m-1} = \sum_{j=1}^{k} \frac{(N_j - np_j)^2}{np_j}$$

ตัวสถิติ Q_{k-m-1} มีความหมายดังนี้

k คือ จำนวนช่วงที่แบ่ง

m คือ จำนวนค่าของพารามิเตอร์ที่จะต้องประมาณค่า โดยอาศัยข้อมูลจากตัวอย่าง เพื่อใช้ในการกำหนดสมมติฐานว่าง*

3.1 ในกรณีทราบค่าพารามิเตอร์ทุกตัว** (m = 0) จะได้ว่า

ภายใต้เงื่อนไขสมมติฐานว่างเป็นจริง ตัวสถิติ Q_{k-1} จะมีการแจกแจงลู่เข้าสู่การแจกแจงไคสแควร์ ที่มีค่าองศาอิสระ k - 1 เมื่อขนาดตัวอย่าง n เข้าใกล้อนันต์

---

*เนื่องจากในการตั้งสมมติฐานว่าง จำเป็นที่จะต้องระบุค่าพารามิเตอร์ของการแจกแจง ในกรณีที่ไม่มีข้อมูลอื่น ๆ มาประกอบ เช่น ข้อมูลจากประสบการณ์ หรือผลงานค้นคว้าของนักวิจัยรุ่นก่อน ๆ จึงมีความจำเป็น ที่จะต้องใช้ค่าประมาณแบบจุดของค่าพารามิเตอร์ที่คำนวณได้จากข้อมูลตัวอย่าง แทนเป็นค่าพารามิเตอร์ในสมมติฐานว่าง

**ตัวอย่างที่ประยุกต์ใช้ในกรณีนี้ ส่วนใหญ่แล้ว จะเป็นเรื่องทางด้านการจำลอง (Simulations) เช่น การทดสอบตัวผลิตเลขสุ่ม (Testing of random number generator) การทดสอบกระบวนการปัวซอง (Poisson Process test) และการทดสอบตัวผลิตค่าตัวแปรสุ่ม (Testing of random variable generator) เป็นต้น

ในการทดสอบสมมติฐานว่าง (H_o) ที่ระดับนัยสำคัญ α มีกฎการตัดสินใจดังนี้

1. ตัดสินใจปฏิเสธสมมติฐานว่าง (H_o) เมื่อค่าสถิติ Q_{k-1} ที่คำนวณได้จากตัวอย่าง มีค่ามากกว่าค่าวิกฤต ซึ่งเป็นค่าของตัวแปรสุ่มไคสแควร์ที่มีองศาอิสระเท่ากับ k - 1 คำนวณที่ระดับนัยสำคัญ α

2. ตัดสินใจไม่ปฏิเสธสมมติฐานว่าง (H_o) เมื่อค่าสถิติ Q_{k-1} มีค่าน้อยกว่าหรือเท่ากับ ค่าวิกฤตในข้อ 1

3.2 ในกรณีที่การกำหนดสมมติฐานว่าง มีความจำเป็นต้องใช้ค่าประมาณแบบจุด ที่คำนวณได้จากข้อมูลตัวอย่าง แทนเป็นค่าของพารามิเตอร์ในสมมติฐาน จำนวนทั้งสิ้น m ตัว (m > 1)

เชอร์นอฟ และ ลีมานท์[5] ได้พิสูจน์ว่า ภายใต้เงื่อนไข 3 ประการ คือ

1. สมมติฐานว่าง (H_o) เป็นจริง
2. ตัวประมาณแบบจุดทั้ง m ตัว เป็นตัวประมาณภาวะน่าจะเป็นสูงสุด (maximum likelihood estimator)
3. ค่าขนาดตัวอย่าง n มีค่าเข้าใกล้อนันต์

จะได้ว่า การแจกแจงของตัวสถิติ Q_{k-m-1} จะลู่เข้าสู่ฟังก์ชันการแจกแจงชนิดหนึ่ง ซึ่งถูกปิดล้อมทางด้านล่าง โดยฟังก์ชันการแจกแจงแบบไคสแควร์ ที่มีองศาอิสระ k - 1 และถูกปิดล้อมทางด้านบน โดยฟังก์ชันการแจกแจงแบบไคสแควร์ ที่มีองศาอิสระ k - m - 1

---

[5]ศึกษารายละเอียดได้ใน H. Chernoff and E.L. Lehmann, The use of maximum likelihood Estimates in X^2 tests for Goodness of Fit (Ann. Math. Stat. Vol. 25, 1954), pp. 579-586.

ในการทดสอบสมมติฐานว่าง (H_o) ที่ระดับนัยสำคัญ α มีกฎการตัดสินใจดังนี้

1. ตัดสินใจปฏิเสธสมมติฐานว่าง (H_o) เมื่อค่าสถิติ Q_{k-m-1} ที่คำนวณได้จากตัวอย่าง มีค่ามากกว่า ค่าของตัวแปรสุ่มไคสแควร์ ที่มีองศาอิสระเท่ากับ k - 1 คำนวณที่ระดับนัยสำคัญ α

2. ตัดสินใจไม่ปฏิเสธสมมติฐานว่าง (H_o) เมื่อค่าสถิติ Q_{k-m-1} มีค่าน้อยกว่า ค่าของตัวสถิติไคสแควร์ ที่มีองศาอิสระ k - m - 1 คำนวณที่ระดับนัยสำคัญ α

3. ทำการสุ่มตัวอย่างเพิ่มเติมให้เพียงพอ ที่จะตัดสินใจปฏิเสธหรือไม่ปฏิเสธสมมติฐานว่าง (H_o) ตามกฎเกณฑ์ในข้อ 1 และข้อ 2 ในกรณีที่ค่าสถิติ Q_{k-m-1} ที่คำนวณได้ ไม่สอดคล้องกับกฎการตัดสินใจในข้อ 1 และ ข้อ 2

2.6 ทฤษฎีการประมาณค่าอินทิกรัล

ในบางครั้ง การคำนวณค่าอินทิกรัล ไม่สามารถทำได้โดยอาศัยเทคนิคการอินทิเกรต เนื่องจากฟังก์ชันที่กำหนดให้ไม่อยู่ในรูปแบบที่จะใช้สูตรใด ๆ ได้ ในกรณีเช่นนี้จึงจำเป็นต้องคำนวณค่าอินทิกรัลโดยการประมาณ วิธีการประมาณง่าย ๆ ก็คือ การประมาณฟังก์ชันด้วยฟังก์ชันซึ่งมีกราฟเป็นเส้นตรงในแต่ละช่วงย่อย ๆ เรียกสูตรการประมาณค่าอินทิกรัลในลักษณะนี้ว่า กฎสี่เหลี่ยมคางหมู (Trapezoidal rule)

กฎสี่เหลี่ยมคางหมู เป็นกฎซึ่งใช้ในการประมาณค่าอินทิกรัล โดยการประมาณค่าฟังก์ชันที่กำหนดให้ด้วยฟังก์ชันซึ่งมีสูตรอยู่ในรูป

$$g(x) = a + bx$$

ฟังก์ชันดังกล่าวมีกราฟเป็นเส้นตรงในแต่ละช่วงย่อย $[x_{i-1}, x_i]$ ใด ๆ ซึ่งอยู่ภายในช่วงของการอินทิเกรต

โดยทั่วไปการประมาณค่าอินทิกรัล $\int_a^b f(x)dx$ เมื่อ f เป็นฟังก์ชันต่อเนื่องบนช่วง $[a, b]$ โดยวิธีกฎสี่เหลี่ยมคางหมู กระทำได้ดังนี้

เริ่มต้นด้วยการแบ่งช่วง $[a, b]$ ออกเป็น n ช่วงเท่า ๆ กัน โดยที่แต่ละช่วงย่อยยาวเท่ากับ $h = (b-a)/n$

ให้ $x_0, x_1, \ldots, x_n$ เป็นจุดแบ่งช่วงดังกล่าว โดยที่

$$a = x_0 < x_1 < \ldots < x_{n-1} < x_n = b$$

ดังนั้นจะได้ช่วงย่อย ๆ จำนวน n ช่วงที่มีความยาวช่วงเท่ากัน ดังนี้

$$[x_0, x_1], [x_1, x_2], [x_2, x_3], \ldots, [x_{n-1}, x_n]$$

และสามารถกำหนดจุด x_i ใด ๆ ได้จากความสัมพันธ์

$$x_i = a + ih \qquad : i = 0, 1, \ldots, n$$

ถ้าให้ $y_o, y_1, \ldots, y_n$ เป็นค่าของฟังก์ชัน f ที่จุด $x_o, x_1, \ldots, x_n$ ตามลำดับกล่าวคือ

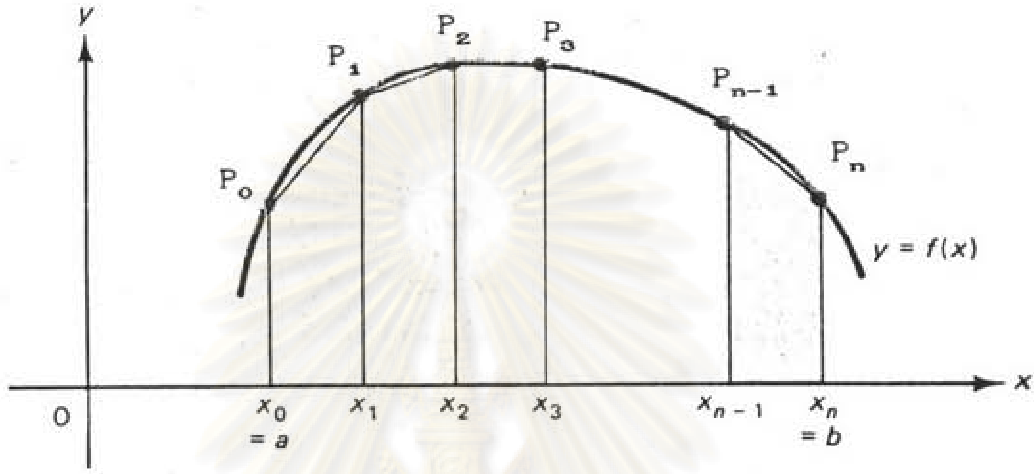
$$\begin{aligned} y_o &= f(x_o) \\ y_1 &= f(x_1) \\ &\ \ \vdots \\ y_{n-1} &= f(x_{n-1}) \\ y_n &= f(x_n) \end{aligned}$$

เราจะสามารถประมาณค่าอินทิกรัล ได้ด้วยสูตร

$$\int_a^b f(x)dx \doteq (h/2)(y_o + 2y_1 + \ldots + 2y_{n-1} + y_n)$$

เรียกสูตรข้างบนว่า กฎสี่เหลี่ยมคางหมู

ในกรณีที่ค่า y_i ต่างมีค่ามากกว่าศูนย์ทุกตัว การประมาณค่าอินทิกรัลด้วยสูตรกฎสี่เหลี่ยมคางหมู มีความหมายทางเรขาคณิต ดังรูปที่ 2.1



รูปที่ 2.1 แสดงความหมายทางเรขาคณิตของกฎสี่เหลี่ยมคางหมู

จากรูปจะเห็นได้ว่า ในการใช้กฎสี่เหลี่ยมคางหมูประมาณค่าอินทิกรัลจำกัดเขต $\int_a^b f(x)dx$ นั้น ทำได้โดยอาศัยการสร้างรูปสี่เหลี่ยมคางหมูหลาย ๆ รูป ภายใต้เส้นโค้งของฟังก์ชัน f ที่ต้องการหาพื้นที่

หลักการของวิธีกฎสี่เหลี่ยมคางหมู อธิบายได้ดังนี้

ในช่วง $[x_0, x_1]$ กราฟของเส้นตรง P_0P_1 มีสูตรเป็น

$$y = a + bx \qquad : \text{เมื่อ } a \text{ และ } b \text{ เป็นค่าคงที่}$$

เราจึงประมาณค่าอินทิกรัลของฟังก์ชัน f ในช่วงนี้จากความสัมพันธ์

$$\int_{x_0}^{x_1} f(x)dx \doteq \int_{x_0}^{x_1} g(x)dx = \int_{x_0}^{x_1} y\,dx$$

$$= \int_{x_0}^{x_1} (a + bx)dx$$

$$= a(x_1 - x_0) + (b/2)(x_1^2 - x_0^2)$$

$$= (1/2)(x_1 - x_0)(2a + bx_1 + bx_0)$$

$$= (1/2)(x_1 - x_0)(a + bx_0 + a + bx_1)$$

$$= (1/2)(x_1 - x_0)(y_0 + y_1)$$

$$= (1/2)(h)(y_0 + y_1)$$

ค่าของ $\int_{x_0}^{x_1} f(x)dx$ คือพื้นที่ระหว่างกราฟของ f กับแกน x ในช่วง $[x_0, x_1]$

ค่า $(1/2)(h)(y_0 + y_1)$ คือพื้นที่สี่เหลี่ยมคางหมู ซึ่งมีด้านคู่ขนานยาว y_0 กับ y_1 และระยะระหว่างด้านคู่ขนานยาว h

g คือฟังก์ชันที่มีกราฟเป็นเส้นตรงที่ใช้ประมาณฟังก์ชัน f

ความสัมพันธ์ดังกล่าว คือการประมาณพื้นที่ใต้โค้งของฟังก์ชัน f ในช่วง $[x_0, x_1]$ ด้วยพื้นที่สี่เหลี่ยมคางหมู

ในทำนองเดียวกัน สำหรับช่วง $[x_1, x_2]$, ..., $[x_{n-1}, x_n]$ จะได้ว่า

$$\int_{x_1}^{x_2} g(x)dx = (h/2)(y_1 + y_2)$$

$$\int_{x_2}^{x_3} g(x)dx = (h/2)(y_2 + y_3)$$

$$\vdots$$

$$\int_{x_{n-1}}^{x_n} g(x)dx = (h/2)(y_{n-1} + y_n)$$

ดังนั้นจึงได้

$$\int_a^b f(x)dx = \int_a^b g(x)dx$$

$$= \int_{x_0}^{x_1} g(x)dx + \int_{x_1}^{x_2} g(x)dx + \dots + \int_{x_{n-1}}^{x_n} g(x)dx$$

$$= (h/2)(y_0 + y_1) + (h/2)(y_1 + y_2) + \dots + (h/2)(y_{n-1} + y_n)$$

$$= (h/2)(y_0 + 2y_1 + 2y_2 + \dots + 2y_{n-1} + y_n)$$

2.7 ผลงานที่เกี่ยวข้องกับการวิจัย

เนื่องจากการวิจัยครั้งนี้ มีเป้าหมายที่จะสร้างตารางสำเร็จรูปแสดงช่วงความเชื่อมั่นสำหรับพารามิเตอร์ p ของการแจกแจงทวินาม เมื่อขนาดตัวอย่าง n มีค่าตั้งแต่ 1 ถึง 30 และพิจารณาความเหมาะสมของของ วิธีการประมาณโดยใช้การแจกแจงปกติ ในการคำนวณช่วงความเชื่อมั่นสำหรับพารามิเตอร์ p ภายใต้เงื่อนไขตัวอย่างขนาดเล็ก (n = 1 ถึง 30)

รายละเอียดที่จะนำเสนอต่อไป เป็นผลงานของนักสถิติในรุ่นก่อน ที่ได้ถ่ายทอดมาสู่ปัจจุบัน โดยจะเป็นผลงานในส่วนที่เกี่ยวกับ วิธีการคำนวณช่วงความเชื่อมั่น 2 วิธี คือ สูตรการประมาณโดยใช้การแจกแจงปกติ และ สูตรที่แท้จริงซึ่งสัมพันธ์กับการแจกแจงเอฟ

2.7.1 สูตรการประมาณโดยใช้การแจกแจงปกติ

ในการตีความเกี่ยวกับค่า p ซึ่งเป็นค่าความน่าจะเป็นของการเกิดความสำเร็จในการทดลองเบอร์นูลี 1 ครั้ง ตามนัยของความน่าจะเป็นในรูปความถี่ ได้ความว่าค่า p ดังกล่าว คือ ค่าความถี่สัมพัทธ์ของจำนวนครั้งที่เกิดความสำเร็จในการทดลองเบอร์นูลี ที่กระทำซ้ำ ๆ กัน เป็นจำนวน n ครั้ง โดยการทดลองแต่ละครั้งเป็นอิสระจากกัน และค่า n มีค่าใกล้อนันต์

ดังนั้นในการหาค่าที่แท้จริงของค่า p ดังกล่าว เราจะต้องทำการทดลองเบอร์นูลี ซ้ำ ๆ กัน โดยการทดลองในแต่ละครั้งจะต้องเป็นอิสระจากกัน ไปเรื่อย ๆ จน จำนวนครั้งของการทดลองมีค่าเข้าใกล้อนันต์ แล้วจึงคำนวณค่า p จากค่าความถี่สัมพัทธ์ของจำนวนครั้งของการทดลองที่เกิดความสำเร็จ ซึ่งจะเห็นได้ว่าวิธีการหาค่า p ตามนัยของความน่าจะเป็นในรูปความถี่ ไม่อาจทำได้ในทางปฏิบัติ เนื่องจากเราไม่ทราบจำนวนครั้งที่แน่นอนของการทดลอง จึงได้มีแนวคิดที่จะนำเอาทฤษฎีทางคณิตศาสตร์ของความน่าจะเป็น (Mathematical Theory of Probability) มาประยุกต์ใช้

กับปัญหาดังกล่าว ผลลัพธ์ที่ได้เป็นที่รู้จักกันในนาม กฎเลขจำนวนมากของเบอร์นูลี* (Bernoulli law of large number) ใจความของกฎเลขจำนวนมากของเบอร์นูลีมีดังนี้

<u>กฎเลขจำนวนมากของเบอร์นูลี</u>

สำหรับการทดลองเบอร์นูลี ที่มีค่าความน่าจะเป็นของการเกิดความสำเร็จเท่ากับ p กำหนดให้

S_n เป็นค่าสังเกต ซึ่งเป็นจำนวนครั้งที่เกิดความสำเร็จในการทดลองเบอร์นูลี n ครั้ง ที่เป็นอิสระจากกัน

f_n เป็นค่าความถี่สัมพัทธ์ของจำนวนครั้งที่เกิดความสำเร็จในการทดลองเบอร์นูลี n ครั้ง และหาได้จากความสัมพันธ์

$$f_n = \frac{S_n}{n}$$

สำหรับจำนวนเต็มบวก $\epsilon > 0$ ใด ๆ ที่กำหนดให้ จะได้ว่า

$$\lim_{n \to \infty} Pr(\ |f_n - p| \leqslant \epsilon\ ) = 1$$

$$\lim_{n \to \infty} Pr(\ |f_n - p| \geqslant \epsilon\ ) = 0$$

---

*กฎเลขจำนวนมากของเบอร์นูลี ถูกค้นพบในปี ค.ศ. ใด ไม่ปรากฎหลักฐานที่แน่ชัด เนื่องจากผลงานดังกล่าวได้ถูกตีพิมพ์ลงในหนังสือ " Ars Conjectandi " (แปลว่าศิลปะของการคาดคะเน) ซึ่งแต่งโดย จาคอบ เบอร์นูลี (Jacob Bernulli: ค.ศ. 1654-1705) ในปี ค.ศ. 1713 หลังจากที่ท่านได้เสียชีวิตไปแล้วถึง 8 ปี

กฎเลขจำนวนมากของเบอร์นูลี กล่าวว่า ในการประมาณค่า p ซึ่งเป็นค่าความน่าจะเป็นของการเกิดความสำเร็จในการทดลองเบอร์นูลี จะต้องทำการทดลองแล้วสังเกตจำนวนครั้งที่เกิดความสำเร็จในการทดลอง ค่าประมาณที่ได้จะมีความถูกต้องเท่ากับที่แท้จริงของ p เมื่อจำนวนครั้งของการทดลองเข้าใกล้ค่าอนันต์

ในปี ค.ศ. 1733 อะบราฮัม เดอร์มัวพ์ (Abraham De Moivre: ค.ศ. 1667-1754) ได้เสนอผลงานชื่อ " Approximato and Summam terminorum Binomii $\overline{a + b}^{n}$ Seriem expansi." ผลงานดังกล่าวเป็นการประมาณค่าฟังก์ชันการแจกแจงแบบทวินามที่มีพารามิเตอร์ (n, p= 1/2) ด้วยฟังก์ชันการแจกแจงของค่าผิดพลาดจากการวัด (Distribution of Error) ซึ่งต่อมาเป็นที่รู้จักกันในชื่อการแจกแจงปกติ (Normal distribution)

ต่อมาในปี ค.ศ. 1812 ปิแยร์ ซิมง เดอ ลาปลาซ (Pierre Simon de Laplace: ค.ศ. 1749-1827) ได้เสนอผลงานชื่อ " Theorie Analytique des Probabilities." ส่วนหนึ่งของผลงานดังกล่าว เป็นการขยายความคิดของเดอร์มัวพ์ในการประมาณค่าฟังก์ชันการแจกแจงแบบทวินาม ที่มีพารามิเตอร์ (n,p) เมื่อ p ไม่เท่ากับ 1/2

ผลงานของท่านทั้งสองต่อมาเป็นที่รู้จักในชื่อ " Demoivre - Laplace Limit Theorem " ซึ่งมีใจความโดยย่อดังต่อไปนี้

<u>ทฤษฎีลิมิตของเดอร์มัวพ์-ลาปลาซ</u>

สำหรับการทดลองเบอร์นูลี ที่มีความน่าจะเป็นของการเกิดความสำเร็จเท่ากับ p กำหนดให้ S_n เป็นค่าสังเกต ซึ่งเป็นจำนวนครั้งที่เกิดความสำเร็จในการทดลองเบอร์นูลี n ครั้ง ที่เป็นอิสระจากกัน

จะได้ว่าเมื่อค่า n เข้าใกล้อนันต์ ตัวแปรสุ่ม S_n จะมีการแจกแจงลู่เข้าสู่การแจกแจงปกติ ที่มีค่าเฉลี่ยเท่ากับ np และ ค่าความแปรปรวนเท่ากับ np(1 - p)

จากทฤษฎีลิมิตของ เดอร์มัวฟ์-ลาปลาซ ได้นำไปสู่ สูตรการประมาณการแจกแจงทวินามโดยใช้การแจกแจงปกติ

$$\sum_{k=a}^{b}\binom{n}{k}p^k q^{n-k} \doteq \frac{1}{\sqrt{2\pi}}\int_{\frac{a-np-\frac{1}{2}}{\sqrt{npq}}}^{\frac{b-np+\frac{1}{2}}{\sqrt{npq}}} e^{-\frac{1}{2}y^2}\,dy$$

$$= \Phi\left(\frac{b-np+\frac{1}{2}}{\sqrt{npq}}\right) - \Phi\left(\frac{a-np-\frac{1}{2}}{\sqrt{npq}}\right)$$

จากสูตรการประมาณการแจกแจงทวินามโดยใช้การแจกแจงปกติ เราจะได้สูตรการประมาณค่าแบบช่วงสำหรับค่าพารามิเตอร์ p ของการแจกแจงทวินาม เป็นช่วง (PL, PU) ที่กำหนดโดย

$$PL = \frac{X}{n} - Z_{1-\alpha/2}\sqrt{\frac{X}{n}\left(1 - \frac{X}{n}\right)/\, n}$$

$$PU = \frac{X}{n} + Z_{1-\alpha/2}\sqrt{\frac{X}{n}\left(1 - \frac{X}{n}\right)/\, n}$$

เมื่อ x คือ ค่าของตัวแปรสุ่มทวินามที่สังเกตได้

n คือ ค่าขนาดตัวอย่าง

$Z_{1-\alpha/2}$ คือ ค่าของตัวแปรสุ่มปกติมาตรฐาน ที่ให้ค่าของฟังก์ชันความน่าจะเป็นสะสมเท่ากับ $1 - \alpha/2$

และจะได้ว่าช่วง (PL, PU) ดังกล่าว เป็นช่วงความเชื่อมั่น $(1 - \alpha)100$ เปอร์เซนต์ สำหรับพารามิเตอร์ p

2.7.2 สูตรที่แท้จริงซึ่งสัมพันธ์กับการแจกแจงเอฟ

ในปี ค.ศ. 1924 คาร์ล เพียร์สัน [6] ได้พิสูจน์ความสัมพันธ์ระหว่าง ผลบวก p เทอมแรกของการกระจายทวินาม $(a + b)^n$ กับฟังก์ชันเบต้าที่ไม่สมบูรณ์ (Incomplete Beta function) ซึ่งสรุปได้ดังนี้

สำหรับการกระจายทวินาม $(a + b)^n$ สามารถกระจายออกเป็น

$$(a + b)^n = C_{0,n}a^0b^n + C_{1,n-1}a^1b^{n-1} + \ldots + C_{n,0}a^nb^0$$

ดังนั้นผลบวก p เทอมแรกของการกระจาย คือ

$$C_{0,n}a^0b^n + C_{1,n-1}a^1b^{n-1} + \ldots + C_{p-1,n-p+1}a^{p-1}b^{n-p+1}$$

เพียร์สันได้แสดงว่า ผลบวก p เทอมแรกดังกล่าว มีค่าเท่ากับ

$$(a + b)^n\left[1 - \frac{B_x(p, q)}{B(p, q)}\right]$$

เมื่อ $x = b/(a + b)$

$q = n - p + 1$

$B(p, q)$ คือ ฟังก์ชันเบต้า (Beta function) ซึ่งกำหนดโดย

$$B(p, q) = \int_0^1 t^{p-1}(1 - t)^{q-1}dt$$

---

[6] Karl Pearson, Note on the Relationship of the Incomplete B-function to the Sum of the first p terms of the Binomial $(a + b)^n$ (Biometika, Vol. 16, 1924), pp. 202-203.

$B_x(p, q)$ คือฟังก์ชันเบต้าที่ไม่สมบูรณ์ (Incomplete Beta function) กำหนดโดย

$$B_x(p, q) = \int_0^x t^{p-1}(1 - t)^{q-1}dt \quad : 0 < x < 1$$

ต่อมาในปี ค.ศ. 1942 ทอมสัน [7] ได้สร้างตารางการแจกแจงเบต้าขึ้น โดยในตอนนั้นเรียกว่า อัตราส่วนของฟังก์ชันเบต้าที่ไม่สมบูรณ์ (Incomplete Beta function ratio) แทนด้วย

$$I_x(p, q) = \frac{B_x(p, q)}{B(p, q)} \quad : 0 < x < 1$$

$$= \frac{1}{B(p, q)} \int_0^x t^{p-1}(1 - t)^{q-1}dt$$

ซึ่งจะเห็นได้ว่า เทอมทางขวามือของสมการดังกล่าว คือ ฟังก์ชันการแจกแจงแบบเบต้าที่รู้จักกันในปัจจุบันนั่นเอง

ในปี ค.ศ. 1952 ฮาล [8] ได้พิสูจน์ความสัมพันธ์ระหว่าง ฟังก์ชันการแจกแจงแบบทวินาม กับ ฟังก์ชันการแจกแจงแบบเบต้า* ดังนี้

---

[7] Catherine M. Thompson, Tables of Percentage Points of the Incomplete BETA-function (Biometrika, Vol. 32, 1941-1942) pp. 151-181.

[8] A. Hald, Statistical Theory with Engineering Applications (New York: John Wiley and sons, 1952), pp. 674.

*ขณะนั้นยังคงเรียกว่า Incomplete Beta function ratio

สำหรับตัวแปรสุ่ม X ซึ่งมีการแจกแจงทวินาม ที่มีพารามิเตอร์ (n,p) จะได้ว่า

$$Pr(X \geq x) = I_p(x, n - x + 1)$$

$$Pr(X \leq x) = I_{1-p}(n - x, x + 1)$$

โดยที่

$$I_{1-p}(n - x, x + 1) = 1 - I_p(x + 1, n - x)$$

และ ฮาล [๑] ยังได้เสนอความสัมพันธ์ระหว่าง ฟังก์ชันการแจกแจงแบบทวินาม กับ ฟังก์ชันการแจกแจงแบบเอฟ ดังนี้

สำหรับตัวแปรสุ่ม X ซึ่งมีการแจกแจงทวินาม ที่มีพารามิเตอร์ (n,p) จะได้ว่า

$$Pr(X \leq x) = 1 - Pr\left(F_1 < \frac{n - x}{x + 1} \frac{p}{1 - p}\right)$$

เมื่อ F_1 เป็นตัวแปรสุ่มเอฟ ที่มีค่าองศาอิสระ (2[x + 1], 2[n - x])
และจะได้ว่า

$$Pr(X \geq x) = Pr\left(F_2 < \frac{n - x + 1}{x} \frac{p}{1 - p}\right)$$

เมื่อ F_2 เป็นตัวแปรสุ่มเอฟ ที่มีค่าองศาอิสระ (2x, 2[n - x + 1])

---

[๑] เรื่องเดียวกัน, หน้า 675 และ 696.

จากความสัมพันธ์ดังกล่าว จะสามารถสร้างช่วงความเชื่อมั่น $(1 - \alpha)100$ เปอร์เซนต์ สำหรับค่าพารามิเตอร์ p ของการแจกแจงทวินาม เป็นช่วง (PL, PU) ซึ่งสอดคล้อง

$$Pr(PL \leqslant p \leqslant PU) = 1 - \alpha$$

ฮาล[10] ได้แสดงว่า สามารถหาค่า PL และ PU ได้จากสมการ

$$\sum_{k=x}^{n} C_{n,k}(PL)^k(1 - PL)^{n-k} = \alpha/2$$

$$\sum_{k=0}^{x} C_{n,k}(PU)^k(1 - PU)^{n-k} = \alpha/2$$

ซึ่งได้คำตอบเป็น

$$PL = \frac{x}{x + (n - x + 1)F_2}$$

$$PU = \frac{(x + 1)F_1}{n - x + (x + 1)F_1}$$

[10] เรื่องเดียวกัน, หน้า 697-698.