

การปรับขนาดตัวอย่างเพิ่มสำหรับผู้ไม่ตอบ แบบสำรวจ

อรุณ จิรวัดนกุล

ภาควิชาชีวสถิติและประชากรศาสตร์ คณะสาธารณสุขศาสตร์ มหาวิทยาลัยขอนแก่น

การศึกษาเชิงพรรณนา/การสำรวจ เพื่อประมาณค่าอุบัติการณ์ ความชุก ความครอบคลุม หรืออัตราต่างๆในการศึกษาถ้าใช้วิธีสุ่มตัวอย่างอย่างง่ายจะใช้สูตรการคำนวณขนาดตัวอย่างเพื่อประมาณค่าสัดส่วนของประชากร $n = \frac{Z^2 P(1-P)}{d^2}$ ขนาดตัวอย่างที่คำนวณได้จะเป็นขนาดตัวอย่างที่น้อยที่สุดที่ใช้ในการศึกษา แต่ในทางปฏิบัติจะมีตัวอย่างที่สุ่มมาได้แต่เก็บข้อมูลไม่ได้ (Non-response) อาจมีสาเหตุจากคนที่สุ่มเป็นตัวอย่างไม่อยู่ขณะเก็บข้อมูล หรือปฏิเสธจะตอบแบบสำรวจ ทำให้เก็บข้อมูลได้ไม่ครบตามจำนวนที่คำนวณไว้ ซึ่งอาจมีผลกระทบต่อความน่าเชื่อถือของผลการศึกษา

เพื่อแก้ไขปัญหาดังกล่าว นักวิจัยจำเป็นต้องมีการคำนวณขนาดตัวอย่างเพื่อไว้เพื่อให้มีจำนวนตัวอย่างที่ศึกษาที่เก็บได้เท่ากับหรือมากกว่าที่คำนวณได้ เช่น ถ้าคาดหวังว่าจะเก็บข้อมูลไม่ได้ร้อยละ 20 นักวิจัยต้องคำนวณขนาดตัวอย่างเพิ่มเพื่อไว้ เพื่อให้เวลาเก็บจริงได้เมื่อหักที่เก็บไม่ได้ร้อยละ 20 แล้วยังมีจำนวนตัวอย่างที่เก็บได้เหลือครบตามจำนวนที่ต้องการ

อัตราการเก็บข้อมูลไม่ได้ควรกำหนดไว้เท่าไร อัตราดังกล่าวจะขึ้นอยู่กับวิธีที่ใช้ในการเก็บข้อมูล ถ้าเป็นการสำรวจทางไปรษณีย์ จำนวนการตอบกลับจะแปรตามความน่าสนใจของเรื่องที่ถาม และความยาวของแบบสอบถาม อัตราเก็บข้อมูลไม่ได้อาจจะอยู่ร้อยละ

25-35

กรณีเก็บข้อมูลโดยพนักงานสัมภาษณ์อัตราเก็บข้อมูลไม่ได้จะไม่เกินร้อยละ 5 แต่ถ้าเลือกเวลาเก็บข้อมูลไม่เหมาะสม เช่น เก็บช่วงเวลาทำงานคนตอบไม่อยู่บ้าน ถึงแม้จะเข้าไปสัมภาษณ์หลายรอบก็จะได้ไม่ครบหรือกรณีในกรุงเทพมหานคร ถึงแม้จะไปเก็บวันหยุดแต่คนตอบต้องการพักผ่อน หรือไม่ว่างก็จะไม่ให้สัมภาษณ์ ผู้วิจัยควรกำหนดอัตราเก็บข้อมูลไม่ได้ไว้อย่างน้อยร้อยละ 10

การปรับขนาดตัวอย่างเพิ่มขึ้นเพื่อแก้ไขปัญหาการเก็บข้อมูลไม่ได้ ที่พบมากคือผู้วิจัยคาดว่าจะมีผู้ที่เก็บข้อมูลไม่ได้ร้อยละ 20 เมื่อคำนวณขนาดตัวอย่างมาได้ 100 คน จะปรับเพิ่มอีก ร้อยละ 20 เป็น 120 คน ($n_{adj} = n + n \times 0.2$) ซึ่งไม่ถูกต้องทั้งนี้เพราะถ้าวางแผนที่จะเก็บ 120 คน เก็บไม่ได้ร้อยละ 20 จำนวนที่จะเก็บได้จริงจะได้เพียง 96 คน $[(120 \times 80)/100=96]$ ซึ่งน้อยกว่าที่คำนวณ ได้ 100 คน การปรับเพิ่มเพื่อทดแทนการเก็บข้อมูลไม่ได้จะทำโดยสูตร $n_{adj} = \frac{n}{1-r}$ โดย r คืออัตราการเก็บข้อมูลไม่ได้ (แทนค่าเป็นสัดส่วน เช่น ร้อยละ 20 ค่าสัดส่วนคือ 0.2) จากตัวอย่างข้างบน $n_{adj} = \frac{100}{1-0.2} = 125$ ขนาดตัวอย่างดังกล่าวเมื่อหักจำนวนคนที่เก็บข้อมูลไม่ได้แล้วจะเหลือขนาดตัวอย่างเท่าที่ต้องการ

การคำนวณขนาดตัวอย่างเพิ่มเพื่อไว้มากจะดีหรือไม่ การเพิ่มเพื่อไว้มากนอกจากจะสิ้นเปลืองค่าใช้จ่ายแล้ว จะมีผลกระทบต่อคุณภาพของข้อมูล และความเป็นตัวแทนของประชากร เพราะคนที่เก็บไม่ได้อาจมีลักษณะแตกต่างจากคนที่เก็บได้ ถ้ามีมากจะทำให้ผลสรุปที่ได้มีอคติ (Non-response bias) ซึ่งอาจต้องใช้วิธีการทางสถิติวิเคราะห์เพื่อลดอิทธิพลของอคติที่เกิดขึ้น

สรุปในการคำนวณขนาดตัวอย่างเพื่อการสำรวจ จำนวนตัวอย่างที่คำนวณได้จะเป็นขนาดตัวอย่างน้อยที่สุดที่ใช้ในการตอบคำถามหลักของงานวิจัย ในการ

เก็บจริงนักวิจัยต้องวางแผนวิธีเก็บข้อมูลให้มีอัตราการเก็บข้อมูลไม่ได้น้อยที่สุด และต้องคำนวณขนาดตัวอย่างเพิ่มเพื่อการเก็บข้อมูลไม่ได้ด้วย โดยต้องใช้วิธีการคำนวณที่ถูกต้อง

เอกสารอ้างอิง

1. อรุณ จิรวัดน์กุล. สถิติทางวิทยาศาสตร์สุขภาพเพื่อการวิจัย. พิมพ์ครั้งที่ 1. กรุงเทพมหานคร: วิทย์พัฒนา; 2552.