

การวิเคราะห์ข้อมูลจากแบบวัดที่มีคำตอบ ไม่ครบทุกข้อ

อรุณ จิรวัดนกุล

ภาควิชาชีวสถิติและประชากรศาสตร์ คณะสาธารณสุขศาสตร์ มหาวิทยาลัยขอนแก่น

แบบวัด (scale) เป็นชุดคำถามเพื่อใช้วัดค่าตัวแปรที่ไม่สามารถวัดค่าได้ด้วยคำถามเพียงข้อเดียว เช่น แบบวัดคุณภาพชีวิตขององค์การอนามัยโลกอย่างย่อ มีคำถามทั้งหมด 26 ข้อ เพื่อใช้วัดคุณภาพชีวิตโดยรวม ในแบบวัดยังแบ่งออกเป็นองค์ประกอบต่าง ๆ 4 ด้าน เช่น องค์ประกอบด้านสุขภาพกาย มีคำถามทั้งหมด 7 ข้อ เป็นข้อคำถามเชิงลบ 2 ข้อเชิงบวก 5 ข้อ ฯลฯ

เมื่อมีคำถามบ้างข้อที่ผู้ตอบไม่ได้ตอบทำให้ไม่สามารถให้ค่ากับคำตอบข้อนั้น ๆ มีผลทำให้คะแนนรวมที่ได้มีคะแนนน้อยลง ถ้านำผลรวมเท่าที่ได้มาใช้เป็นค่าของตัวแปรที่วัด ทำให้ได้ผลรวมที่มีอคติ มีค่าน้อยกว่าความจริง

ความไม่เหมาะสมที่พบบ่อยในการวิเคราะห์ข้อมูลแบบวัดที่มีข้อมูลบ้างข้อสูญหายจากการตอบคำถามไม่ครบ คือการนำค่าเฉลี่ยของข้อที่มีคำตอบในชุดเดียวกันมาแทนค่าข้อมูลที่สูญหาย โดยไม่พิจารณาเหตุของการที่ไม่ตอบ ทำให้ได้ผลรวมที่มีอคติ

ข้อมูลสูญหายที่เกิดจากการตอบคำถามในแบบวัดไม่ครบแบ่งออกเป็นสองประเภทคือ สูญหายเกิดอย่างสุ่ม และสูญหายเกิดอย่างไม่สุ่ม

ข้อมูลสูญหายที่เกิดอย่างสุ่ม เกิดจากผู้ตอบไม่ได้ตั้งใจที่จะไม่ตอบข้อนั้น แต่อาจตอบข้ามบางข้อด้วยความสะเพร่า หรือด้วยความรีบร้อน

ข้อมูลสูญหายเกิดอย่างไม่สุ่ม เกิดจากผู้ตอบไม่ต้องการตอบคำถามบางข้อ เช่น แบบประเมินคุณภาพ

ชีวิตขององค์การอนามัยโลกด้านความสัมพันธ์ทางสังคม มีหนึ่งข้อประเมินความพอใจในชีวิตทางเพศ คำถามข้อนี้ผู้หญิงจะลำบากใจในการตอบ โดยเฉพาะผู้ที่ไม่เคยมีเพศสัมพันธ์ จึงมีโอกาสไม่ตอบข้อนี้

ข้อมูลสูญหายเกิดอย่างไม่สุ่มจึงมีสาเหตุมาจากคำถามที่ผู้ตอบรู้สึกลำบากใจ หรืออับอายที่จะตอบ ส่วนใหญ่จะเกี่ยวกับการถามเรื่อง เพศ รูปลักษณ์ ความสามารถ และระดับสติปัญญา

อีกสาเหตุหนึ่งเกิดจากคำถามไม่ดีผู้ตอบไม่เข้าใจ หรือคำถามไม่เหมาะสมกับบริบทของผู้ตอบทำให้ผู้ตอบไม่สามารถตอบได้ เช่น ถามผู้สูงอายุว่า “ท่านมีปัญหาในการเดินขึ้นลงบันไดในการการดำเนินชีวิตประจำวันหรือไม่” ถ้าผู้สูงอายุที่อาศัยอยู่ในบ้านชั้นเดียว หรือการดำเนินชีวิตปกติอยู่แต่ชั้นล่าง จะตอบคำถามข้อนี้ไม่ได้ การวิเคราะห์ข้อมูลแบบวัดที่มีข้อมูลบ้างข้อสูญหายอาจทำได้โดย

1. ตัดชุดที่มีข้อมูลไม่ครบออกจากการวิเคราะห์
2. ใช้วิธีการทางสถิติคำนวณค่าที่ใช้แทนค่าข้อมูลที่สูญหาย

การใช้วิธีการทางสถิติคำนวณค่าแทนข้อมูลที่สูญหายใช้ได้ดีกับการสูญหายเกิดอย่างสุ่ม ถ้าข้อมูลสูญหายเกิดอย่างไม่สุ่ม ค่าที่คำนวณแทนอาจทำให้ได้ผลรวมที่มีอคติ ดังนั้นการใช้วิธีการทางสถิติคำนวณค่าแทนจึงควรมีการพิจารณาลักษณะข้อมูลที่สูญหายว่าเกิดอย่างสุ่มหรือไม่ดังนี้

1) ลักษณะของกลุ่มผู้ที่ไม่ตอบข้อนี้มีลักษณะเหมือนกลุ่มผู้ตอบหรือไม่ เช่น การไม่ตอบคำถามความพอใจในชีวิตทางเพศ ถ้าพบว่ากลุ่มไม่ตอบเป็นหญิงโสดมากกว่ากลุ่มที่ตอบ การจะนำค่าเฉลี่ยของผู้ตอบที่ส่วนใหญ่เป็นคนที่ม่ประสบการณ์ทางเพศมาเป็นค่าแทนกลุ่มที่ไม่ตอบ ทำให้ผลรวมที่ได้มีอคติ

2) ข้อคำถามที่มีผู้ไม่ตอบมากกว่าข้ออื่น อาจเกิดจากเขียนคำถามไม่ดีมีผู้ตอบส่วนหนึ่งไม่เข้าใจเลย ไม่ตอบ หรือรูปแบบการจัดพิมพ์ไม่ดีทำให้มีโอกาสตอบข้ามข้อนี้ไปโดยไม่ตั้งใจ

3) การไม่ตอบข้อคำถามอาจมีปัจจัยเฉพาะในเรื่องที่ศึกษา เช่น ในการศึกษาคุณภาพชีวิตของผู้ป่วย โดยผู้ให้การรักษาพยาบาล การไม่ตอบของผู้ป่วยแต่ละคนอาจมีเหตุผลที่ต่างกัน ดังนั้นการใช้ค่าเฉลี่ยของกลุ่มเป็นค่าแทนข้อมูลที่สูญหาย จะมีโอกาสเกิดได้ผลรวมที่มีอคติ

ในกรณีที่คำถามที่ไม่มีคำตอบ มีไม่เกินร้อยละหนึ่งของคำถามทั้งหมด และนักวิจัยพิจารณาแล้วเห็นว่าการสูญหายน่าจะเกิดอย่างสุ่ม สามารถใช้ค่าเฉลี่ยของกลุ่มที่ตอบในข้อเดียวกันแทนได้ ไม่ควรใช้ค่าเฉลี่ยของคำตอบที่ได้ของชุดแบบวัดเดียวกัน เพราะในแบบวัดแต่ละชุดจะมีองค์ประกอบหลายด้าน และข้อคำถามมีทั้งที่เป็นเชิงบวกและลบ การนำค่าเฉลี่ยของคำตอบในชุดเดียวกันมาแทนค่าที่สูญหายย่อมทำให้ได้ผลรวมที่มี

อคติ

ในการใช้วิธีการทางสถิติคำนวณค่าของข้อมูลสูญหายมีหลายวิธี ความเหมาะสมของแต่ละวิธีจะขึ้นอยู่กับลักษณะของการไม่ตอบ ความสัมพันธ์ของตัวแปรที่มีข้อมูลสูญหายกับตัวแปรอื่นที่เก็บมาศึกษา และจำนวนข้อที่สูญหาย ในการดำเนินการควรปรึกษานักสถิติในการเลือกวิธีการคำนวณค่าแทนข้อมูลสูญหาย

สรุป การวัดผลโดยใช้แบบวัดเมื่อมีข้อมูลสูญหาย บางข้อจากการไม่ตอบ เพื่อไม่ให้ได้ผลรวมที่มีอคติ นักวิจัยต้องพิจารณาว่าการไม่ตอบเกิดอย่างสุ่มหรือไม่ ก่อนที่จะเลือกใช้วิธีการทางสถิติที่เหมาะสมในการคำนวณค่าแทนข้อมูลสูญหาย

เอกสารอ้างอิง

1. Fayers PM, Curran D, Machin D. Incomplete quality of life data in randomized trials: missing items. *Statistics in Medicine* 1998; 17:679-96.
2. Fayers PM, Machin D. *Quality of life: assessment, analysis and interpretation*. West Sussex: John Wiley & Son; 2007.
3. Fielding S, Fayers PM, McDonald A, McPherson G, Campbell MK. Simple imputation methods were inadequate for missing not at random (MNAR) quality of life data. *Health and Quality of Life Outcome* 2008; 6:57-65.