

## การวัดความสอดคล้องด้วยสถิติ

### อรุณ จิรวัดเนกุล

ภาควิชาชีวสถิติและประชากรศาสตร์ คณะสาธารณสุขศาสตร์ มหาวิทยาลัยขอนแก่น

ในการประเมินความสอดคล้องระหว่างผู้ประเมิน 2 คน หรือมากกว่า ในกรณีที่ผลการประเมินเป็นข้อมูลต่อเนื่อง หรือ คะแนนที่วัดจากเครื่องมือวัดที่เป็น summated rating scale (ชุดคำถาม) เช่น เครื่องมือวัดคุณภาพชีวิต เครื่องมือวัดความรู้ทัศนคติและการปฏิบัติตัว ฯลฯ

การประเมินความสอดคล้องในกรณีที่ตัวแปรผลเป็นข้อมูลต่อเนื่องจะมีการวิเคราะห์ที่แตกต่างกันไปตามวิธีการประเมิน การสุ่มหรือไม่สุ่มผู้ประเมิน ในบทความนี้จะขอยกกรณีของการประเมินตัวอย่างเดียวกันโดยผู้ประเมิน 2 คน ซึ่งใช้ในโครงการวิจัยที่จำเป็นต้องมีผู้ตรวจวิเคราะห์ หรือผู้ประเมินสองคน และต้องการดูว่าผู้ตรวจวิเคราะห์/ประเมินทั้งสองให้ผลการตรวจที่สอดคล้องกันหรือไม่

สถิติที่ใช้ในกรณีนี้คือ Intraclass correlation coefficient (ICC) โดยที่ ICC มีค่าน้อยสุดเท่ากับศูนย์ (0) มากที่สุดเท่ากับ 1 วิธีการอ่านแปรผลค่า ICC ให้ใช้เกณฑ์เดียวกับการอ่านผลค่า Kappa ที่ได้เสนอบนบทความเรื่องการวัดความสอดคล้องด้วยสถิติ Kappa แนวคิดของการคำนวณ ค่า ICC จะคำนวณจากค่าความแปรปรวนที่เกิดจากหน่วยตัวอย่าง กับความแปรปรวนระหว่างผู้ประเมินดังสูตรข้างล่างนี้

$$ICC = \frac{\text{Variance due to subjects (patients)}}{\text{Variance due to subjects} + \text{Variance due to Raters} + \text{Residual Variance}}$$

จากสูตร ถ้าความแปรปรวนระหว่างผู้ประเมิน (Variance due to Raters) มีค่าน้อย ๆ ค่า ICC จะมีค่าใกล้ 1 ถ้าความแปรปรวนระหว่างผู้ประเมินที่มีค่ามากๆ ค่า ICC จะมีค่าใกล้ศูนย์ ตัวอย่างในงานวิจัยเพื่อประเมินทักษะของแม่ในการดูแลลูกป่วย โดยมีพยาบาลที่จะเป็นผู้ประเมิน 2 คน ทีมวิจัยได้จัดอบรมผู้ประเมิน และได้ทดสอบความสอดคล้องของผู้ประเมินทั้งสอง ได้ผลดังตารางข้างล่างนี้

ลำดับที่	ผลการประเมิน	
	ผู้ประเมิน 1	ผู้ประเมิน 2
1	9	8
2	6	3
3	8	8
4	7	6
5	5	6
6	10	9
7	6	7
8	5	5
9	4	4
10	5	7

ใช้ Two-way ANOVA วิเคราะห์หาความแปรปรวนระหว่างหน่วยตัวอย่าง ระหว่างผู้ประเมิน และจากการสุ่ม (Residual) เพื่อนำไปแทนค่าในสูตรคำนวณค่า ICC

โปรแกรมสถิติบางโปรแกรมจะมีคำสั่งวิเคราะห์หาค่า ICC โดยตรง เช่น SPSS จะวิเคราะห์ด้วยคำสั่ง Reliability (ในเมนูเลือก Scale → Reliability analysis) นำข้อมูลจากตารางไปคำนวณค่า ICC ด้วยโปรแกรม SPSS ได้ค่า ICC เท่ากับ 0.752 โดยมีค่า 95% ช่วงเชื่อมั่นอยู่ระหว่าง 0.266 ถึง 0.932

จากค่า ICC ที่ได้แสดงว่าผลการประเมินความสอดคล้องของพยาบาลทั้งสองคนมีความสอดคล้องอยู่ในระดับดีมาก

ข้อที่นักวิจัยพึงตระหนักในการดำเนินการประเมินความสอดคล้องของผู้ประเมิน คือ การฝึกอบรม และวิธีการดำเนินการประเมิน (SOP) ซึ่งมีผลโดยตรงต่อคุณภาพการประเมินในกรณีผลการประเมินมีความสอดคล้องกันน้อย ควรพิจารณาว่าวิธีการประเมินรวมทั้งคู่มือการประเมินว่ามีข้อบกพร่องหรือไม่ ถ้ามีเมื่อดำเนินการแก้ไขให้ผู้ประเมินมีความเข้าใจตรงกัน เมื่อทำการทดสอบใหม่ ค่า ICC จะสูงขึ้น แต่ถ้าวิธีการประเมินชัดเจนดี แต่ผู้ประเมินมีทักษะ หรือการยินยอมทำตามขั้นตอนประเมินที่แตกต่างกัน ควรเปลี่ยนผู้ประเมิน

ในกรณีที่คำถามหลักของงานวิจัยเป็นการประเมินความสอดคล้อง นักวิจัยจะต้องมีการคำนวณขนาดตัวอย่าง 4 เพื่อให้ได้ช่วงเชื่อมั่นที่กระชับ

สรุป การประเมินความสอดคล้องในกรณีข้อมูลที่วัดได้เป็นข้อมูลต่อเนื่อง หรือคะแนนจากเครื่องมือวัด (ชุดคำถาม) ด้วยค่าสถิติ ICC ซึ่งมีค่าระหว่าง 0-1 ในกรณีที่ผลการประเมินมีความสอดคล้องกันน้อยควรพิจารณาแก้ไขตามสาเหตุ และถ้าคำถามความสอดคล้องเป็นคำถามหลักต้องมีการคำนวณขนาดตัวอย่าง

#### บรรณานุกรม

1. Fleiss JL. Statistical Methods for Rate and Proportions. New York: John Wiley & Sons; 2003.
2. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977; 33:159-74.
3. Roberts AR, Yeager KR. Evidence-based practice manual: research and outcome measures in health and human services. New York: Oxford University Press; 2003.
4. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. Stat Med 1998; 17:101-10.