

# ไขปัญหาเกี่ยวกับสถิติ วัตถุประสงค์ และวิจัย ตอนที่ 1

รศ.ดร. บุญชม ศรีสะอาด

มีข้อสงสัยเกี่ยวกับการใช้สถิติในการวิเคราะห์ข้อมูล การวัตถุประสงค์และวิจัย หลาย ๆ เรื่องที่น่าจะทำความเข้าใจให้ชัดเจนในประเด็นปัญหาเหล่านั้น จึงได้เปิดคอลัมน์ "ไขปัญหาเกี่ยวกับสถิติ วัตถุประสงค์ และวิจัย" ขึ้น สำหรับครั้งนี้จะไขปัญหาเกี่ยวกับจำนวนกลุ่มตัวอย่าง การสรุปผลการทดสอบสมมติฐาน และเครื่องหมายของค่า  $t$  โดยในแต่ละเรื่องจะยกสถานการณ์ ประเด็นปัญหา และคำตอบ ตามลำดับ

## 1. เกี่ยวกับจำนวนกลุ่มตัวอย่าง

**สถานการณ์** ผู้วิจัยกำหนดจำนวนกลุ่มตัวอย่าง โดยใช้หลักเกณฑ์ใดหลักเกณฑ์หนึ่ง (เช่น ใช้ตารางของเครทซี และมอร์แกน) พบว่า จะต้องใช้กลุ่มตัวอย่าง 250 คน แต่จากการส่งแบบสอบถามให้กลุ่มตัวอย่าง 250 คน ตอบ เมื่อครบกำหนด ได้รับแบบสอบถามคืนมาเพียง 100 ชุด หลังจากทำการติดตามได้รับคืนเพิ่มอีก 120 ชุด รวมแบบสอบถามที่ได้รับคืนมาทั้งหมด 220 ชุด ผู้วิจัยทำการวิเคราะห์ข้อมูลจากแบบสอบถาม 220 ชุด ที่ได้รับคืนมานั้น แล้วแปลผลและสรุปผลการวิจัย

**ประเด็นปัญหา** จำนวนกลุ่มตัวอย่าง ในการวิจัยครั้งนี้มี 250 คน หรือ 220 คน

**คำตอบ** จำนวนกลุ่มตัวอย่างในการวิจัยครั้งนี้มี 220 คน ทั้งนี้เพราะผู้วิจัยวิเคราะห์ข้อมูลจาก 220 คน ไม่ได้วิเคราะห์จาก 250 คน ผู้วิจัยตั้งใจจะใช้ 250 คน ตามตารางของเครทซี และมอร์แกน แต่ไม่สามารถเก็บรวบรวมข้อมูลให้ครบได้ กลุ่มตัวอย่างจึงมีเพียง 220 คน ซึ่งเป็นร้อยละ 88 ของจำนวนที่กำหนดไว้ ผู้วิจัยเห็นว่ามีความเหมาะสมพอ อนุโลมตามควรเก็บรวบรวมข้อมูลให้ได้ครบ 250 คน ซึ่งจะถูกต้องดีกว่า

ในการเขียนรายงานการวิจัยทั้งในบทคัดย่อและในส่วนอื่น ๆ ต้องระบุให้สอดคล้องตรงกันว่า กลุ่มตัวอย่างในการวิจัยครั้งนี้มี 220 คน ดังนั้นในบทที่ 1 ที่ว่าด้วยขอบเขตของการวิจัย เมื่อกล่าวถึงกลุ่มตัวอย่างในการวิจัยต้องระบุว่า มี 220 คน บทที่ 5 สรุปผล กล่าวถึงกลุ่มตัวอย่างว่ามี 220 คน และในบทที่ 3 วิเคราะห์ผลการ กล่าวถึงกลุ่มตัวอย่างว่ามี 220 คน และระบุถึงที่มาว่า ในการกำหนดกลุ่มตัวอย่างผู้วิจัยใช้ตารางของเครทซี และมอร์แกน พบว่า มีจำนวน 250 คน จากการส่งแบบสอบถาม 250 ชุด ให้กับกลุ่มตัวอย่างที่สุ่มได้ 250 คน ได้รับคืน 100 ชุด ผู้วิจัยติดตามได้อีก 120 ชุด รวมแบบสอบถามที่ได้รับคืนทั้งหมด 220 ชุด กลุ่มตัวอย่างในการวิจัยครั้งนี้จึงมี 220 คน ในส่วนที่กล่าวถึงการเก็บรวบรวมข้อมูลก็ระบุในรายละเอียดด้วย

อย่างไรก็ตามการเก็บรวบรวมข้อมูลให้ครบ 250 คน ตามที่กำหนดไว้จะถูกหลักการมากกว่า ถ้าสุดิวสัยที่จะเก็บรวบรวมข้อมูลกับผู้ที่สุ่มได้ ที่ยังไม่ตอบแบบสอบถาม (30 คน) ก็ควรสุ่มใหม่ในจำนวนที่เหลือ 30 คน นั้น แล้วส่งแบบสอบถามและเก็บรวบรวมข้อมูลจากคนที่สุ่มเพิ่มใหม่นั้น

## 2. เกี่ยวกับการสรุปผลการทดสอบสมมติฐาน

**สถานการณ์** ในการทดสอบสมมติฐานโดยใช้สถิติอ้างอิง (Inferential Statistics) เช่น t-test, F-test หรือ  $\chi^2$ -test เมื่อค่าสถิติทดสอบไม่มีนัยสำคัญทางสถิติ ซึ่งเกิดขึ้นในกรณีที่ค่าสถิติที่ได้จากการวิเคราะห์จากข้อมูลที่รวบรวมมา มีค่าน้อยกว่าค่าวิกฤต ค่า p มากกว่าระดับนัยสำคัญ ( $\alpha$ ) ที่กำหนดไว้ ผู้วิจัยเขียนแปลผลสรุปว่า "ไม่แตกต่างกัน" หรือแปลผลสรุปว่า "แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ"

**ประเด็นปัญหา** ควรแปลผลอย่างไรจึงจะเหมาะสมถูกต้องที่สุด (แปลผลสรุปว่า "ไม่แตกต่างกัน" หรือแปลผลสรุปว่า "แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ")

**คำตอบ** ก่อนที่จะตอบว่าควรแปลผลอย่างไรดี ขออธิบายเหตุผลที่แปลผลสรุปในลักษณะนั้นก่อน ดังนี้

ที่แปลผลสรุปว่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ มาจากฐานความคิดที่ว่า มีความแตกต่างกันจริง แต่ความแตกต่างนั้นน้อยกว่าระดับนัยสำคัญเกณฑ์ที่กำหนด กล่าวคือ มี probability ที่จะเกิด type I error มากกว่าระดับนัยสำคัญ ( $\alpha$ ) ที่กำหนดไว้

ตามหลักการแล้ว ถ้าค่าสถิติที่คำนวณจากข้อมูลที่รวบรวมมามีค่าเท่ากับหรือมากกว่าค่าวิกฤต แสดงว่าความแตกต่างนั้น มีนัยสำคัญทางสถิติ ผู้วิจัยก็จะสรุปว่า **แตกต่างกันอย่างมีนัยสำคัญทางสถิติ** ที่ระดับนัยสำคัญ ( $\alpha$ ) ที่กำหนดไว้ นั่น ถ้าค่าสถิติที่คำนวณมีค่าน้อยกว่าค่าวิกฤต แสดงว่า ความแตกต่างนั้นไม่มีนัยสำคัญทางสถิติที่ระดับนัยสำคัญดังกล่าว ดังนั้นเมื่อพบว่าค่าสถิติที่คำนวณน้อยกว่าค่าวิกฤต ผู้วิจัยจึงสรุปว่า **แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ** การสรุปดังกล่าวนี้ ผู้อ่านบางคนไม่เข้าใจ สงสัยว่าสรุปผลวิจัยได้เช่นไรแตกต่างกันหรือไม่แตกต่างกัน? ผู้วิจัยก็คงตอบว่าแตกต่างกันแต่ไม่มีนัยสำคัญทางสถิติ ซึ่งผู้อ่าน (คนเดิม) ก็ยังไม่คลายสงสัย

ที่แปลผลสรุปว่า **ไม่แตกต่างกัน** เนื่องจากเห็นว่า เป็นเรื่องของการทดสอบสมมติฐานสมมติฐานที่ทดสอบด้วยสถิติ คือ สมมติฐานแบบ Null Hypothesis เช่น  $H_0 : \mu_E = \mu_C$  (สัญลักษณ์  $\mu$  แทนค่าเฉลี่ยของประชากร E มาจาก Experiment หมายถึงกลุ่มทดลอง C มาจาก Control ซึ่งหมายถึงกลุ่มควบคุม) เมื่อคำนวณค่าสถิติทดสอบโดยใช้ข้อมูลที่รวบรวมมา เช่น คำนวณค่า t ได้ค่าที่คำนวณน้อยกว่าค่าวิกฤต (เช่น t คำนวณมีค่า 0.80 ค่า t วิกฤตจากการเปิดตาราง t ในตำราสถิติ เท่ากับ 1.96) แสดงว่า ค่า t ไม่มีนัยสำคัญทางสถิติ ตามหลักเมื่อไม่มี

นัยสำคัญทางสถิติ ผู้วิจัยก็จะยอมรับ (สรุปตาม)  $H_0$  คือ สรุปว่า  $\mu_E = \mu_C$  นั่นคือ สรุปว่า ค่าเฉลี่ยของประชากรในกลุ่มทดลองไม่แตกต่างกับค่าเฉลี่ยของประชากรในกลุ่มควบคุม จากแนวคิดดังกล่าวจึงแปลผลสรุปว่า **ไม่แตกต่างกัน** ซึ่งการสรุปเช่นนี้เหมาะสมกว่าสรุปว่า **แตกต่างกัน** อย่างไม่มีนัยสำคัญทางสถิติ

หมายเหตุ การสรุปว่า **ไม่แตกต่างกัน** ลอย ๆ เป็นการสรุปที่ไม่ชัดเจนเท่าที่ควร ถ้าจะให้ดียิ่งขึ้น ผู้วิจัยควรระบุระดับนัยสำคัญ ( $\alpha$ ) ไว้ด้วย โดยระบุในตอนต้น ดังตัวอย่างการแปลสรุปผลดังนี้

“จากการทดสอบสมมติฐานด้วย t-test โดยกำหนดระดับนัยสำคัญที่ระดับ .05 พบว่า นักเรียนชายและนักเรียนหญิงมีค่าเฉลี่ยไม่แตกต่างกัน”

### 3. เกี่ยวกับเครื่องหมายของค่า t

สถานการณ์ ผู้วิจัยทำการทดสอบสมมติฐาน โดยใช้การทดสอบค่าที (t-test) ผลปรากฏดังในตาราง

ผู้เรียน	ค่าเฉลี่ยก่อนทดลอง	ค่าเฉลี่ยหลังทดลอง	t
หน่วยที่ 1	2.34	2.57	3.4514
หน่วยที่ 2	2.41	3.01	4.2931
หน่วยที่ 3	2.64	3.25	2.3362

**ประเด็นปัญหา** ผลการวิเคราะห์ถูกต้องหรือไม่

**คำตอบ** การที่จะตอบว่าผลการวิเคราะห์ถูกต้องหรือไม่ ก็ต้องตรวจสอบจากข้อมูลอีกครั้งหนึ่งลองวิเคราะห์ตามสูตรที่กำหนด แล้วดูว่าได้ค่าตรงกันหรือไม่ อย่างไรก็ตามแม้ว่าคำนวณได้ค่า t ตรงกับในตารางทุกค่า ก็ยังจัดว่าไม่ถูกต้อง เพราะการเขียนเช่นนี้แสดงว่าตัวตั้งคือ ค่าเฉลี่ยก่อนการทดลอง ตัวลบคือค่าเฉลี่ยหลังการทดลอง ค่าตัวตั้งน้อยกว่าค่าตัวลบ ค่า t ต้องมีค่าติดลบ

ผู้วิจัยน่าจะสลับคอลัมน์กล่าวคือ ใช้ค่าเฉลี่ยหลังการทดลองเป็นตัวตั้ง ค่าเฉลี่ยก่อนการทดลองเป็นตัวลบ ซึ่งจะได้ค่า t เป็นบวกดังที่รายงานไว้ การแปลผลจะต้องระมัดระวังในเรื่องนี้ด้วย และในการกำหนดสมมติฐาน ควรกำหนดแบบมีทิศทางคือ **"ค่าเฉลี่ยหลังการทดลองมากกว่าค่าเฉลี่ยก่อนการทดลอง"** ซึ่งจะต้องทดสอบแบบหางเดียวด้วย

-----\*\*\*\*\*-----