www.wphat.com

เหมาะสำหรับบุคลากรและนักวิจัยในสาขาวิทยาศาสตร์สุขภาพ
รวมถึงนักวิจัยในสาขาอื่นที่ต้องการใช้สถิติในงานวิจัยให้ถูกต้อง

สถิติในงานวิจัย
เลือกใช้อย่างไรให้เหมาะสม

Statistics in Research: Appropriate Selection

- ตัวอย่างการใช้สถิติในการสรุปคำตอบคำถามวิจัยต่างๆ
- ชี้ประเด็นการใช้ที่ไม่ถูกต้องพร้อมคำอธิบายเหตุผลทางสถิติ
- รวบรวมข้อผิดพลาดต่างๆของการใช้สถิติในบทความวิจัย
- แสดงการเลือกใช้สถิติให้สอดคล้องกับคำถามวิจัย

รองศาสตราจารย์อรุณ จิรวัฒน์กุล

สถิติในงานวิจัย
เลือกใช้อย่างไรให้เหมาะสม

Statistics in Research: Appropriate Selection

รองศาสตราจารย์อรุณ จิรวัฒน์กุล

การทำสำเนา ลอกเลียน ดัดแปลงหนังสือเล่มนี้ ไม่ว่าเพียงบางส่วนหรือทั้งหมด
เป็นการละเมิดลิขสิทธิ์ มีความผิดทั้งทางแพ่งและอาญา ผู้ละเมิดลิขสิทธิ์จะถูกดำเนินคดีจนถึงที่สุด

ผู้ใดให้เบาะแสของผู้ละเมิดลิขสิทธิ์ จะได้รับรางวัลเป็นเงิน 10,000 บาท หลังจากผู้ละเมิดลิขสิทธิ์ถูกจับกุมแล้ว
และจะได้รับเงินอีก 20,000 บาท หลังจากการดำเนินคดีถึงที่สุดโดยศาลพิพากษาลงโทษในทางอาญาผู้ละเมิดลิขสิทธิ์แล้ว

Using Statistics in Reseach 1_1 .indd 1 10/9/2557 15:58:54

สถิติในงานวิจัย เลือกใช้อย่างไรให้เหมาะสม

รองศาสตราจารย์อรุณ จิรวัฒน์กุล

ฉบับพิมพ์ที่ 1 พิมพ์ครั้งแรก กันยายน 2557
พิมพ์ซ้ำครั้งที่ 2 เมษายน 2558

สงวนสิทธิ์ตามพระราชบัญญัติลิขสิทธิ์ พ.ศ. 2537

ห้ามทำซ้ำ ดัดแปลง คัดลอก ลอกเลียน หรือนำไปเผยแพร่ในสื่อทุกประเภท ไม่ว่าส่วนใดส่วนหนึ่งของหนังสือเล่มนี้ ตลอดจนห้ามมิให้สแกนหนังสือหรือคัดลอกส่วนใดส่วนหนึ่งเพื่อสร้างฐานข้อมูลอิเล็กทรอนิกส์ นอกจากจะได้รับอนุญาตเป็นลายลักษณ์อักษรจากบริษัทวิทยพัฒน์ จำกัด

จัดทำรูปเล่ม จัดพิมพ์ และจำหน่ายโดย

บริษัทวิทยพัฒน์ จำกัด
52/103-104 บางกะปิสแควร์ ถนนรามคำแหง
เขตบางกะปิ กรุงเทพฯ 10240
โทรศัพท์ 02 3749915 (หลายคู่สาย)
โทรสาร 02 3746495
ที่อยู่อีเมล contact@wphat.com

พิมพ์ที่ บริษัท ส. เอเซียเพรส (1989) จำกัด
143, 145 ซอยรามคำแหง 42 แขวงหัวหมาก เขตบางกะปิ กรุงเทพฯ 10240

ราคา 130 บาท

ข้อมูลทางบรรณานุกรมของหอสมุดแห่งชาติ

อรุณ จิรวัฒน์กุล.
สถิติในงานวิจัย เลือกใช้อย่างไรให้เหมาะสม.--กรุงเทพฯ: วิทยพัฒน์, 2558.
208 หน้า.
1. วิจัย. I. ชื่อเรื่อง.
001.4

ISBN 978-616-7136-71-4

ท่านที่ต้องการสั่งซื้อหนังสือเล่มนี้ กรุณาสอบถามหรือสั่งซื้อได้ที่ บริษัทวิทยพัฒน์ จำกัด โทร. 02 3749915 หรือตามที่อยู่ด้านบน

หากท่านมีข้อติชม หรือคำแนะนำเกี่ยวกับหนังสือหรือบริการของบริษัทฯ กรุณาส่งจดหมายถึงผู้จัดการฝ่ายลูกค้าสัมพันธ์ตามที่อยู่ด้านบน หรือส่งอีเมลที่ admin@wphat.com จักเป็นพระคุณยิ่ง

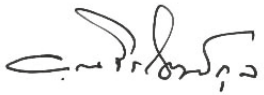
คำนำ

งานวิจัยด้านวิทยาศาสตร์สุขภาพที่เป็นงานวิจัยเชิงปริมาณ ส่วนใหญ่เป็นการศึกษาจากตัวอย่างแล้วใช้สถิติอนุมานผลการศึกษาไปเป็นคำตอบของประชากร ความถูกต้องของการใช้สถิติในขั้นตอนต่าง ๆมีผลต่อความถูกต้องของข้อสรุปที่ได้จากงานวิจัย ผู้เขียนมีประสบการณ์มากกว่า 20 ปีในฐานะผู้ทำวิจัย ผู้สอนวิชาสถิติในงานวิจัย ผู้ให้คำปรึกษาด้านการใช้สถิติในงานวิจัย ผู้ประเมินโครงการวิจัยเพื่อขอทุน ผู้ประเมินบทความวิจัยเพื่อขอลงตีพิมพ์ และผู้วิพากษ์การนำเสนอผลงานวิจัย ได้พบเห็นข้อผิดพลาด ความไม่เหมาะสม และความเข้าใจผิดที่เกี่ยวข้องกับการใช้สถิติในงานวิจัยด้านวิทยาศาสตร์สุขภาพจำนวนมาก จึงได้เขียนบทความวิชาการเพื่อชี้ให้เห็นปัญหาการใช้และวิธีการแก้ไขลงในคอลัมน์ *มุมสถิติ* ในวารสารวิชาการสาธารณสุข กระทรวงสาธารณสุข ตั้งแต่ปี พ.ศ. 2547 จนถึงปัจจุบัน

"Learn from the mistakes of others, you can't make them all yourself." Confession No. 3 ของ Usher นักร้องชาวอเมริกัน ตรงกับใจที่ผู้เขียนอยากให้เป็นมากที่สุด การที่นักวิจัยได้เรียนรู้ความผิดพลาดของการใช้สถิติในงานวิจัยที่ผ่านมาจะช่วยให้ไม่ต้องไปทำผิดซ้ำอีก จึงได้รวบรวมบทความดังกล่าวมาปรับปรุงให้ทันสมัยมากขึ้น แล้วมาจัดทำเป็นหนังสือ **สถิติในงานวิจัย เลือกใช้อย่างไรให้เหมาะสม** เล่มนี้

ในหนังสือเล่มนี้ผู้เขียนจัดกลุ่มบทความออกเป็น 3 กลุ่ม กลุ่มแรกอยู่ในขั้นออกแบบงานวิจัย การคำนวณขนาดตัวอย่าง แผนการวิเคราะห์ บางส่วนเกี่ยวกับแบบงานวิจัยและการสร้างเครื่องมือ กลุ่มที่สองเป็นปัญหาหรือความไม่เหมาะสมในการวิเคราะห์ และกลุ่มสุดท้ายเป็นเรื่องของการสรุปแปลผลและการนำเสนอผลวิจัย

สุดท้ายขอบคุณครอบครัวที่สนับสนุนและเป็นกำลังใจ ขอบคุณอาจารย์นิตยา จันทร์เรือง มหาผล บรรณาธิการวารสารสาธารณสุขที่ให้โอกาสและสนับสนุนให้เขียนคอลัมน์ *มุมสถิติ* ขอขอบคุณกองบรรณาธิการและสำนักพิมพ์วิทยพัฒน์ที่ดำเนินการแก้ไขปรับปรุงและจัดพิมพ์ ท้ายสุดขอมอบประโยชน์ที่พึงเกิดจากหนังสือเล่มนี้เป็นคารวะบูชาแด่ครูผู้ประสาทวิชาและผู้ที่มีส่วนเป็นเจ้าของปัญหาทุกท่าน



(รองศาสตราจารย์อรุณ จิรวัฒน์กุล)

สารบัญ

สถิติในขั้นออกแบบงานวิจัย 7

1 การเลือกตัวอย่างกับการอนุมานด้วยสถิติ 8

2 การเลือกตัวอย่างแบบเจาะจงกับการอ้างอิงทางสถิติ 12

3 ขนาดตัวอย่างกับจริยธรรมการวิจัยในมนุษย์ 15

4 การเลือกตัวอย่างผู้ป่วยที่เข้ารับการรักษาในโรงพยาบาล เป็นการเลือกตัวอย่างแบบใด 19

5 วิธีการคำนวณขนาดตัวอย่างและความเหมาะสมของการใช้ขนาดตัวอย่างจากตารางของยามาเน่ 22

6 การกำหนดขนาดตัวอย่างกลุ่มละ 30 คน เพียงพอจริงหรือ 28

7 การใช้ตารางคำนวณขนาดตัวอย่างของยามาเน่อย่างถูกต้อง 32

8 บทบาทของตัวแปรผลในการคำนวณขนาดตัวอย่าง 36

9 การเลือกค่าพารามิเตอร์สำหรับการคำนวณขนาดตัวอย่างเพื่อการประมาณค่าสัดส่วน 39

10 การกำหนดค่า e ในการคำนวณขนาดตัวอย่างเพื่อการศึกษาความชุก 45

11 สถิติและการคำนวณขนาดตัวอย่างงานวิจัยเชิงสำรวจภาคตัดขวาง 50

12 การคำนวณขนาดตัวอย่างในการศึกษาความชุกและปัจจัยเสี่ยงพร้อมกัน 54

13 ความเหมาะสมของการใช้การสุ่มตัวอย่างแบบกลุ่มในการศึกษาปัจจัยเสี่ยง 58

14 การสุ่มตัวอย่างเข้าในการศึกษากับการสุ่มตัวอย่างเข้ากลุ่มการทดลอง 62

15 การปรับขนาดตัวอย่างเพิ่มสำหรับผู้ไม่ตอบแบบสำรวจ 65

16 การปรับขนาดตัวอย่างเพิ่มสำหรับการไม่ปฏิบัติตามแผนการทดลอง 68
17 การกำหนดประชากรศึกษา การสุ่มตัวอย่าง และการจัดสรรตัวอย่าง 72
18 ความหมายของ univariate & multivariate และ
การเขียนแผนการวิเคราะห์.....76
19 การตั้งวัตถุประสงค์งานวิจัยกับการวิเคราะห์ข้อมูล.....80
20 การเขียนวิธีการวิเคราะห์ทางสถิติในโครงร่างวิจัย.....84
21 ความถูกต้องในการถามเหตุการณ์ย้อนหลัง.....88

สถิติในการวิเคราะห์ข้อมูลวิจัย 91

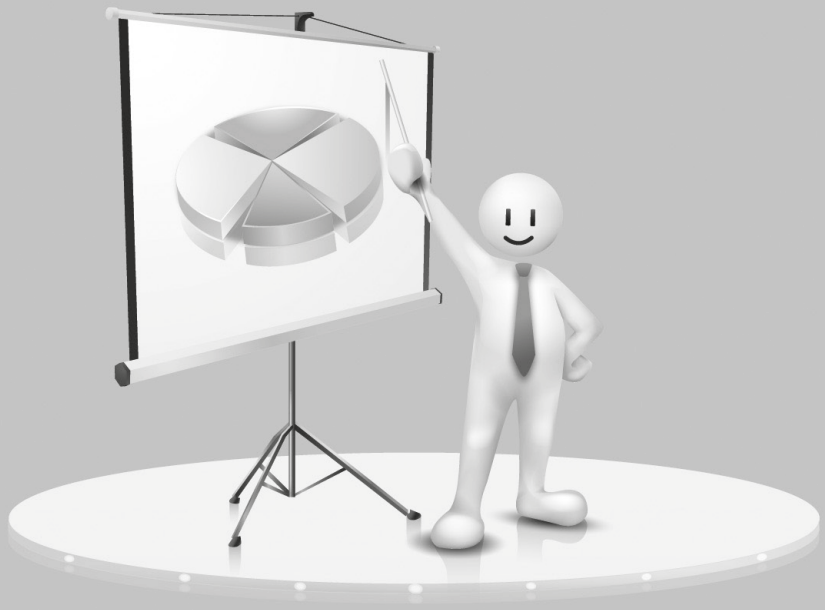
22 หน่วยวิเคราะห์กับความถูกต้องของผลสรุป 92
23 สถิติที่ใช้ในการวิเคราะห์ปัจจัยการเกิดโรค 96
24 การใช้ adjusted OR วิเคราะห์ค่าปัจจัยเสี่ยง 101
25 ผลที่เกิดจากการแบ่งกลุ่มตัวแปรต่อเนื่อง 106
26 การวิเคราะห์ข้อมูลจากแบบวัด.....111
27 การใช้สถิติสัมประสิทธิ์สหสัมพันธ์อย่างไม่เหมาะสม 116
28 การวิเคราะห์ทางสถิติสำหรับข้อมูลที่ศึกษาทั้งประชากร..... 120
29 ความไม่เหมาะสมของการทดสอบสมมติฐานกับตัวอย่างแบบไม่สุ่ม 124
30 การทดสอบความแตกต่างกับการสรุปความสัมพันธ์ระหว่างตัวแปร..... 127
31 การวิเคราะห์ข้อมูลจากแบบวัดที่มีคำตอบไม่ครบทุกข้อ 130
32 การปรับอิทธิพลของตัวแปรกวนในงานวิจัยแบบกึ่งทดลอง 134
33 การวิเคราะห์ผลการทดลองในกรณีมีผู้ออกระหว่างการทดลอง..... 137
34 การวัดความสอดคล้องด้วยสถิติ Kappa 140

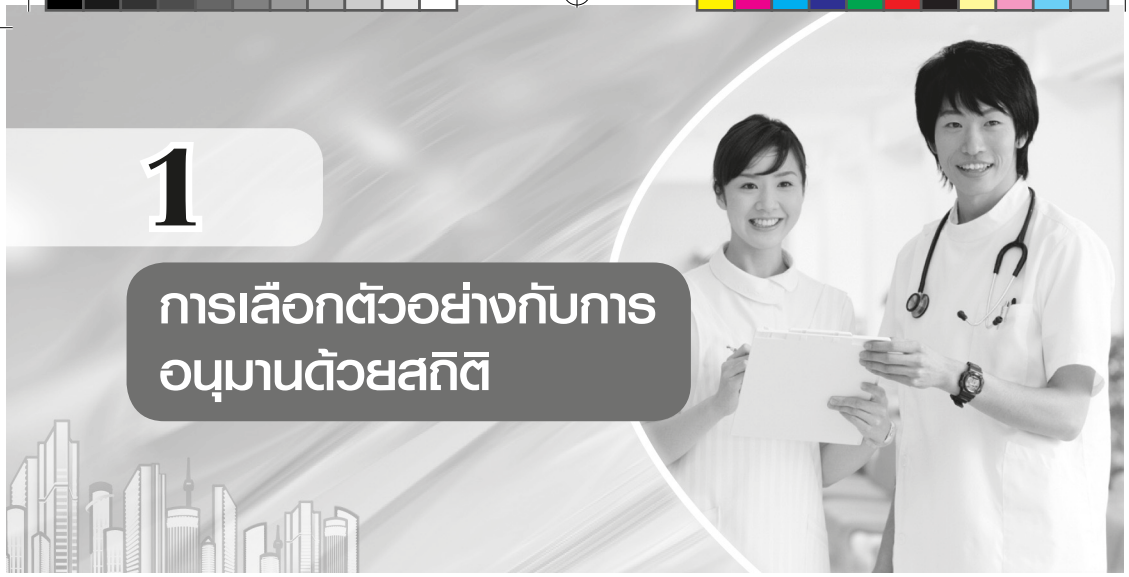
35 การวัดความสอดคล้องด้วยสถิติ Intraclass Correlation Coefficient ... 145
36 การวิจัยเชิงทดลองการสอนควรใช้ผลประเมินความรู้ช่วงใดสรุปผลวิจัย 149

สถิติในการสรุปแปลผลและนำเสนอผลวิจัย 152

37 ความแตกต่างอย่างมีนัยสำคัญ สำคัญจริงหรือ 153
38 การแปลความหมาย P value ที่ไม่ถูกต้อง 157
39 การนำเสนอผลต่างของการสอนด้วยขนาดอิทธิพล 160
40 ประโยชน์ของการนำเสนอช่วงเชื่อมั่นของผลการเปรียบเทียบ 163
41 ทำไมต้องรายงานช่วงเชื่อมั่นในรายงานวิจัย 167
42 การนำเสนอข้อมูลพื้นฐานในบทความวิจัย 170
43 ผลการทดสอบสมมติฐานไม่แตกต่างกันอย่างมีนัยสำคัญ ควรสรุปอย่างไร 174
44 การแปลผลการทดสอบสมมติฐานของค่าเฉลี่ย 177
45 การนำเสนอข้อมูลพื้นฐานในงานวิจัยเชิงทดลอง 181
46 การแปลผลการทดสอบสมมติฐานของค่าสัดส่วน 185
47 ทำไมจึงควรนำเสนอช่วงเชื่อมั่นร่วมกับอัตราหรือค่าเฉลี่ย ในการนำเสนอผลการสำรวจด้วยตัวอย่าง 189
48 การแปลผลการทดสอบทางสถิติของปัจจัยเสี่ยง 192
49 การแปลผลความสัมพันธ์ข้อมูลจากแบบวัด 196
50 การพิจารณานำงานวิจัยไปใช้ 200

สถิติในขั้นออกแบบงานวิจัย





1 การเลือกตัวอย่างกับการอนุมานด้วยสถิติ

งานวิจัยทุกประเภทที่ศึกษาด้วยกลุ่มตัวอย่างจะใช้ตรรกะหรือวิธีการทางสถิติอนุมานสรุปคำตอบจากตัวอย่างไปเป็นคำตอบของประชากร การจะใช้ตรรกะหรือวิธีการทางสถิติสรุปผลได้ถูกต้องขึ้นอยู่กับวิธีการได้มาของตัวอย่างที่ใช้ในการศึกษา

การเลือกตัวอย่างในงานวิจัย แบ่งได้เป็น 2 ลักษณะ คือ

1) **การเลือกตัวอย่างแบบไม่ใช้ความน่าจะเป็น** (non-probability sampling) เช่น เลือกแบบโควตา (quota sampling) เลือกแบบบังเอิญ (accidental sampling) และเลือกแบบเจาะจง (purposive sampling)
2) **การเลือกตัวอย่างแบบใช้ความน่าจะเป็น** (probability sampling) หรือแบบสุ่ม (random) การเลือกตัวอย่างโดยวิธีนี้สามารถคำนวณค่าความน่าจะเป็นที่หน่วยศึกษาถูกเลือกเป็นตัวอย่าง และอาศัยความน่าจะเป็นไปใช้ในการอนุมานค่าพารามิเตอร์ของประชากร การเลือกตัวอย่างด้วยวิธีนี้ ได้แก่ การสุ่มแบบง่าย (simple random sampling) การสุ่มแบบระบบ

(systematic sampling) การสุ่มแบบกลุ่ม (cluster sampling) เป็นต้น

การอ้างอิงผลการศึกษา

1) **การสรุปผลจากตัวอย่างที่เลือกแบบไม่ใช้ความน่าจะเป็น ตัวอย่างที่นำมาศึกษาถูกเลือกมาโดยเงื่อนไขไม่ได้เลือกอย่างสุ่ม** ทำให้ไม่สามารถคำนวณความน่าจะเป็นที่หน่วยศึกษาถูกเลือกมาเป็นตัวอย่าง จึงไม่สามารถใช้วิธีการทางสถิติในการประมาณค่าหรือทดสอบสมมติฐานได้ การศึกษาที่เลือกตัวอย่างโดยวิธีนี้ ส่วนใหญ่ต้องการบรรยายความคิดเห็น ทรรศนะ การรับรู้ และลักษณะอื่นๆของกลุ่มตัวอย่าง โดยใช้ตรรกะในการแปลและอนุมานผล นิยมใช้ในการสำรวจเร่งด่วนที่ต้องการผลเบื้องต้นเร็ว และในงานวิจัยเชิงคุณภาพ
2) **การสรุปผลจากตัวอย่างที่เลือกแบบใช้ความน่าจะเป็นที่แต่ละหน่วยศึกษาถูกเลือกมาอย่างสุ่ม** มีวิธีการทางสถิติสำหรับอนุมานผลจากตัวอย่างไปเป็นข้อสรุปของประชากร 2 วิธีใหญ่ๆ ดังนี้
    2.1) การอนุมานด้วยแนวคิดเดิม Frequentist approach วิธีนี้จะใช้การแจกแจงของค่าข้อมูล การแจกแจงค่าสถิติของตัวอย่าง หรือการแจกแจงตำแหน่งของข้อมูล (rank) คำนวณความน่าจะเป็นสำหรับการประมาณค่าและทดสอบสมมติฐาน
        ก) การอนุมานแบบพาราเมตริก จะประมาณค่าและทดสอบสมมติฐานโดยอาศัยการแจกแจงของค่าสถิติของตัวอย่าง เช่น t-test, Z-test และ F-test เป็นต้น

Using Statistics in Reseach 1_1 .indd 9 dummy 4 27/8/2557 11:48:24

ข) *การอนุมานแบบไม่ใช้พารามิเตอร์* วิธีนี้จะไม่ใช้การแจกแจงของค่าสถิติของตัวอย่างในการประมาณค่าและทดสอบสมมติฐาน ทำได้ใน 2 ลักษณะ คือ

- ▲ แปลงค่าข้อมูลให้เป็นตำแหน่ง (rank) แล้วจึงนำค่าตำแหน่งที่ได้มาคำนวณค่าสถิติ เช่น Wilcoxon rank-sum test, Wilcoxon signed-rank test และ Kruskal-Wallis test เป็นต้น ค่า P value ของการทดสอบด้วยค่าตำแหน่งคำนวณโดยใช้ความน่าจะเป็นหรือกรณีขนาดตัวอย่างใหญ่จะคำนวณโดยใช้การแจกแจงความน่าจะเป็น เช่น Wilcoxon rank-sum test เมื่อตัวอย่างขนาดใหญ่สถิติของค่าตำแหน่งมีการแจกแจงแบบทวินาม จึงคำนวณค่า P value โดยใช้ความน่าจะเป็นของการแจกแจงแบบทวินาม
- ▲ แบบ exact test ทำโดยการคำนวณค่าความน่าจะเป็นของการเกิดเหตุการณ์จากข้อมูลของตัวอย่างโดยตรง เช่น Fisher exact test ฯลฯ

2.2) *วิธีการอนุมานด้วยแนวคิดใหม่ Bayesian approach* วิธีการประมาณค่าและการทดสอบสมมติฐานด้วยวิธีการนี้จะใช้ prior probability ของการเกิดเหตุการณ์ที่ทราบมาก่อน พิจารณาร่วมกับความน่าจะเป็นของการ

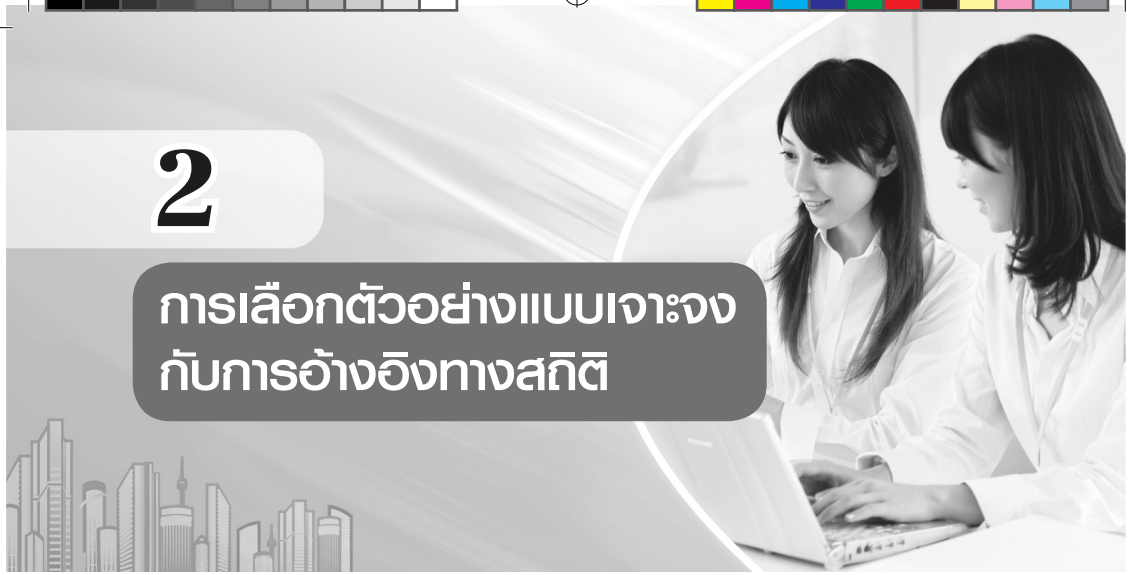
Using Statistics in Reseach 1_1 .indd 10 dummy 4 27/8/2557 11:48:24

เกิดเหตุการณ์ที่พบจากตัวอย่างที่ศึกษา Bayesian approach เริ่มมีบทบาทมากขึ้นในการวิเคราะห์ข้อมูลงานวิจัย โดยเฉพาะงานวิจัยทางคลินิก

สรุป

การเลือกตัวอย่างแบบไม่ใช้ความน่าจะเป็น นักวิจัยใช้สถิติพรรณนาลักษณะข้อมูลของกลุ่มตัวอย่างได้เท่านั้น โดยไม่สามารถใช้สถิติในการประมาณค่าหรือทดสอบสมมติฐาน ในการสรุปผลการศึกษาจะใช้ตรรกะในการอนุมานผล สำหรับการเลือกตัวอย่างแบบใช้ความน่าจะเป็นจะมีวิธีการทางสถิติแบบพาราเมตริก แบบไม่ใช้พาราเมตริก และแบบ Bayesian ในการประมาณค่าและทดสอบสมมติฐาน

Using Statistics in Reseach 1_1 .indd 11 dummy 4 30/6/2557 11:29:24



2

การเลือกตัวอย่างแบบเจาะจงกับการอ้างอิงทางสถิติ

การเลือกตัวอย่างแบบเจาะจง (purposive sampling) เป็นวิธีการเลือกตัวอย่างแบบไม่ใช้ความน่าจะเป็น (non-probability sampling) มีการนำมาใช้ในการเลือกตัวอย่างสำหรับงานวิจัยที่ต้องการผลเร็ว หรือในการวิจัยเชิงคุณภาพ ตัวอย่างที่ศึกษาไม่ได้สุ่มให้เป็นตัวแทนของประชากร ในการเลือกตัวอย่างแบบเจาะจง นักวิจัยจะต้องกำหนดลักษณะและคุณสมบัติของผู้ที่คาดว่าจะสามารถให้ข้อมูลที่ดีแล้วเลือกผู้ที่มีคุณสมบัติดังกล่าวมาเป็นตัวอย่าง เช่น ในการศึกษาพฤติกรรมวัยรุ่นเรื่องการดื่มสุรา นักวิจัยอาจตัดสินใจเลือกครูที่ปรึกษา ผู้จัดการผับร้านอาหาร หรือสถานบันเทิง หัวหน้านักเรียน หรือนักเรียนที่มีพฤติกรรมชอบเที่ยว มาเป็นตัวอย่างของการศึกษา เหตุที่ตัดสินใจเลือกเก็บข้อมูลจากบุคคลดังกล่าว เพราะคิดว่าเป็นผู้ที่ใกล้ชิดหรือเป็นผู้ให้บริการที่ทราบเรื่องเหล่านี้ดี เป็นข้อมูลของวัยรุ่นบางกลุ่มไม่ได้เป็นตัวแทนศึกษาทั้งหมด

การสรุปข้อมูลที่ได้จากการเลือกตัวอย่างโดยวิธีนี้จะขึ้นอยู่กับว่าตัวอย่างที่เลือกเป็นผู้ที่มีความรู้เรื่องเหล่านั้นมากหรือน้อย ถ้าสามารถ

เลือกผู้ที่รู้เรื่องดังกล่าวมากจะสามารถบรรยายลักษณะพฤติกรรมได้หลากหลาย แต่เนื่องจากไม่ได้เป็นข้อมูลที่เก็บจากวัยรุ่นแต่ละคนโดยตรง ทำให้ไม่สามารถคำนวณความถี่ของพฤติกรรมการดื่มสุราแต่ละแบบได้ อาจได้เพียงความเห็นว่าพบพฤติกรรมใดมากน้อยเป็นลำดับอย่างไร

การศึกษาที่เลือกตัวอย่างแบบเจาะจงจะไม่สามารถใช้สถิติอนุมานในการประมาณค่าหรือทดสอบสมมติฐานเป็นผลสรุปของประชากร (วัยรุ่นทั้งหมดในพื้นที่ศึกษา)

โดยสรุป ข้อมูลจากการเลือกตัวอย่างโดยวิธีนี้สามารถแสดงให้เห็นลักษณะพฤติกรรมของวัยรุ่นเรื่องการดื่มสุราว่ามีลักษณะพฤติกรรมอะไรบ้างเท่านั้น จะไม่สามารถสรุปในเชิงปริมาณ และไม่สามารถใช้วิธีการทางสถิติในการประมาณค่าหรือทดสอบสมมติฐาน

การเลือกแบบไม่ใช้ความน่าจะเป็นที่ใช้ในงานวิจัยเชิงสำรวจมากมี 2 แบบ คือ การเลือกแบบบังเอิญหรือการเลือกแบบสะดวก (accidental or convenience sampling) และการเลือกแบบโควตา (quota sampling)

การเลือกตัวอย่างแบบบังเอิญ จะเลือกตัวอย่างจากประชากรที่มีคุณสมบัติที่ต้องการ โดยผู้วิจัยจะกำหนดสถานที่และช่วงเวลาที่จะเก็บข้อมูล คนทุกคนที่มีคุณสมบัติตามที่นักวิจัยกำหนดมาที่อยู่ในสถานที่และในช่วงเวลาที่นักวิจัยกำหนดจะถูกเลือกเป็นตัวอย่าง เก็บข้อมูลไปเรื่อยๆ จนครบตามจำนวนคนที่ต้องการ เช่น ต้องการศึกษาเรื่องการดื่มสุราของผู้หญิงอายุ 15 ปีขึ้นไปที่อาศัยอยู่ในเขตเทศบาล โดยส่งผู้เก็บข้อมูลไปยืนดักถามสตรีที่เข้ามาในศูนย์การค้าที่ตั้งอยู่ในเขตเทศบาล การเก็บโดยวิธีนี้ทำได้ง่ายและเร็วกว่าการสุ่มครัวเรือนแล้วส่งคนไปถามตามบ้าน ผู้หญิงที่เข้ามาในศูนย์การค้าในช่วงเวลาเก็บข้อมูลย่อมไม่สามารถใช้เป็นตัวแทนผู้หญิงทั้งหมดในเขตเทศบาล เพราะผู้หญิงที่ไม่เคยมาศูนย์การค้าหรือมาในช่วงเวลาอื่นก็จะไม่มีโอกาสถูกเลือกเป็นตัวอย่าง

Using Statistics in Reseach 1_1 .indd 13 dummy 4 30/6/2557 11:29:24

การเลือกตัวอย่างแบบโควตา เป็นวิธีการแบ่งตัวอย่างที่จะเลือกมาศึกษาไปตามลักษณะกลุ่มย่อยของประชากร เพื่อให้มีความมั่นใจว่ามีตัวอย่างครอบคลุมทุกกลุ่มย่อยในประชากร เช่น จากตัวอย่างเรื่องการดื่มสุราของผู้หญิงอายุ 15 ปีขึ้นไป ถ้าคิดว่าอายุของผู้หญิงเป็นปัจจัยที่สำคัญ ข้อมูลที่ได้ควรครอบคลุมผู้หญิงทุกกลุ่มอายุ ถ้าเลือกตัวอย่างแบบบังเอิญอาจจะได้ตัวอย่างที่ไม่ครอบคลุมทุกกลุ่มอายุ ควรเปลี่ยนวิธีการเลือกตัวอย่างเป็นแบบโควตา โดยกำหนดจำนวน (โควตา) ตัวอย่างที่จะเก็บของแต่ละกลุ่มอายุ เมื่อกำหนดโควตาตามกลุ่มอายุแล้ว การเลือกตัวอย่างในแต่ละกลุ่มอายุยังใช้วิธีการเลือกแบบบังเอิญ ข้อมูลที่ได้จึงไม่ได้เป็นตัวแทนที่ดีของประชากร และไม่สามารถใช้สถิติในการประมาณค่าและทดสอบสมมติฐานได้

สรุป

งานวิจัยที่ใช้วิธีการเลือกตัวอย่างแบบเจาะจงหรือการเลือกตัวอย่างแบบไม่ใช้ความน่าจะเป็นแบบต่างๆ ข้อมูลที่ได้สามารถนำมาสรุปลักษณะของตัวอย่างที่ศึกษา นำไปบรรยายได้ว่าพบลักษณะอะไรบ้าง มีจำนวนมากน้อยเท่าไร สามารถนำข้อมูลดังกล่าวมาหาค่าเฉลี่ยและร้อยละได้ แต่เนื่องจากตัวอย่างไม่เป็นแบบเลือกจึงไม่สามารถใช้สถิติอนุมานในการประมาณค่าหรือทดสอบสมมติฐานของประชากร

นักวิจัยที่ใช้วิธีการเลือกตัวอย่างแบบไม่ใช้ความน่าจะเป็นควรตระหนักถึงข้อจำกัด และระวังในการสรุปผลการศึกษา

Using Statistics in Reseach 1_1 .indd 14 dummy 4 27/8/2557 11:49:56

3

ขนาดตัวอย่างกับจริยธรรมการวิจัยในมนุษย์

การทำวิจัยที่เกี่ยวข้องกับคน โครงร่างวิจัยจำเป็นต้องได้รับความเห็นชอบ (รับรอง) จากคณะกรรมการจริยธรรมการวิจัยในมนุษย์ ซึ่งเป็นแนวปฏิบัติที่เป็นมาตรฐานสากล งานวิจัยที่ไม่ได้รับการรับรองจะไม่ได้รับการตีพิมพ์เผยแพร่ผลงานในวารสารวิชาการนานาชาติ สำหรับประเทศไทยถูกกำหนดเป็นหลักเกณฑ์ที่ต้องปฏิบัติ ทั้งในพระราชบัญญัติสุขภาพแห่งชาติ สภาวิชาชีพ และในสถาบันการศึกษา รวมทั้งถูกกำหนดเป็นจรรยาบรรณนักวิจัยโดยสภาวิจัยแห่งชาติ

เกณฑ์ที่ใช้พิจารณารับรองจริยธรรมข้อหนึ่งคือ งานวิจัยต้องสามารถให้คำตอบที่เป็นประโยชน์ และวิธีดำเนินการวิจัยที่ใช้ต้องสามารถให้คำตอบที่ถูกต้องเชื่อถือได้ ดังนั้นการคำนวณขนาดตัวอย่างจะเป็นส่วนหนึ่งของเกณฑ์การพิจารณา

หลักจริยธรรมการวิจัยในมนุษย์ มีอยู่ 3 ประการ คือ

1) หลักเคารพศักดิ์ศรีของความเป็นมนุษย์
2) หลักประโยชน์
3) หลักความยุติธรรม

ในการพิจารณารับรองจริยธรรม การเลือกตัวอย่างมีส่วนเกี่ยวข้องกับหลักจริยธรรมการวิจัยในมนุษย์ดังนี้ ขนาดตัวอย่างที่ใช้ศึกษาจะเกี่ยวข้องกับหลักประโยชน์ และวิธีเลือกอาสาสมัคร (ผู้ที่ถูกเชิญเข้าร่วมให้ข้อมูลหรือเข้ารับการทดลอง) เข้ากลุ่มการทดลองจะเกี่ยวกับหลักความยุติธรรม ประเด็นความยุติธรรมจะพิจารณาจากวิธีการเลือกตัวอย่างอาสาสมัครเข้ากลุ่มทดลอง วิธีการที่ใช้เลือกต้องทำให้อาสาสมัครที่เข้าร่วมงานวิจัยทุกคนมีโอกาสได้รับประโยชน์/ความเสี่ยงเท่าเทียมกัน ตัวอย่างเช่น การทดลองเปรียบเทียบวิธีการสอนสุขศึกษา โดยกลุ่มทดลองสอนวิธีใหม่ และกลุ่มควบคุมสอนวิธีเดิม นักวิจัยเชื่อว่าสอนด้วยวิธีใหม่ต้องได้ผลดีกว่าวิธีการสอนแบบเดิม ดังนั้นในการเลือกให้อาสา-สมัครคนใดเข้าอยู่ในกลุ่มทดลองหรือกลุ่มควบคุมจะต้องมีความยุติธรรม โดยอาสาสมัครทุกคนต้องมีโอกาสอยู่ในกลุ่มทดลองหรืออยู่ในกลุ่มควบคุมเท่า ๆกัน วิธีการเลือกอาสาสมัครเข้ากลุ่มทดลองหรือกลุ่มควบคุมจึงต้องใช้วิธีการสุ่ม (random allocation) ในโครงร่างวิจัยจึงต้องอธิบายว่านักวิจัยใช้วิธีการอย่างไรในการสุ่มตัวอย่างให้แก่กลุ่มทดลองหรือกลุ่มควบคุม

การพิจารณาเรื่องของประโยชน์ในส่วนที่เกี่ยวข้องกับขนาดตัวอย่าง ถ้าขนาดตัวอย่างเล็กเกินไปคำตอบที่ได้จากงานวิจัยจะไม่น่าเชื่อถือ เมื่อผลงานวิจัยได้คำตอบที่ไม่น่าเชื่อถือจะถือว่างานวิจัยนั้นไม่มีประโยชน์ การเก็บข้อมูลจากอาสาสมัครแล้วไม่ได้ประโยชน์ถือว่าเป็นการไปรบกวนอาสาสมัคร ซึ่งเป็นการกระทำที่ไม่ชอบด้วยจริยธรรม การคำนวณขนาดตัวอย่างจึงมีความจำเป็นเพื่อใช้แสดงว่าขนาดตัวอย่างที่ใช้มีความเหมาะสม สามารถให้คำตอบที่เชื่อถือได้

ในกรณีที่นักวิจัยต้องการเพิ่มขนาดตัวอย่างให้มากๆ เช่น เป็นสองเท่าของที่คำนวณ จะทำได้หรือไม่ ประเด็นการพิจารณาจะอยู่ที่ว่าถ้าเก็บได้ครบตามจำนวนที่คำนวณได้ก็สามารถให้คำตอบที่เชื่อถือได้

Using Statistics in Reseach 1_1 .indd 16 dummy 4 30/6/2557 11:29:24

แล้ว จำนวนตัวอย่างที่เพิ่มขึ้นจึงไม่มีผลต่อการเปลี่ยนแปลงความน่าเชื่อถือของผลการวิจัย ส่วนที่เกินมาจึงเป็นการรบกวนอาสาสมัครโดยไม่มีประโยชน์อะไรเพิ่มขึ้น แต่ถ้ามีการเพิ่มขนาดตัวอย่างเผื่อการตกสำรวจ หรืออาสาสมัครออกระหว่างการทดลอง (drop out) สามารถทำได้

การพิจารณาความเหมาะสมของขนาดตัวอย่างจะพิจารณาจากความเหมาะสมของสูตรที่ใช้ และข้อมูลที่ใช้แทนค่าในสูตรซึ่งต้องมีการอ้างอิงแหล่งที่มาและเหตุผลในการเลือกใช้ค่าดังกล่าว สำหรับสูตรที่ใช้คำนวณขนาดตัวอย่างพิจารณาจาก

1) สูตรที่ใช้มีความสอดคล้องกับคำถามหรือวัตถุประสงค์หลัก ถ้าเป็นงานวิจัยเพื่อหาปัจจัยเสี่ยง วิธีการคำนวณขนาดตัวอย่างต้องใช้สูตรเพื่อหาปัจจัยเสี่ยง ไม่ใช่สูตรหาความชุก
2) สูตรที่ใช้ต้องเหมาะกับประเภทข้อมูลของตัววัดผลหลัก (main outcome) เช่น ข้อมูลกลุ่มต้องใช้สูตรคำนวณขนาดตัวอย่างที่เป็นค่าสัดส่วน และข้อมูลต่อเนื่องใช้สูตรคำนวณขนาดตัวอย่างที่เป็นค่าเฉลี่ย
3) วิธีการสุ่มตัวอย่างแบบต่างๆ เช่น สุ่มอย่างง่าย สุ่มแบบกลุ่ม สุ่มแบบชั้นภูมิ และเงื่อนไขการสุ่มตัวอย่าง เป็นอิสระหรือไม่เป็นอิสระต่อกัน จะใช้สูตรคำนวณขนาดตัวอย่างที่แตกต่างกัน

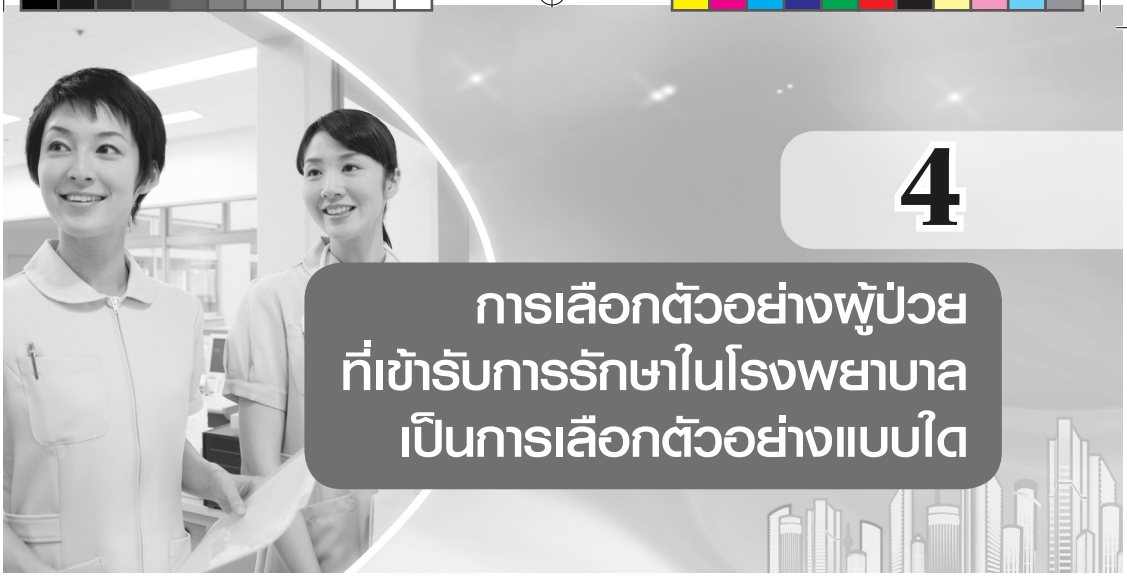
ดังนั้นในการอธิบายวิธีการคำนวณขนาดตัวอย่างในโครงร่างวิจัยต้องเลือกใช้สูตรการคำนวณขนาดตัวอย่างให้ถูกต้องสอดคล้องกับวัตถุประสงค์ ประเภทข้อมูล วิธีการสุ่ม และเงื่อนไขการสุ่มตัวอย่าง นอกจากนี้แล้วควรระบุแหล่งที่มาของค่าต่างๆที่ใช้ในการคำนวณด้วย

Using Statistics in Reseach 1_1 .indd 17 dummy 4 30/6/2557 11:29:25

สรุป

การสุ่มและการคำนวณขนาดตัวอย่างเกี่ยวข้องกับความน่าเชื่อถือของผลการวิจัย และยังมีผลต่อความเท่าเทียมของอาสาสมัครที่จะได้รับประโยชน์และความเสี่ยงจากการเข้าร่วมในโครงการวิจัย นักวิจัยต้องใช้วิธีการคำนวณขนาดตัวอย่างที่ถูกต้องสอดคล้องกับวัตถุประสงค์ ประเภทตัววัดผลหลัก วิธีการและเงื่อนไขการสุ่ม จึงจะเชื่อมั่นได้ว่างานวิจัยที่จะทำมีประโยชน์คุ้มกับการเสียเวลาของอาสาสมัครที่เข้าร่วมในงานวิจัย

Using Statistics in Reseach 1_1 .indd 18 dummy 4 30/6/2557 11:29:25



4

การเลือกตัวอย่างผู้ป่วยที่เข้ารับการรักษาในโรงพยาบาลเป็นการเลือกตัวอย่างแบบใด

มีงานวิจัยทางสุขภาพจำนวนมากที่ประชากรเป็นผู้ป่วยที่เข้ามารับการรักษาในโรงพยาบาล (hospital base) นักวิจัยกำหนดให้โรงพยาบาลเป็นสถานที่ศึกษา เลือกตัวอย่างจากผู้ป่วยที่มีคุณสมบัติตามเกณฑ์ที่เข้ามารับการรักษาในโรงพยาบาลตามระยะเวลาที่กำหนด

การเลือกตัวอย่างด้วยวิธีดังกล่าวมีบางรายงานวิจัยระบุว่าเป็นการเลือกตัวอย่างแบบเจาะจง บางรายงานวิจัยระบุว่าเป็นการเลือกตัวอย่างแบบสุ่ม ซึ่งข้อโต้แย้งเรื่องวิธีการเลือกตัวอย่างมีผลต่อการใช้สถิติในการประมาณค่าหรือการทดสอบสมมติฐานเพื่อสรุปผลการวิจัย เพราะการเลือกแบบเจาะจงจะไม่สามารถใช้สถิติดังกล่าวในการสรุปผล

ปกตินักวิจัยจะกำหนดให้ประชากรศึกษาคือผู้ป่วยที่มีคุณลักษณะตามเกณฑ์คัดเข้า/เกณฑ์คัดออก ที่จะเข้ามารับการรักษา ณ โรงพยาบาลที่ทำการวิจัย ซึ่งหมายถึงผู้ป่วยในอนาคตที่มีจำนวนอนันต์ ไม่สามารถนับจำนวนได้ครบถ้วน เพราะจะมีผู้ป่วยดังกล่าวเข้ามารับการรักษาทุกวันตลอดไป

การเลือกตัวอย่างจากประชากรลักษณะนี้จะเลือกโดยใช้กรอบเวลาเป็นหน่วยสุ่ม ผู้ป่วยแต่ละคนที่เข้ามารับการรักษาที่โรงพยาบาลใน

แต่ละวันจะเข้ามาแบบสุ่ม เพราะไม่มีใครสามารถกำหนดได้ว่าผู้ป่วยรายใดจะเข้ามารับการรักษาในวันใด การนำผู้ป่วยทั้งหมดในช่วงเวลาที่ศึกษามาเป็นตัวอย่างจึงถือว่าเป็นการเลือกตัวอย่างแบบสุ่ม เพื่อเป็นตัวแทนผู้ป่วยทั้งหมดที่จะเข้ามารับการรักษาในอนาคต

เพื่อให้ท่านผู้อ่านได้เข้าใจวิธีการสุ่มในลักษณะนี้ได้ชัดเจนยิ่งขึ้น จึงขอยกตัวอย่างวิธีการเก็บตัวอย่างน้ำจากลำธารที่มีการปนเปื้อนสารหนู ในกรณีที่พบว่าประชาชนในหมู่บ้านที่ใช้น้ำจากลำธารมีอาการป่วยด้วยโรคพิษจากสารหนู ทางสำนักงานสาธารณสุขจังหวัดจึงส่งนักวิชาการสาธารณสุขไปเก็บน้ำตัวอย่างจากลำธารมาตรวจหาสาเหตุ โดยสุ่มเก็บน้ำที่ไหลผ่านหมู่บ้าน ณ บริเวณที่ชาวบ้านตักน้ำไปใช้บริโภค น้ำตัวอย่างที่เก็บมาตรวจถ้าพบว่ามีสารหนู จะอนุมานว่าน้ำที่ไหลผ่านหน้าหมู่บ้านในอนาคตจะมีการปนเปื้อนสารหนูเช่นเดียวกันกับน้ำในช่วงเวลาที่สุ่มเก็บตัวอย่างไปตรวจ จึงแนะนำไม่ให้ชาวบ้านดื่มน้ำจากลำธารเพราะจะทำให้ป่วยด้วยโรคพิษจากสารหนู

ผู้ป่วยที่มีคุณลักษณะตามเกณฑ์คัดเข้า/เกณฑ์คัดออกเป็นการกำหนดลักษณะประชากร ไม่ใช่เกณฑ์ในการเลือกตัวอย่าง การพิจารณาว่าการเป็นตัวอย่างแบบสุ่มหรือไม่ จะพิจารณาจากวิธีการนำผู้ป่วยตามเกณฑ์คัดเข้า/เกณฑ์คัดออกเข้ามาเป็นตัวอย่าง ถ้าเป็นการเลือกโดยวิธีสุ่มจะทำได้ 3 แบบตามเงื่อนไขของปริมาณผู้ป่วยและข้อจำกัดของการทำวิจัย ดังนี้

1) **กรณีที่อาสาสมัคร (ผู้ที่ถูกเชิญเข้าร่วมให้ข้อมูลหรือเข้ารับการทดลอง) ที่มีคุณสมบัติตามเกณฑ์คัดเข้า/เกณฑ์คัดออกที่เข้ามารับการรักษาต่อวันไม่มาก** นักวิจัยสามารถเก็บข้อมูลได้ทุกราย ให้นำผู้ป่วยทุกรายที่มีคุณลักษณะตามเกณฑ์เข้ามาเป็นตัวอย่างในการศึกษาจนได้จำนวนครบตามที่กำหนด

Using Statistics in Reseach 1_1 .indd 20 dummy 4 30/6/2557 11:29:25

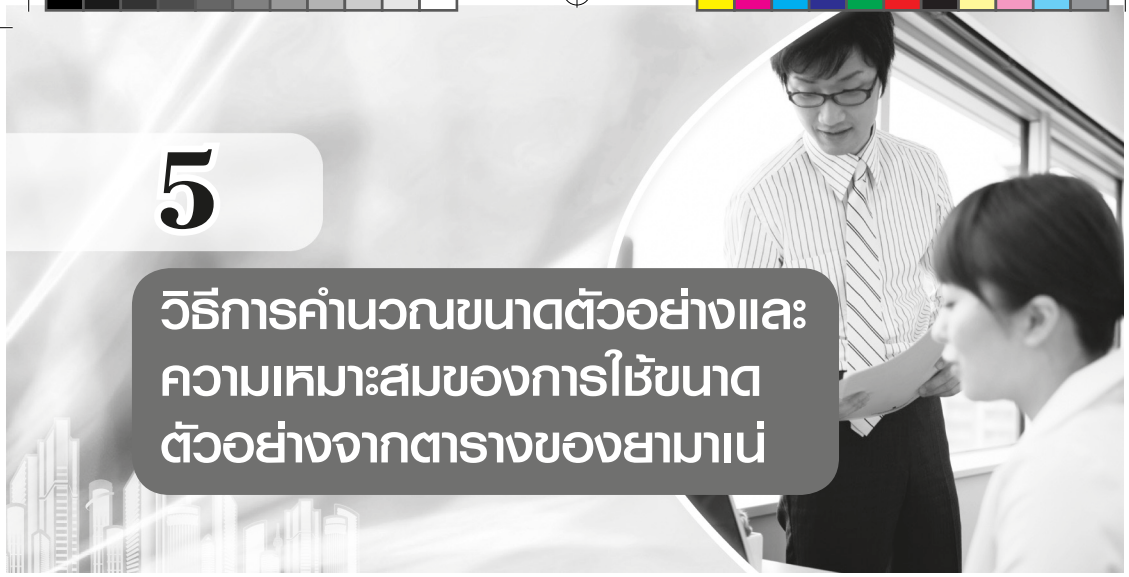
2) **กรณีที่อาสาสมัครที่มีคุณสมบัติตามเกณฑ์คัดเข้า/เกณฑ์คัดออกที่เข้ามารับการรักษาต่อวันมีจำนวนมาก** นักวิจัยไม่สามารถเก็บข้อมูลของผู้ป่วยได้ทั้งหมด (ทำไม่ทัน) นักวิจัยสามารถกำหนดวิธีการสุ่มตามจำนวนที่สามารถทำได้ต่อวัน เช่น ทำได้วันละ 10 รายก็สุ่มมาวันละ 10 ราย ทำเช่นนี้จนได้จำนวนครบตามที่กำหนด การจะกำหนดว่าควรเป็นวันละกี่รายไม่มีเกณฑ์กำหนด ขึ้นอยู่กับเวลาที่ใช้ในการเก็บข้อมูลของนักวิจัยหรือความครอบคลุมช่วงเวลา เช่น ต้องการที่จะให้มีผู้ป่วยกระจายเท่า ๆกันตลอดทั้งปีในกรณีที่การป่วยโรคนี้มีการเปลี่ยนแปลงตามเวลา (seasonal variation)

3) **กรณีที่เป็นการทดลองทางคลินิกที่มีการเปรียบเทียบระหว่างกลุ่ม** นักวิจัยต้องการให้แต่ละกลุ่มมีสัดส่วนผู้ป่วยที่มีอาการหนักเบาเท่า ๆกัน ในกรณีนี้จะต้องแยก (stratified) ผู้ป่วยตามอาการหนักเบาเป็นกลุ่มย่อย ๆ ในแต่ละกลุ่มย่อยจะมีการสุ่มกลุ่มทดลองให้แก่อาสาสมัครแต่ละคนที่นำเข้าในการศึกษา นำผู้ป่วยเข้าศึกษาไปเรื่อย ๆจนมีจำนวนผู้ป่วยทุกกลุ่มย่อยรวมกันครบตามจำนวนขนาดตัวอย่างที่กำหนด

สรุป

การเลือกตัวอย่างผู้ป่วยที่เข้ารับการรักษาในโรงพยาบาลในช่วงเวลาที่กำหนดเป็นการเลือกตัวอย่างแบบสุ่ม โดยมีวัตถุประสงค์ที่จะนำผลวิจัยที่ได้จากผู้ป่วยในช่วงเวลาที่ศึกษาไปอนุมานลักษณะและผลการรักษาพยาบาลผู้ป่วยที่จะเข้ามารับการรักษาในอนาคต

Using Statistics in Reseach 1_1 .indd 21 27/8/2557 11:49:56
dummy 4



5

วิธีการคำนวณขนาดตัวอย่างและความเหมาะสมของการใช้ขนาดตัวอย่างจากตารางของยามาเน่

การคำนวณขนาดตัวอย่างสำหรับงานวิจัยเป็นขั้นตอนที่สำคัญขั้นตอนหนึ่งของการทำวิจัย เพราะขนาดตัวอย่างที่เหมาะสมจะช่วยให้ผลการประมาณค่ามีความเที่ยงได้ตามระดับที่กำหนด หรือการทดสอบสมมติฐานมีอำนาจการทดสอบเพียงพอที่จะระบุความต่าง นอกจากนี้ขนาดตัวอย่างที่คำนวณได้ยังใช้พิจารณาความเป็นไปได้ของการทำวิจัย เช่น จำนวนอาสาสมัคร ค่าใช้จ่าย ระยะเวลาที่ใช้ดำเนินการ และความเป็นไปได้ในการบริหารจัดการ

การคำนวณขนาดตัวอย่างสำหรับงานวิจัยต้องอาศัยความรู้ในเรื่องระเบียบวิธีการวิจัย ความรู้ทางสถิติ และข้อมูลจากผลงานวิจัยที่ผ่านมา การเลือกสูตรที่จะใช้คำนวณขนาดตัวอย่างจะพิจารณาจากวัตถุประสงค์ของงานวิจัย แบบงานวิจัย สถิติที่ใช้ในการวิเคราะห์คำตอบตัวแปรผล และวิธีการสุ่มตัวอย่าง

1) **วัตถุประสงค์ของงานวิจัย** จะช่วยให้ทราบว่าจะใช้วิธีการประมาณค่าหรือทดสอบสมมติฐาน อนุมานข้อมูลจากตัวอย่างไปเป็นค่าสรุปของประชากร เช่น วัตถุประสงค์การศึกษาต้องการประมาณอัตราการเกิดโรค หรือต้องการเปรียบเทียบอัตราการ

เกิดโรคระหว่างผู้ที่ได้รับและผู้ที่ไม่ได้รับปัจจัยเสี่ยง สูตรการคำนวณขนาดตัวอย่างที่ใช้จะแตกต่างกัน

2) **แบบงานวิจัย** จะช่วยให้ทราบว่ามีประชากรศึกษากี่ประชากร วิธีการได้มาของตัวอย่างเป็นอิสระต่อกันหรือไม่ ซึ่งจะใช้สูตรการคำนวณขนาดตัวอย่างที่แตกต่างกัน

3) **สถิติที่ใช้ในการวิเคราะห์คำตอบตัวแปรผลหลัก** ใช้สถิติอะไร เช่น ค่าสัดส่วน RR OR ค่าเฉลี่ย ระยะเวลาการปลอดเหตุการณ์ (survival time) ค่าสัมประสิทธิ์สหสัมพันธ์ ฯลฯ สูตรคำนวณขนาดตัวอย่างจะแตกต่างไปตามสถิติที่ใช้ในการวิเคราะห์ตัวแปรผลหลัก

4) **วิธีการสุ่มตัวอย่างของหน่วยศึกษา** เช่น ใช้วิธีสุ่มอย่างง่าย การสุ่มแบบกลุ่ม ซึ่งการสุ่มแต่ละวิธีจะใช้สูตรคำนวณขนาดตัวอย่างที่ต่างกัน

การพิจารณาเลือกสูตรคำนวณขนาดตัวอย่างจะต้องนำข้อมูลทั้ง 4 ข้อมาร่วมพิจารณา หลังจากได้สูตรที่ต้องการแล้วจะนำข้อมูลที่ได้จากรายงานวิจัยที่ผ่านมากำหนดค่าต่าง ๆในสูตร

ตัวอย่าง ในการศึกษาเพื่อหาว่าประชาชนอำเภอแห่งหนึ่งมีการออกกำลังกายสม่ำเสมอร้อยละเท่าไร ถ้าจะทำการศึกษาโดยวิธีสำรวจจะต้องใช้ตัวอย่างขนาดเท่าไร

จากวัตถุประสงค์เพื่อหาว่าประชาชนอำเภอแห่งนี้มีการออกกำลังกายสม่ำเสมอร้อยละเท่าไร ในการศึกษาจะสุ่มตัวอย่างหาอัตราการออกกำลังกายในตัวอย่าง แล้วนำอัตราการออกกำลังกายของตัวอย่างไปประมาณค่าอัตราการออกกำลังกายของประชากร

Using Statistics in Reseach 1_1 .indd 23 dummy 4 30/6/2557 11:29:25

สถิติที่ใช้ในการวิเคราะห์คำตอบคือค่าสัดส่วนของการออกกำลังกายและอนุมานค่าสัดส่วนของการออกกำลังกายของประชากรด้วยสถิติ Z แบบงานวิจัยเป็นการศึกษาเชิงพรรณนาจึงมีการศึกษาในประชากรเดียว ถ้าใช้วิธีการสุ่มตัวอย่างแบบง่าย (simple random sampling) จะมีสูตรที่ใช้ในการคำนวณขนาดตัวอย่างดังนี้

$$n = \frac{Z^2_{\alpha/2}\, p(1-p)}{e^2} \qquad \text{.....❶}$$

โดยที่ n = ขนาดตัวอย่าง

α = ความผิดพลาดของการสุ่มตัวอย่าง

Z = confidence coefficient ได้จากความเชื่อมั่นที่กำหนด $(1 - \alpha)$

p = สัดส่วน (อัตราการออกกำลังกาย)

e = ความเที่ยงของการประมาณค่า (precision of estimation) หรือความคลาดเคลื่อนที่ยอมรับได้ (allowance error)

ถ้านำข้อมูลอัตราการออกกำลังกายสม่ำเสมอจากรายงานวิจัยที่ผ่านมาเป็นค่า p สำหรับแทนในสูตร และกำหนดค่า α และค่า e ตามที่ต้องการก็จะสามารถคำนวณขนาดตัวอย่างสำหรับการสำรวจดังกล่าวได้

สูตรคำนวณขนาดตัวอย่างสูตรที่ ❶ ใช้กับการสุ่มตัวอย่างแบบง่าย โดยที่ในการสุ่มแต่ละครั้งจะต้องใส่ตัวอย่างที่ถูกเลือกแล้วกลับคืน

Using Statistics in Reseach 1_1 .indd 24 dummy 4 30/6/2557 11:29:25

ในทางปฏิบัติเมื่อเลือกตัวอย่างใดได้แล้วนักวิจัยจะไม่ใส่คืน เพราะเกรงว่าคนคนเดียวอาจถูกสุ่มมาได้มากกว่า 1 ครั้ง ในการกระทำดังกล่าว ถ้าประชากรมีขนาดใหญ่จะไม่มีผลกระทบต่อขนาดของตัวอย่าง แต่จะมีผลกระทบต่อประชากรขนาดเล็ก การพิจารณาขนาดประชากรว่าใหญ่หรือเล็กจะพิจารณาจากค่า n ที่คำนวณได้จากสูตรที่ ❶ ถ้ามีค่าน้อยกว่าร้อยละ 5 ของประชากร แสดงว่าประชากรมีขนาดใหญ่ ในทางตรงกันข้าม ถ้าขนาดตัวอย่างที่คำนวณได้มากกว่าร้อยละ 5 ของประชากร แสดงว่าประชากรมีขนาดเล็ก จะต้องนำค่าขนาดตัวอย่างที่คำนวณได้มาปรับด้วยค่า finite correction factor $\sqrt{\frac{N-n}{N-1}}$ โดยมีสูตรคำนวณขนาดตัวอย่างที่ปรับด้วยค่า finite correction factor ดังนี้

$$n_0 = n\sqrt{\frac{N-n}{N-1}}$$

โดยที่ n = ขนาดตัวอย่างที่คำนวณโดยสูตรที่ ❶
N = จำนวนประชากร
n_0 = ขนาดตัวอย่างที่ปรับแล้ว

ในหนังสือสถิติหลายเล่มจะนำสูตรดังกล่าวมาจัดใหม่เพื่อความสะดวกในการคำนวณดังแสดงในสูตรที่ ❷

$$n_0 \frac{Z^2_{\alpha/2}\, Np(1-p)}{Z^2_{\alpha/2}\, p(1-p) + (N-1)e^2} \quad \text{.....❷}$$

โดยปกติในการคำนวณ ผู้คำนวณจะต้องแทนค่าต่าง ๆลงในสูตรและคำนวณด้วยเครื่องคิดเลขหรือเครื่องคอมพิวเตอร์ แต่เมื่อ 40 กว่าปีที่แล้วเครื่องคอมพิวเตอร์มีขนาดใหญ่และไม่ได้มีใช้อย่างแพร่หลาย คน

Using Statistics in Reseach 1_1 .indd 25 dummy 4 30/6/2557 11:29:26

ทั่วไปจึงไม่มีโอกาสได้ใช้ ดังนั้นเพื่ออำนวยความสะดวกในการคำนวณขนาดตัวอย่าง ผู้แต่งตำราสถิติบางท่านจึงได้คำนวณขนาดตัวอย่างโดยกำหนดค่าต่างๆในสูตร แล้วนำผลการคำนวณจัดทำเป็นตารางเพื่อให้นักวิจัยเลือกใช้ได้โดยไม่ต้องคำนวณเอง ทาโร ยามาเน่ (Taro Yamane) ได้ทำตารางคำนวณขนาดตัวอย่างสำหรับการประมาณค่าสัดส่วนของประชากรในกรณีประชากรขนาดเล็กตามสูตรที่ ❷ ใส่ลงในหนังสือ Statistics: An Introductory Analysis* หน้าที่ 1088 ตารางที่ 6 โดยกำหนดให้ $p = 0.5$ ค่า $\alpha = 0.05$ เมื่อนำค่าตัวเลขทั้งสองแทนลงในสูตร จัดทำเป็นสูตรใหม่ได้ดังนี้ $n = \frac{N}{1 + Ne^2}$ นำสูตรดังกล่าวมาสร้างเป็นตารางคำนวณขนาดตัวอย่างตามค่าต่างๆของ N และ e

ในประเทศไทยมีผู้ใช้ตารางคำนวณขนาดตัวอย่างของยามาเน่ไปใช้คำนวณขนาดตัวอย่างสำหรับงานวิจัยอย่างแพร่หลาย จากประสบการณ์ของผู้เขียนในฐานะผู้ทบทวนรายงานวิจัยและโครงร่างงานวิจัย พบว่ามีการใช้ที่ไม่เหมาะสมและไม่ถูกต้องจำนวนมาก ทั้งนี้ อาจมีสาเหตุมาจากการที่ผู้ใช้ไม่ทราบเงื่อนไขของการคำนวณตัวอย่างของตารางดังกล่าว ความผิดพลาดที่พบมีดังนี้

1) การคำนวณขนาดตัวอย่างในตารางยามาเน่ใช้ได้กับงานวิจัยเชิงพรรณนาที่ต้องการประมาณค่าสัดส่วนของประชากรเท่านั้น แต่มีนักวิจัยจำนวนมากที่วัตถุประสงค์งานวิจัยเป็นการเปรียบเทียบหรือหาความสัมพันธ์ แต่ใช้ตารางดังกล่าวคำนวณขนาดตัวอย่าง ซึ่งไม่ถูกต้อง
2) มีการใช้กับสถิติที่เป็นค่าเฉลี่ยหรือค่าสหสัมพันธ์ ซึ่งใช้ไม่ได้

---

* Taro Yamane. **Statistics: An Introductory Analysis.** 3rd ed. New York: Harper International edition, 1973.

Using Statistics in Reseach 1_1 .indd 26 dummy 4 27/8/2557 11:49:56

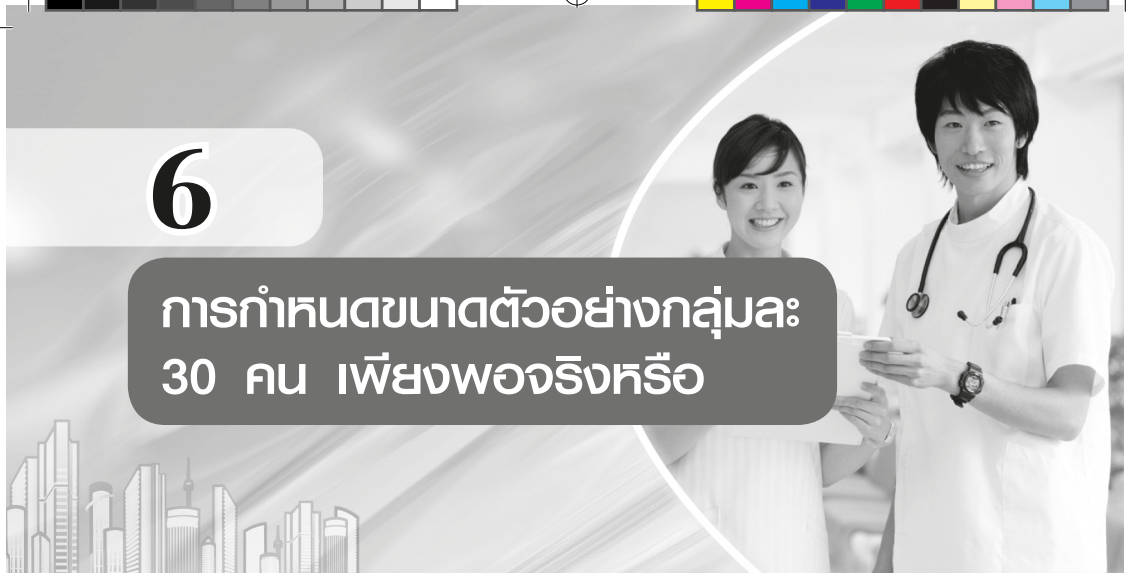
3) มีการใช้กับวิธีการสุ่มแบบกลุ่ม (cluster sampling) ซึ่งใช้ไม่ได้ จะใช้กับวิธีสุ่มแบบง่ายเท่านั้น
4) มีการใช้ระดับนัยสำคัญ (α) 0.05 ไปแทนเป็นค่า e ในตาราง ซึ่งเป็นค่าความเที่ยงของการประมาณค่า
5) ในการศึกษาที่ค่าสัดส่วนมีค่าน้อย เช่น อัตราการเป็นโรคขาดสารอาหารระดับ 2 และ 3 มีอยู่ประมาณร้อยละ 2 ($p = 0.02$) หรืออัตราการป่วยด้วยโรคหัวใจและหลอดเลือด 10 ต่อ 10,000 คน หรือร้อยละ 0.1 ($p = 0.001$) แต่ขนาดตัวอย่างจากตารางดังกล่าวกำหนดจากอัตราร้อยละ 50 ($p = 0.5$) ซึ่งต่างกันมาก ทำให้ได้ขนาดตัวอย่างไม่เหมาะสม

จะเห็นได้ว่า การใช้ตารางคำนวณขนาดตัวอย่างของยามาเน่ใช้ได้จำกัดตามเงื่อนไขของการสร้างตาราง จึงขอให้มีความระมัดระวังในการใช้ตารางดังกล่าว

สรุป

การคำนวณขนาดตัวอย่างเป็นเรื่องที่นักวิจัยควรศึกษาวิธีการคำนวณให้เข้าใจและสามารถคำนวณได้ด้วยตนเอง เพราะจะช่วยให้สามารถชี้แจงความเหมาะสมของตัวอย่างให้แก่ผู้ให้ทุนหรือคณะกรรมการจริยธรรม วิธีการคำนวณขนาดตัวอย่างสำหรับงานวิจัยมีอยู่ในหนังสือชีวสถิติ หนังสือระเบียบวิธีวิจัย และแหล่งความรู้บนอินเทอร์เน็ต นอกจากนี้ยังมีโปรแกรมคอมพิวเตอร์จำนวนมากให้นักวิจัยใช้ในการคำนวณ การนำเสนอบทความนี้เพื่อให้เป็นข้อระมัดระวังในการใช้ตารางคำนวณขนาดตัวอย่าง และเป็นแนวคิดให้หาวิธีการที่เหมาะสมในการคำนวณขนาดตัวอย่างสำหรับงานวิจัย

Using Statistics in Reseach 1_1 .indd 27 dummy 4 27/8/2557 11:49:56



6

การกำหนดขนาดตัวอย่างกลุ่มละ 30 คน เพียงพอจริงหรือ

ในการกำหนดขนาดตัวอย่าง มีคำแนะนำว่าขนาดตัวอย่างที่ศึกษาควรเท่ากับหรือมากกว่า 30 เช่น ในการกำหนดขนาดตัวอย่างของงานวิจัยเชิงทดลองที่กำหนดให้กลุ่มทดลองและกลุ่มควบคุมมีขนาดตัวอย่างกลุ่มละ 30 คน โดยเข้าใจว่ามีขนาดตัวอย่างตั้งแต่ 30 คนเป็นขนาดตัวอย่างที่เพียงพอ

จากทฤษฎี Central limit theorem ได้พิสูจน์ว่าถ้าขนาดตัวอย่างที่สุ่มมีตั้งแต่ 30 คนขึ้นไป การแจกแจงค่าสถิติของตัวอย่างจะมีการแจกแจงแบบปกติ ในกรณีตัวแปรผลเป็นข้อมูลต่อเนื่องขนาดตัวอย่าง 30 คนขึ้นไปจะทำให้การแจกแจงค่าเฉลี่ยของตัวอย่างมีการแจกแจงแบบปกติ จึงใช้การแจกแจงแบบปกติคำนวณความน่าจะเป็นในการทดสอบสมมติฐานหรือการประมาณค่าประชากร โดยไม่เกี่ยวข้องกับขนาดตัวอย่างที่ใช้ในการระบุความต่างในการทดสอบสมมติฐานของค่าเฉลี่ยหรือขนาดตัวอย่างที่ใช้ในการประมาณค่าเฉลี่ย

ในการคำนวณขนาดตัวอย่าง สูตรที่ใช้ในการคำนวณจะสร้างมาจากสูตรการประมาณค่าหรือสูตรการทดสอบสมมติฐาน ซึ่งส่วนใหญ่จะมีข้อตกลงเบื้องต้นว่าการแจกแจงค่าสถิติของตัวอย่างมีการแจกแจงแบบ

ปกติ ดังนั้นจึงใช้ค่าสถิติ Z ในการคำนวณความน่าจะเป็นของการเกิดความผิดพลาด α และ β ตัวอย่างเช่น ในการคำนวณขนาดตัวอย่างเพื่อเปรียบเทียบผลการสอนสุขศึกษาระหว่างกลุ่มทดลองและกลุ่มควบคุม สูตรการคำนวณขนาดตัวอย่างเพื่อเปรียบเทียบค่าเฉลี่ยของ 2 กลุ่มมีดังนี้

$$n = \frac{2(Z_{\alpha/2} + Z_{\beta})^2\sigma^2}{(\mu_1 - \mu_2)}$$

จากสูตรการคำนวณขนาดตัวอย่าง ข้อมูลที่ใช้ในการคำนวณมี 3 ค่า คือ

σ^2 = ความแปรปรวนร่วมของคะแนนความรู้ที่ได้จากการสอนสุขศึกษา (จากงานวิจัยที่ผ่านมา)

$(\mu_1 - \mu_2)$ = ผลต่างของคะแนนความรู้เฉลี่ยระหว่างกลุ่มทดลองและกลุ่มควบคุมที่คาดว่าจะพบจากการศึกษา

$(Z_{\alpha/2} + Z_{\beta})^2$ = ความน่าจะเป็นของการเกิดความผิดพลาด α และ β

จากสูตรดังกล่าวขนาดตัวอย่างจะแปรตามความแปรปรวนร่วมของคะแนนความรู้ ถ้ามีความแปรปรวนร่วมมากขนาดตัวอย่างจะใหญ่ (ตัวคูณใหญ่) ขนาดตัวอย่างยังแปรตามผลต่างของคะแนนความรู้เฉลี่ยระหว่างกลุ่มทดลองและกลุ่มควบคุม ถ้าค่าเฉลี่ยความรู้ต่างกันมาก (ตัวหารใหญ่) ขนาดตัวอย่างจะเล็ก และแปรตามความน่าจะเป็นของการเกิดความผิดพลาด α และ β

ตัวอย่างของการคำนวณขนาดตัวอย่างในการเปรียบเทียบคะแนนความรู้ระหว่างกลุ่มทดลองและกลุ่มควบคุม

Using Statistics in Reseach 1_1 .indd 29 dummy 4 30/6/2557 11:29:26

ตัวอย่างที่ 1 ถ้าความแปรปรวนร่วมเท่ากับ 5.8 โดยมี $\mu_1 = 60$, $\mu_2 = 55$, $\alpha = 0.05$ และ $\beta = 0.2$ จากสูตรข้างต้นจะคำนวณขนาดตัวอย่างได้ $(n) = 21$ ต่อกลุ่ม

ตัวอย่างที่ 2 ถ้าความแปรปรวนร่วมเท่ากับ 10.2 โดยมี $\mu_1 = 70$, $\mu_2 = 65$, $\alpha = 0.05$ และ $\beta = 0.2$ จะคำนวณขนาดตัวอย่างได้ $(n) = 65$ ต่อกลุ่ม

จากตัวอย่างที่ 1 และ 2 จะเห็นได้ว่าเมื่อความแปรปรวนมีค่ามากขึ้น (ตัวอย่างที่ 2) ขนาดตัวอย่างจะใหญ่ขึ้น และขนาดตัวอย่างที่คำนวณได้อาจจะมากหรือน้อยกว่า 30 ก็ได้

ดังนั้นการกำหนดให้ขนาดตัวอย่างในการศึกษาเท่ากับ 30 อาจจะบอกได้เพียงแต่ว่าการแจกแจงของค่าเฉลี่ยของตัวอย่างมีการแจกแจงแบบปกติเท่านั้น แต่ไม่สามารถระบุความพอเพียงของขนาดตัวอย่างที่จะทดสอบสมมติฐานหรือประมาณค่าได้

นอกจากนี้ถ้าเป็นการประมาณค่าหรือทดสอบสมมติฐานค่าสัดส่วน (ตัวแปรผลเป็นตัวแปรกลุ่มแบบมีสองค่า) ค่าสัดส่วนจะมีการแจกแจงแบบปกติ เมื่อ $np > 5$ และ $n(1 - p) > 5$ เมื่อนักวิจัยกำหนดค่า $n = 30$ ถ้า p มีค่าน้อยกว่า 0.17 จะทำให้การแจกแจงค่าสัดส่วนของตัวอย่างไม่แจกแจงแบบปกติ

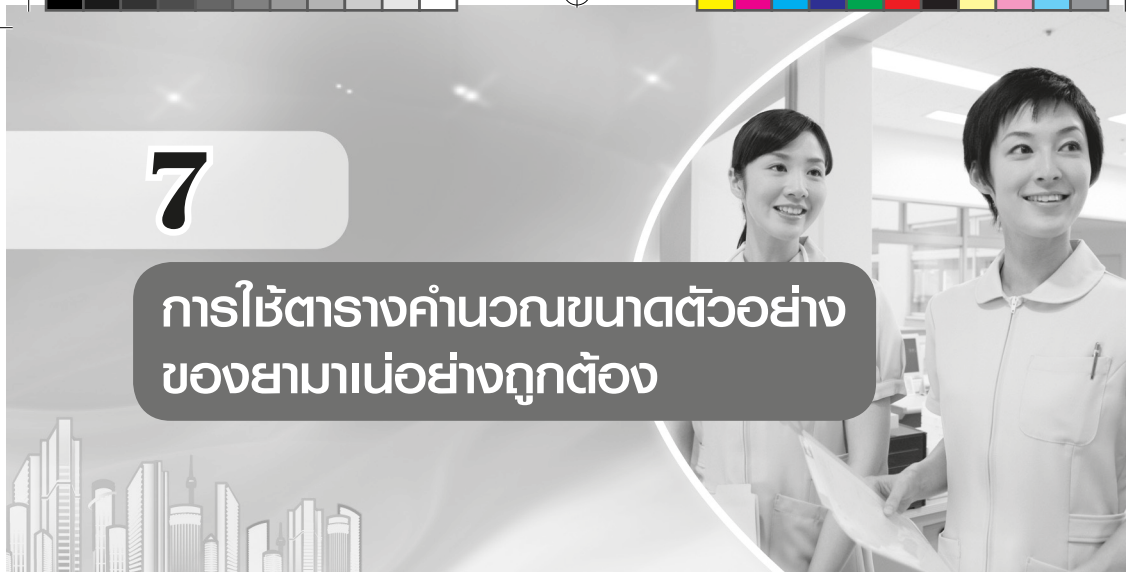
Using Statistics in Reseach 1_1 .indd 30 dummy 4 27/8/2557 11:49:56

สรุป

การกำหนดขนาดตัวอย่างเท่ากับ 30 สามารถบอกได้แต่เพียงว่าการแจกแจงของค่าเฉลี่ยของตัวอย่างมีการแจกแจงแบบปกติเท่านั้น ไม่สามารถใช้ระบุความเหมาะสมของขนาดตัวอย่างที่ใช้ในการประมาณค่าหรือทดสอบสมมติฐาน การกำหนดขนาดตัวอย่างเท่ากับ 30 ต่อกลุ่ม อาจจะมากเกินไปหรือไม่เพียงพอสำหรับการศึกษาก็ได้ ขึ้นอยู่กับความแปรปรวนและผลต่างของค่าเฉลี่ยระหว่างกลุ่ม

ในกรณีที่เป็นการทดสอบสมมติฐานหรือประมาณค่าสัดส่วน จะใช้ขนาดตัวอย่างที่ใหญ่กว่าค่าเฉลี่ย การกำหนดขนาดตัวอย่าง 30 คน ยิ่งมีโอกาสได้ขนาดตัวอย่างที่ไม่เหมาะสมสำหรับการศึกษา

Using Statistics in Reseach 1_1 .indd 31 dummy 4 30/6/2557 11:29:26



7

การใช้ตารางคำนวณขนาดตัวอย่างของยามาเน่อย่างถูกต้อง

ผู้เขียนได้เคยกล่าวถึงความไม่เหมาะสมหรือความไม่ถูกต้องของการใช้ตารางคำนวณขนาดตัวอย่างของยามาเน่ไปแล้ว ครั้งนี้จะอธิบายว่า หากจะใช้ให้ถูกต้องควรจะใช้กับงานวิจัยอะไร มีเงื่อนไขในการกำหนดค่าในการเปิดตารางอย่างไร

ยามาเน่ได้ทำตารางคำนวณขนาดตัวอย่างสำหรับการประมาณค่าสัดส่วนในประชากรขนาดเล็ก (ทราบขนาดของประชากร) โดยใช้สูตรคำนวณขนาดตัวอย่างสูตรที่ ❶

$$n_0 \frac{Z^2_{\alpha/2}\, Np(1-p)}{Z^2_{\alpha/2}\, p(1-p) + (N-1)e^2} \qquad \text{.....❶}$$

จากสูตรที่ ❶ ยามาเน่ได้กำหนดค่า $p = 0.5$ ค่า $\alpha = 0.05$ (ได้ค่า $Z^2_{\alpha/2} = 1.96^2 = 2^2$) เมื่อนำค่า p และ $Z^2_{\alpha/2}$ แทนค่าลงในสูตรที่ ❶ จะได้สูตรใหม่ ❷ โดยเหลือตัวแปรที่ใช้คำนวณ 2 ตัว คือ N และ e

$$n = \frac{N}{1 + Ne^2} \qquad \text{.....❷}$$

แล้วจึงนำสูตรที่ ❷ มาสร้างเป็นตารางคำนวณขนาดตัวอย่างตามค่าต่างๆ ของ N และ e

การใช้ตารางยามาเน่ช่วยคำนวณขนาดตัวอย่างได้ถูกต้อง นักวิจัยจำเป็นต้องทราบที่มาของสูตรที่ใช้ในการคำนวณ แบบงานวิจัย และตัวแปรผล ดังนี้

1) **ขนาดตัวอย่าง** ในตารางยามาเน่คำนวณมาจากสูตรการประมาณค่าสัดส่วนของประชากร จึงใช้สำหรับงานวิจัยเชิงพรรณนาที่คำถามหลักต้องการประมาณค่าสัดส่วนในประชากร และขนาดตัวอย่างในตารางคำนวณมาจากค่าสัดส่วน 0.5 ($p = 0.5$) ดังนั้นการนำตารางไปใช้หาขนาดตัวอย่างสำหรับประมาณค่าสัดส่วนที่มีค่าน้อยกว่า 0.5 จะทำให้ได้ขนาดตัวอย่างที่ไม่ถูกต้อง โดยเฉพาะในการศึกษาโรคที่มีความชุกต่ำค่าขนาดตัวอย่างที่คำนวณได้จากตารางจะน้อยกว่าขนาดตัวอย่างที่คำนวณจากสูตรมาก
2) **ตัวแปรผลที่ใช้ตอบคำถามหลัก** จะต้องเป็นตัวแปรกลุ่มแบบมี 2 ค่า (dichotomous outcome) เช่น หาย/ไม่หาย เห็นด้วย/ไม่เห็นด้วย ถ้าเป็นตัวแปรผลที่มีกลุ่มคำตอบมากกว่า 2 กลุ่ม หรือตัวแปรต่อเนื่องจะต้องมีการแปลงข้อมูลจัดให้เป็น 2 กลุ่ม เช่น ชอบ/ไม่ชอบ ผ่าน ($\geq$ 60%)/ไม่ผ่าน ($<$ 60%)
3) **ค่าสถิติที่จะใช้เป็นตัวกะประมาณ** (estimator) ต้องเป็นค่าสัดส่วน (p) เท่านั้น จะไปใช้กับค่าเฉลี่ย ($\bar{x}$) หรือค่าสหสัมพันธ์ (r) ไม่ได้
4) **N เป็นขนาดของประชากรที่ศึกษา** (study population) เช่น ต้องการหาความชุกของการเป็นโรคความดันโลหิตสูงของผู้มีอายุ 40 ปีขึ้นไปของตำบลศาลา ค่า N คือประชากรอายุ 40 ปีขึ้นไปทุกคนที่อาศัยอยู่ในตำบลศาลา

Using Statistics in Reseach 1_1 .indd 33 dummy 4 30/6/2557 11:29:26

5) **e คือค่าความเที่ยงของการประมาณค่า** หมายถึงความต่างระหว่างค่าสัดส่วนของประชากรที่นักวิจัยประมาณจากข้อมูลตัวอย่างกับค่าสัดส่วนของประชากร (ได้จากการศึกษาที่ผ่านมา) นักวิจัยเป็นผู้กำหนดค่าความเที่ยงโดยพิจารณาว่าต้องการให้ผลการศึกษามีค่าใกล้เคียงค่าสัดส่วนของประชากรมากน้อยเท่าใด

การกำหนดค่า e

ตัวอย่างที่ 1 จากการศึกษาที่ผ่านมาพบว่าค่าสัดส่วนความชุกของการเป็นโรคหนอนพยาธิเท่ากับ 0.32 (ร้อยละ 32) ในการสำรวจที่ตำบลศาลา ถ้ากำหนดให้ค่าความเที่ยงเท่ากับ 4% (e = 0.04) หมายความว่า ถ้าความชุกของประชากรเท่ากับร้อยละ 32 ผลการศึกษาจากขนาดตัวอย่างที่กำหนดให้ e = 0.04 อาจพบความชุกที่ร้อยละ 28 (32 – 4) หรืออาจมีค่ามากถึง 36 (32 + 4)

ตัวอย่างที่ 2 จากการศึกษาที่ผ่านมาพบว่าค่าสัดส่วนความชุกของการป่วยด้วยโรคหัวใจและหลอดเลือดเท่ากับ 10 ต่อ 10,000 คน (p = 0.001) ในการศึกษากำหนดให้ค่าความเที่ยงเท่ากับ 0.05 (e = 0.05) ค่าประมาณความชุกจากการศึกษาเป็น –0.049 (0.001 – 0.05) หรือ 0.051 (0.001 + 0.05) การที่ได้ค่าประมาณของ p เป็นลบเพราะกำหนดค่า e ให้มีขนาดใหญ่กว่าความชุก (p) แต่ค่าความชุกต่ำสุดคือ 0 มีค่าติดลบไม่ได้ ดังนั้นการให้ค่าความเที่ยงมากกว่าความชุกจึงไม่ถูกต้อง

Using Statistics in Reseach 1_1 .indd 34 dummy 4 30/6/2557 11:29:27

ในการกำหนดค่าความเที่ยงนักวิจัยต้องทราบว่าตัวแปรใดเป็นตัวแปรผล และค่าสัดส่วนตัวแปรผลที่เคยพบจากงานวิจัยที่ผ่านมามีค่าเท่าไร จึงจะสามารถกำหนดค่าความเที่ยงได้เหมาะสม

การใช้ตารางยามาเน่กำหนดขนาดตัวอย่าง

ตัวอย่างที่ 3 ผู้วิจัยสนใจว่าครัวเรือนที่มีฝาปิดภาชนะเก็บน้ำป้องกันยุงลายครบถ้วน ในตำบลศาลาซึ่งมี 1,500 หลังคาเรือน มีอยู่ร้อยละเท่าไร จากการสำรวจของจังหวัดพบว่าครัวเรือนที่มีฝาปิดภาชนะเก็บน้ำป้องกันยุงลายครบถ้วนในพื้นที่อื่นมีอยู่ร้อยละ 50 ($p = 0.50$) จึงมีเงื่อนไขสอดคล้องกับการคำนวณขนาดตัวอย่างด้วยตารางยามาเน่ ผู้วิจัยจึงกำหนดให้ค่าความเที่ยงเท่ากับร้อยละ 4 ($e = 0.04$) เปิดตารางยามาเน่ที่ $N = 1,500$ และ $e = 0.04$ ได้ขนาดตัวอย่างในการศึกษาเท่ากับ 441 ครัวเรือน

ตารางคำนวณขนาดตัวอย่างของยามาเน่สามารถใช้คำนวณขนาดตัวอย่างในงานวิจัยที่มีวัตถุประสงค์หลักเพื่อการประมาณค่าสัดส่วนของประชากร และการใช้ต้องเป็นไปตามเงื่อนไขข้อ 1) ถึง 5) ที่ได้กล่าวไปแล้ว ปัจจุบันโปรแกรมคำนวณขนาดตัวอย่างมีให้ใช้แพร่หลายและฟรี การคำนวณจากสูตรโดยตรงจะได้ขนาดตัวอย่างที่เหมาะสมสำหรับการตอบคำถามงานวิจัย

Using Statistics in Reseach 1_1 .indd 35 dummy 4 30/6/2557 11:29:27

8

บทบาทของตัวแปรผลในการคำนวณขนาดตัวอย่าง

ในการทำวิจัย นักวิจัยต้องมีความมั่นใจว่าจำนวนตัวอย่างที่ศึกษาเพียงพอที่จะให้คำตอบที่เชื่อถือได้ กรณีที่เป็นบทความวิจัยขนาดตัวอย่างจะช่วยให้ผู้อ่านใช้พิจารณาว่าผลการศึกษามีความน่าเชื่อถือหรือไม่

ดังนั้นการนำเสนอขนาดตัวอย่างที่ใช้โดยระบุเพียงแต่สูตรที่ใช้และขนาดตัวอย่างที่ได้ย่อมไม่เพียงพอต่อการพิจารณาความเหมาะสมของขนาดตัวอย่างที่ใช้ศึกษา นักวิจัยต้องระบุค่าตัวแปรที่ใช้แทนค่าในสูตรด้วย ในบทความนี้จะอธิบายความสำคัญของตัวแปรผลที่ต้องใช้ในการคำนวณขนาดตัวอย่างสำหรับการประมาณค่า $n = \frac{Z^2_{\alpha/2}\, p(1-p)}{e^2}$

การคำนวณขนาดตัวอย่างเพื่อการประมาณค่าสัดส่วน ค่า p เป็นตัวแปรหนึ่งที่ต้องใช้แทนค่าในสูตร ค่า p เป็นค่าสัดส่วนของเหตุการณ์ที่สนใจ ซึ่งหมายถึงตัวแปรผลที่ใช้ตอบวัตถุประสงค์หลัก เช่น ในการศึกษาเรื่องของผู้สูงอายุที่เดินด้วยตนเองไม่ได้ โดยมีวัตถุประสงค์หลักเพื่อหาอัตราผู้สูงอายุที่เดินด้วยตนเองไม่ได้ และมีวัตถุประสงค์รองเพื่อหาอัตราผู้สูงอายุที่เดินด้วยตนเองไม่ได้จำแนกตามเพศ ในกรณีนี้ ค่า p คือค่าสัดส่วนของผู้สูงอายุที่เดินด้วยตนเองไม่ได้ โดยปกติในแต่ละงานวิจัยจะคำนวณขนาดตัวอย่างจากตัวแปรผลหลักเท่านั้น เพราะขนาดตัวอย่างที่

ใช้ศึกษาอย่างน้อยต้องเพียงพอที่จะตอบวัตถุประสงค์หลักได้ ในกรณีที่ผู้วิจัยต้องการตอบคำถามรองที่เชื่อถือได้จะต้องคำนวณขนาดตัวอย่างเพื่อตอบวัตถุประสงค์รองด้วย โดยใช้ค่า p ของตัวแปรผลรองของแต่ละวัตถุประสงค์มาคำนวณขนาดตัวอย่าง เช่น คำนวณขนาดตัวอย่างจากค่า p (ผู้สูงอายุที่เดินด้วยตนเองไม่ได้) ของชายและของหญิง เมื่อได้ขนาดตัวอย่างที่คำนวณจากตัวแปรผลทั้งหมดแล้ว จึงเลือกขนาดตัวอย่างที่ใหญ่ที่สุดสำหรับการศึกษา จึงจะทำให้งานวิจัยมีขนาดตัวอย่างเพียงพอสำหรับคำตอบทุกวัตถุประสงค์

ในโครงร่างวิจัยและรายงานผลวิจัย ความไม่ถูกต้องที่พบคือการไม่ระบุว่าคำนวณขนาดตัวอย่างด้วยตัวแปรผลใด โดยเฉพาะผู้ที่ใช้ตารางยามาเน่และตารางของ Krejcie and Morgan ในการกำหนดขนาดตัวอย่าง ซึ่งค่าขนาดตัวอย่างในตารางคำนวณมาจากค่า p = 0.5 ถ้าผู้วิจัยไม่เข้าใจว่าต้องใช้ค่า p ของตัวแปรผลหลักในการคำนวณขนาดตัวอย่าง ก็จะไม่ทราบว่าการกำหนดขนาดตัวอย่างจากตารางเหมาะสมกับงานวิจัยของตนหรือไม่ เช่น ในการวิจัยเพื่อหาความชุกของผู้ป่วยโรคความดันโลหิตสูง จากรายงานความชุกของโรคความดันโลหิตสูงเท่ากับ 14/1,000 คน ถ้าไปเปิดตารางยามาเน่ N = 50,000 คน และค่า e = 5% ขนาดตัวอย่างในตารางเท่ากับ 397 คน ถ้าใช้ตัวอย่าง 400 คนในการศึกษาอาจพบผู้เป็นโรคความดันโลหิตสูงประมาณ 3–4 คน ซึ่งขนาดตัวอย่างที่ได้ไม่พอที่จะตอบคำถาม ทั้งนี้ เพราะค่า p ของความชุกของโรคความดันโลหิตสูงจากรายงานเท่ากับ 0.014 (14/1,000) ซึ่งต่างจากค่า p = 0.5 ที่ใช้คำนวณขนาดตัวอย่างในตารางมาก ในกรณีที่ความชุกยิ่งน้อยขนาดตัวอย่างที่ใช้ศึกษายิ่งต้องมีขนาดใหญ่ ดังนั้นถ้านักวิจัยจะใช้ตารางในการกำหนดขนาดตัวอย่าง จะต้องพิจารณาเลือกใช้ตารางที่สร้างจากค่า p ที่ตรงหรือใกล้เคียงกับขนาดของค่า p ของตัวแปรผลหลักที่จะศึกษา จึงจะได้ขนาดตัวอย่างที่ถูกต้อง

Using Statistics in Reseach 1_1 .indd 37 dummy 4 30/6/2557 11:29:27

ความชุก	p	e (30% p)	α	n
10/100 คน	0.1	0.03	0.05	385
1/100 คน	0.01	0.003	0.05	4,226
1/1,000 คน	0.001	0.0003	0.05	42,642

จากตาราง ถ้ากำหนดค่าความกระชับในการประมาณค่า (e) ให้คงที่เท่ากับ 30% ของค่า p และค่า α เท่ากับ 0.05 จะพบว่า ขนาดตัวอย่าง (n) เพิ่มขึ้นเมื่อค่าสัดส่วนหรือความชุกมีค่าน้อยลง ยิ่งความชุกต่ำๆยิ่งต้องใช้ตัวอย่างขนาดใหญ่ ปัจจุบันโรคต่างๆมีความชุกค่อนข้างต่ำ การจะใช้ตารางขนาดตัวอย่างที่คำนวณจากค่า p (0.5) เพียงค่าเดียวจะไม่ได้ขนาดตัวอย่างที่ถูกต้องสำหรับงานวิจัย

กรณีประมาณค่าเฉลี่ยก็ต้องใช้ส่วนเบี่ยงเบนมาตรฐานของตัวแปรผลที่ใช้ตอบวัตถุประสงค์หลักในการคำนวณขนาดตัวอย่างเช่นกัน

สรุป

ในการคำนวณขนาดตัวอย่างเพื่อการประมาณค่าสัดส่วน ค่า p ที่ใช้แทนค่าในสูตรการคำนวณจะเป็นค่าสัดส่วนของตัวแปรผลที่ใช้ตอบวัตถุประสงค์หลัก หรืออาจคำนวณจากตัวแปรผลที่ใช้ตอบวัตถุประสงค์รอง ถ้าต้องการตอบคำถามรองที่เชื่อถือได้ด้วย

ขนาดตัวอย่างจะแปรไปตามขนาดของค่าสัดส่วนของตัวแปรผล ค่าสัดส่วนของตัวแปรผลที่มีค่าน้อย (ความชุกต่ำ) จะคำนวณได้ตัวอย่างขนาดใหญ่ การทบทวนวรรณกรรมเพื่อให้ได้ค่า p ที่เหมาะกับบริบทที่ศึกษาจะช่วยให้ได้ค่า p ที่เหมาะสมสำหรับการคำนวณขนาดตัวอย่าง ซึ่งจะทำให้งานวิจัยได้คำตอบที่เชื่อถือได้

Using Statistics in Reseach 1_1 .indd 38 dummy 4 30/6/2557 11:29:27

9

การเลือกค่าพารามิเตอร์สำหรับการคำนวณขนาดตัวอย่างเพื่อการประมาณค่าสัดส่วน

การคำนวณขนาดตัวอย่างเพื่อการประมาณค่ามีการใช้มากในการสำรวจและการวิจัยเพื่อประเมินสถานการณ์ทางด้านสาธารณสุข ในการคำนวณขนาดตัวอย่าง สิ่งที่นักวิจัยมีคำถามหรือไม่มั่นใจมากที่สุด คือ เรื่องการเลือกค่าพารามิเตอร์ที่ใช้แทนค่าในสูตรการคำนวณ

ในการคำนวณขนาดตัวอย่างเพื่อการประมาณค่าสัดส่วน มีสูตรที่ใช้คำนวณขนาดตัวอย่างดังนี้

$$n = \frac{Z^2_{\alpha/2}\, p(1 - p)}{e^2}$$

จากสูตร ค่าพารามิเตอร์ที่นักวิจัยต้องแทนลงในสูตรมีอยู่ 3 ค่าด้วยกัน คือ

- α = ค่าความผิดพลาดของการสุ่มตัวอย่าง นำไปคำนวณค่า $Z^2_{\alpha/2}$
- p = ค่าสัดส่วนของตัวแปรผล เช่น อัตราความครอบคลุมของบริการ อัตราการเกิดโรค หรืออัตราการหาย ฯลฯ
- e = ค่าความเที่ยงของการประมาณค่า (precision of estimation) หรือความคลาดเคลื่อนที่ยอมให้เกิด (allowance error)

โดยมีหลักในการเลือกค่า α, p และ e สำหรับแทนค่าลงในสูตรดังนี้

α เป็นค่าความน่าจะเป็นของความผิดพลาดที่เกิดจากการสุ่มตัวอย่าง เพื่อสรุปลักษณะของประชากร ความผิดพลาดดังกล่าวนักวิจัยสามารถกำหนดได้ว่าจะให้มีโอกาสเกิดมากน้อยเท่าใด นักวิจัยส่วนใหญ่นิยมกำหนดให้ α มีค่าเท่ากับ 0.05 ซึ่งเป็นค่าที่ทุกคนรู้สึกว่าเป็นขนาดความผิดพลาดที่ไม่มากเกินไป ในกรณีที่ต้องการให้มีโอกาสเกิดความผิดพลาดน้อยๆ นิยมกำหนดให้ α มีค่าเท่ากับ 0.01

p เป็นค่าสัดส่วนของตัวแปรผลหลัก (ใช้ตอบวัตถุประสงค์หลักของงานวิจัย) นำมาจากรายงานวิจัยที่ผ่านมา เช่น ในการหาอุบัติการณ์ของโรคพยาธิใบไม้ตับ นักวิจัยจะค้นหาอัตราการเกิดโรคพยาธิใบไม้ตับจากรายงานวิจัยหรือรายงานการสำรวจ อัตราที่ได้จะเป็นจำนวนคนที่เป็นโรค ต่อ 100 หรือต่อ 1,000 คน เช่น 8/1,000 คน เปลี่ยนอัตราที่ได้ให้เป็นค่าสัดส่วนโดยคำนวณเป็นค่าต่อประชากร 1 คน ซึ่งคำนวณค่า p ได้เท่ากับ 0.008

ในกรณีค้นพบว่าในเรื่องที่ศึกษามีหลายรายงานวิจัย หรือมีเฉพาะในรายงานวิจัยในต่างประเทศ จะมีหลักเกณฑ์ในการเลือกหรือปรับค่า p ที่ได้จากรายงานก่อนนำไปแทนค่าในสูตรดังนี้

1) **ในกรณีที่มีหลายรายงานวิจัย** เช่น ในการศึกษาอัตราการสวมใส่หมวกนิรภัยในการขับขี่รถจักรยานยนต์ พบว่ามีรายงานวิจัยที่ทำในจังหวัดขอนแก่น จังหวัดนครสวรรค์ และจังหวัดกาญจนบุรี ผู้วิจัยต้องหาอัตราการสวมใส่หมวกนิรภัยของจังหวัดพระนครศรี-อยุธยา จะเลือกใช้ค่า p ของรายงานวิจัยใด ในการพิจารณา

Using Statistics in Reseach 1_1 .indd 40 dummy 4 30/6/2557 11:29:27

นักวิจัยควรเริ่มต้นด้วยการพิจารณาคุณภาพของวิธีการศึกษา ถ้ารายงานทั้ง 3 เรื่องมีวิธีการศึกษาที่เชื่อถือได้ ผลงานวิจัยก็น่าเชื่อถือ นำไปใช้ได้ หลังจากนั้นจะพิจารณาว่าจังหวัดใดที่มีปัจจัยที่มีผลต่อการสวมใส่หมวกนิรภัยใกล้เคียงกับจังหวัดพระนครศรี-อยุธยามากที่สุด (เช่น สภาพถนน สภาพการจราจร ความเข้มงวดในการบังคับใช้กฎหมายจราจร) ให้ใช้ค่า p ของรายงานวิจัยฉบับนั้น

2) **ในกรณีที่มีเฉพาะรายงานวิจัยที่ทำในต่างประเทศ** เช่น อัตราการติดเชื้อจากการพักรักษาตัวในโรงพยาบาล มีรายงานวิจัยที่ทำในโรงพยาบาลในประเทศอังกฤษ สิงคโปร์ และสหรัฐอเมริกา ในกรณีนี้จะใช้หลักการเดียวกันกับกรณีในข้อ 1) คือเลือกรายงานวิจัยที่มีคุณภาพ และปัจจัยที่มีผลต่อการติดเชื้อใกล้เคียงกับสถานการณ์ของสถานที่ที่จะทำวิจัย

ถ้าผู้วิจัยประเมินแล้วยังพบว่าปัจจัยที่มีผลต่อการติดเชื้อจากรายงานวิจัยที่เลือกมาใกล้เคียงที่สุดแล้วแต่ก็ยังดีกว่าของโรงพยาบาลที่จะทำวิจัย ผู้วิจัยสามารถปรับค่า p จากรายงานก่อนนำไปแทนค่า เช่น พบว่าอัตราการติดเชื้อจากรายงานมี 12% นักวิจัยพิจารณาจากปัจจัยที่มีผลต่อการติดเชื้อแล้วคาดว่า ที่ดีน้อยกว่าน่าจะมีผลทำให้อัตราการติดเชื้อมากกว่าอีก 20% (ของ 12) ดังนั้นจึงปรับค่า p เพิ่มอีก 20% เป็น 14.4% นำค่า p = 0.144 ไปแทนค่าในสูตร

3) **ในกรณีรายงานวิจัยที่มีอยู่เป็นของหลายปีมาแล้ว** ในปัจจุบันสภาพการณ์ต่างๆเปลี่ยนไป เช่น ในการศึกษาเรื่องอัตราการคาดเข็มขัดนิรภัย นักวิจัยสามารถนำข้อมูลที่ศึกษาเมื่อ 7 ปีที่แล้วมาเป็นค่าเริ่มต้นแล้วประมาณค่าเพิ่มขึ้นหรือลดลงจากการ

Using Statistics in Reseach 1_1 .indd 41 dummy 4 30/6/2557 11:29:28

ให้บริการที่ดีขึ้น เช่น อัตราการคาดเข็มขัดนิรภัย มีผู้ที่ทำการศึกษาไว้เมื่อ 7 ปีที่แล้วมีอยู่ร้อยละ 30 ในปัจจุบันคาดว่ามีการบังคับใช้กฎหมายมากขึ้น และผู้ขับขี่มีความใส่ใจในความปลอดภัยมากขึ้น นักวิจัยคาดว่าจำนวนผู้คาดเข็มขัดนิรภัยจะเพิ่มขึ้นอีกเท่าตัว จึงปรับค่า p เพิ่มอีก 1 เท่าจาก 30% เป็น 60% แล้วจึงนำไปคำนวณค่าสัดส่วน (p = 0.6) ก่อนแทนในสูตรเพื่อคำนวณขนาดตัวอย่าง

การปรับค่า p ก่อนแทนลงในสูตรคำนวณขนาดตัวอย่างต้องทำด้วยความระมัดระวัง ถ้าค่าที่ปรับได้ต่ำกว่าที่เป็นจริง จะมีผลทำให้ต้องใช้ตัวอย่างในการศึกษามากขึ้น ทำให้เสียเงินและเวลาเพิ่มขึ้น แต่ถ้าค่าที่ปรับน้อยกว่าค่าที่ควรจะเป็น จะทำให้ค่าความเที่ยงของการประมาณมีขนาดมากกว่าที่กำหนด

e เป็นความเที่ยงของการประมาณค่า เป็นความแตกต่างของค่าสัดส่วนของประชากรที่นักวิจัยประมาณจากข้อมูลตัวอย่างกับค่าสัดส่วนของประชากร (ได้จากการศึกษาที่ผ่านมา) ถ้าต้องการให้ผลการประมาณใกล้เคียงมากค่า e จะเล็ก การกำหนดค่า e จะพิจารณาจากเงื่อนไขการใช้งานและข้อจำกัดของทรัพยากรในการทำวิจัย

ตัวอย่างที่ 1 อัตราการป่วยวัณโรคในชุมชนเป็น 5/10,000 คน ถ้าในการศึกษาได้ค่าประมาณ 3/10,000 คน หรือ 7/10,000 ค่า e เท่ากับ 2/10,000 ก็นับว่าใกล้เคียงพอแล้ว เพราะอัตราการป่วยวัณโรคที่ต่ำกว่า 10/10,000 คน จะใช้มาตรการในการป้องกันเหมือนกัน

Using Statistics in Reseach 1_1 .indd 42 dummy 4 30/6/2557 11:29:28

ตัวอย่างที่ 2 ในการหาอัตราการใช้ยาแก้ปวดที่ไม่เหมาะสม ทางจังหวัดกำหนดไว้ว่า ถ้ามีอัตราการใช้ยาแก้ปวดที่ไม่เหมาะสมร้อยละ 10 จะใช้มาตรการปกติในการให้ความรู้แก่ประชาชน ถ้าร้อยละ 25 จะเพิ่มวิธีการให้เข้มขึ้น และถ้ามากกว่าร้อยละ 40 จะต้องมีการจัดรณรงค์ ข้อมูลในรายงานวิจัยที่ผ่านมาพบว่ามีการใช้ยาแก้ปวดที่ไม่เหมาะสมร้อยละ 24 ในการกำหนดค่า e นักวิจัยได้หารือกันมีข้อสรุปว่า ถ้าในการศึกษาครั้งนี้กำหนดให้ e มีค่าร้อยละ 4 จากการสำรวจได้ค่า 95% ช่วงเชื่อมั่นร้อยละ 8–27 ซึ่งคร่อมอยู่ 2 มาตรการ ทำให้ไม่สามารถระบุได้ว่าจะใช้มาตรการใด ถ้ากำหนดค่า e ให้เล็กลงจะได้ขนาดตัวอย่างที่ใหญ่ขึ้น 95% ช่วงเชื่อมั่นจะแคบ ทำให้ไม่คร่อม 2 มาตรการ

ข้อควรระวังในการกำหนดค่า e มีนักวิจัยจำนวนหนึ่งกำหนดค่า e เท่ากับ 0.05 โดยเข้าใจผิดว่าเป็นค่าเดียวกันกับความผิดพลาดในการสุ่มตัวอย่าง (α) ทำให้การศึกษากรณีมีค่า p เล็กๆน้อยกว่าร้อยละ 5 จะมีความเที่ยงของการประมาณค่ามากกว่าอัตราที่ต้องการหา เช่น ต้องการหาอัตราความชุกที่คาดว่าจะพบร้อยละ 3 แต่กำหนดให้มีความเที่ยงของการประมาณค่าร้อยละ 5 ซึ่งไม่ถูกต้อง เพราะมีผลทำให้ค่าสัดส่วนที่ประมาณได้มีค่าเป็นลบ

Using Statistics in Reseach 1_1 .indd 43 dummy 4 27/8/2557 11:49:57

สรุป

ในการคำนวณขนาดตัวอย่าง เมื่อการประมาณค่าสัดส่วน สามารถคำนวณด้วยโปรแกรมคำนวณขนาดตัวอย่าง โดยนักวิจัยนำข้อมูลจากงานวิจัยที่ผ่านมาแทนค่าพารามิเตอร์ต่างๆลงในสูตร วิธีการกำหนดค่าพารามิเตอร์ทั้ง 3 ค่าได้เสนออยู่ในเรื่องนี้เพื่อใช้เป็นแนวทางพิจารณาค่าที่เหมาะสม ซึ่งจะทำให้ได้ขนาดตัวอย่างที่เหมาะสมในการศึกษา

Using Statistics in Reseach 1_1 .indd 44 dummy 4 27/8/2557 11:49:57

10

การกำหนดค่า e ในการคำนวณขนาดตัวอย่างเพื่อการศึกษาความชุก

การคำนวณขนาดตัวอย่างสำหรับงานวิจัยที่ศึกษาความชุกหรืออุบัติการณ์จะคำนวณขนาดตัวอย่างด้วยสูตรการประมาณค่าสัดส่วน

$$n = \frac{Z^2_{\alpha/2}\, p(1-p)}{e^2}$$

ในการคำนวณจะกำหนดค่า α จากระดับความเชื่อมั่น และค่า p จากงานวิจัยที่ผ่านมา หรือจากการศึกษานำร่อง สำหรับการกำหนดค่า e (ความเที่ยงของการประมาณค่า) นักวิจัยต้องกำหนดจากเงื่อนไขการใช้งานและข้อจำกัดของทรัพยากรในการทำวิจัย

ความไม่เหมาะสมของการกำหนดค่า e ที่พบมากที่สุด คือ การกำหนดค่า e เท่ากับ 0.05 ซึ่งเป็นค่าเดียวกับค่า α ทำให้การศึกษาโรคที่มีความชุกต่ำกว่าร้อยละ 10 มีค่าความผิดพลาดของการประมาณค่าสูง เช่น ความชุกของการเป็นโรคหลอดเลือดสมองเท่ากับ 3/1,000 คน ถ้ากำหนดค่า e เท่ากับ 0.05 จะเป็นการประมาณค่าความชุกที่คาดว่าจะพบ 3/1,000 คน แต่ยอมให้ค่าประมาณผิดพลาดได้ 50/1,000 คน ซึ่งไม่ถูกต้อง

ปัญหาที่ผู้เขียนถูกถามมากที่สุดในเรื่องการคำนวณขนาดตัวอย่างในกรณีนี้ คือ มีหลักกำหนดหรือไม่ว่าค่า e อย่างต่ำควรเป็นเท่าไร หรือควรตั้งเป็นร้อยละเท่าไรของค่า p และจะรู้ได้อย่างไรว่าที่นักวิจัยกำหนดมีความเหมาะสม

การกำหนดค่า e ของแต่ละการศึกษาจะขึ้นอยู่กับเงื่อนไขการนำผลการศึกษาไปใช้ ความต้องการขนาดตัวอย่างในการตอบคำถามรองทรัพยากรและเวลาที่ใช้ในการศึกษา ซึ่งจะแตกต่างกันไปในแต่ละการศึกษา

ตัวอย่างที่ 1 ในการศึกษาความชุกเพื่อประเมินว่าสามารถควบคุมโรคได้ตามแผนที่วางไว้หรือไม่ เช่น แผนกำหนดไว้ว่าความชุกของการสูบบุหรี่เป็นประจำในคนอายุ 15 ปีขึ้นไปไม่เกินร้อยละ 20 การคำนวณขนาดตัวอย่างเพื่อประเมินผลว่าการดำเนินงานได้ผลสำเร็จตามแผนหรือไม่ จะกำหนดค่า p = 0.2 สำหรับค่า e จะพิจารณาว่าถ้าผลการสำรวจที่ได้พบว่ามีความชุกเกินกว่าที่กำหนด (ร้อยละ 20) จำนวนเท่าไรที่พอจะอนุโลมว่าน่าจะยอมรับได้ เช่น กำหนดให้เป็นร้อยละ 2 (ค่า e = 0.02) หมายถึงถ้าค่าความชุกเป็นร้อยละ 20 จริง ผลการศึกษามีโอกาสที่จะได้ค่าประมาณสูงถึงร้อยละ 22 (ซึ่งไม่เกินจำนวนที่ยอมรับได้) ดังนั้นในกรณีนี้การกำหนดค่า e จะพิจารณาว่าขนาดความชุกที่เกินกว่าแผนเท่าไรที่ยอมรับได้เป็นสำคัญ ในทำนองเดียวกันสามารถใช้หลักนี้ในการพิจารณากำหนดค่า e ในการศึกษาความครอบคลุมของกิจกรรมบริการต่างๆ

Using Statistics in Reseach 1_1 .indd 46 dummy 4 30/6/2557 11:29:28

ตัวอย่างที่ 2 ในกรณีที่เป็นการศึกษาความชุกโดยอยากทราบว่าโรคที่ศึกษามีความชุกเท่าไร เพื่อจะได้ทราบขนาดของปัญหานำไปกำหนดวิธีป้องกันโรค เช่น จากรายงานอุบัติการณ์โรคความดันโลหิตสูงเท่ากับ 14/1,000 คน ในการสำรวจถ้ากำหนดค่า e = 5/1,000 คน ผลการศึกษาอาจพบความชุก 9/1,000 (14 – 5) คน หรือ 19/1,000 (14 + 5) คน ที่เลือกค่า e ค่อนข้างใหญ่เพราะข้อมูลทางการบริหารเพื่อตัดสินใจกำหนดกิจกรรมไม่ต้องการข้อมูลที่กระชับมาก เช่น ในกรณีนี้ ถ้าความชุกของโรคความดันโลหิตสูงมีไม่เกิน 20/1,000 จะไม่เปลี่ยนวิธีการดำเนินการป้องกันโรค

แต่ถ้าเป็นการสำรวจความชุกและหาปัจจัยเสี่ยงจากตัวอย่างข้างต้น ถ้ากำหนดความเชื่อมั่น 95% จะคำนวณขนาดตัวอย่างได้ 2,122 คน ขนาดตัวอย่างจำนวนนี้จะให้คำตอบอุบัติการณ์ได้ตามระดับความกระชับที่กำหนด แต่จากขนาดตัวอย่างที่คำนวณได้ ข้อมูลที่เก็บมาอาจพบจำนวนคนที่เป็นโรคเพียง 30 คน ซึ่งไม่เพียงพอในการตอบคำถามรองในเรื่องปัจจัยเสี่ยงของการเกิดโรค

ดังนั้นในกรณีที่มีคำถามรองเรื่องของปัจจัยเสี่ยง นักวิจัยควรกำหนดความกระชับให้เล็กเพื่อให้ได้ขนาดตัวอย่างที่ใหญ่ที่สุดเท่าที่จะทำได้ นอกจากจะได้ค่าอุบัติการณ์หรือความชุกที่มีความน่าเชื่อถือได้เพิ่มขึ้นแล้ว ยังจะมีจำนวนผู้เป็นโรคเพียงพอที่จะตอบคำถามเรื่องปัจจัยเสี่ยง

ในกรณีต้องการคำตอบของปัจจัยเสี่ยงที่เชื่อถือได้ด้วย ควรทำการคำนวณขนาดตัวอย่างสำหรับคำถามรอง ทำให้ได้ขนาดตัวอย่างที่เหมาะสมที่ตอบได้ทั้งคำถามหลักและคำถามรอง

Using Statistics in Reseach 1_1 .indd 47 dummy 4 27/8/2557 11:49:57

ตัวอย่างที่ 3 ในการศึกษาโรคที่มีความชุกหรืออุบัติการณ์ต่ำๆที่ต้องใช้ตัวอย่างขนาดใหญ่ ซึ่งต้องใช้ทรัพยากรและเวลาในการดำเนินการศึกษา แต่ผู้วิจัยมีค่าใช้จ่ายในการดำเนินการจำกัดหรือต้องทำให้เสร็จในเวลาที่กำหนด เช่น ในการศึกษาอุบัติการณ์ของโรคความดันโลหิตสูงในตัวอย่างที่ 2 การคำนวณขนาดตัวอย่างในกรณีนี้นักวิจัยควรคำนวณขนาดตัวอย่างที่ค่า e ขนาดต่างๆแล้วนำมาพิจารณาดูความเป็นไปได้ในการทำ

อุบัติการณ์ = 14/1,000, α = 0.05	
ค่า e/1,000	**ขนาดตัวอย่าง**
7	1,083
6	1,473
5	2,122
4	3,315
3	5,892

จากตาราง ถ้านักวิจัยมีเงินและเวลาทำการศึกษาได้ประมาณ 1,100 คน ค่าความกระชับจะเท่ากับ 7/1,000 คน ค่าความกระชับที่ได้ค่อนข้างใหญ่ และผลการศึกษาอาจพบความชุก 21/1,000 (14 + 7) คน ซึ่งเกิน 20/1,000 ของเกณฑ์ที่ใช้พิจารณาเปลี่ยนวิธีการดำเนินการป้องกันโรค จึงไม่เหมาะสมในการใช้ e = 7/1,000 ควรไปหางบประมาณเพิ่มเพื่อศึกษาด้วยตัวอย่างที่มากกว่านี้

Using Statistics in Reseach 1_1 .indd 48 dummy 4 30/6/2557 11:29:28

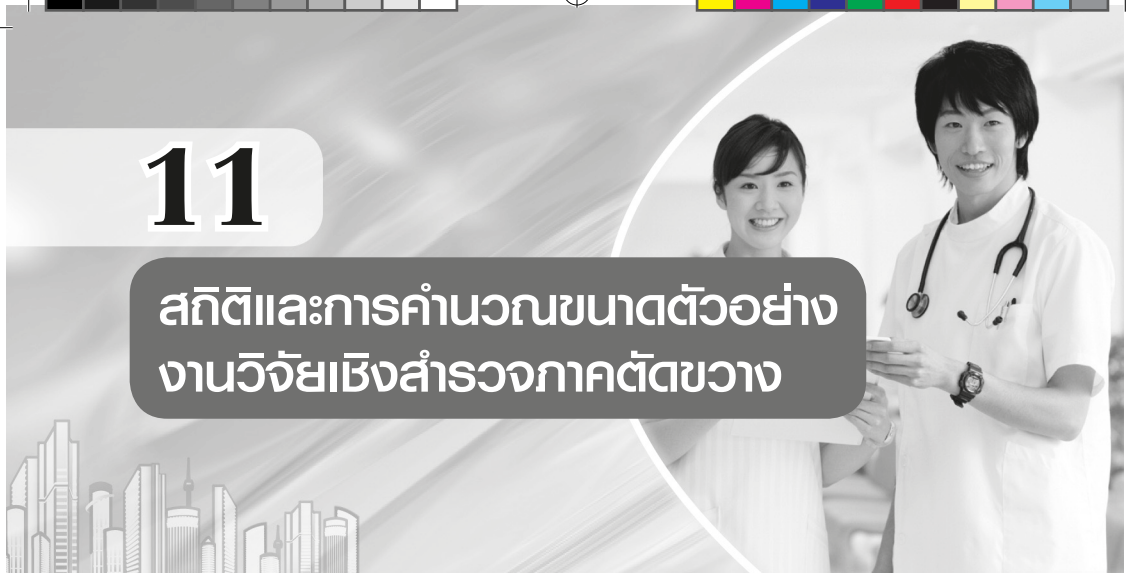
ในการตัดสินใจด้วยเงื่อนไขทางวิชาการว่าควรจะศึกษาด้วยตัวอย่างขนาดเล็กหรือไม่ จะพิจารณาว่าการศึกษาที่จะทำจะให้ความรู้อะไรเพิ่มจากเดิมที่นักวิจัยอื่นเคยทำมาแล้วหรือไม่ ถ้าตัวอย่างที่จะศึกษาน้อยกว่าการศึกษาที่ผ่านมาผลการศึกษาที่ได้จะมีความน่าเชื่อถือต่ำกว่า ถ้าไม่ได้ความรู้เพิ่มขึ้นจากเดิมก็ไม่ควรทำ แต่ถ้างานวิจัยที่ผ่านมามีขนาดตัวอย่างใหญ่กว่า แต่วิธีการศึกษามีจุดบกพร่องที่มีผลต่อความถูกต้องและความน่าเชื่อถือของผลการศึกษา ก็ควรพิจารณาทำด้วยวิธีการที่ดีกว่าซึ่งอาจได้ความรู้ที่มีความน่าเชื่อถือมากกว่า

ในกรณีที่มีข้อจำกัดเรื่องขนาดตัวอย่างที่ใช้ในการศึกษา โอกาสที่จะมีข้อมูลเพียงพอที่จะตอบคำถามรองจะน้อยมาก แต่นักวิจัยยังสามารถนำข้อมูลมาวิเคราะห์เพื่อตอบคำถามปัจจัยเสี่ยงได้ แต่ควรมีความระมัดระวังในการสรุปผลการวิเคราะห์

สรุป

การกำหนดค่า e จะพิจารณาจากการใช้ประโยชน์ของผลการสำรวจ การใช้ประโยชน์แต่ละด้านจะมีแนวคิดในการกำหนดค่า e แตกต่างกัน การทบทวนองค์ความรู้ที่ผ่านมาจะช่วยให้กำหนดค่า e ได้อย่างเหมาะสม

Using Statistics in Reseach 1_1 .indd 49 dummy 4 30/6/2557 11:29:28



11

สถิติและการคำนวณขนาดตัวอย่างงานวิจัยเชิงสำรวจภาคตัดขวาง

การศึกษาวิจัยโดยวิธีสำรวจภาคตัดขวาง (cross-sectional survey) ใช้แพร่หลายในงานระบาดวิทยาและการจัดบริการสาธารณสุข การวิจัยเชิงสำรวจภาคตัดขวางใช้กับงานวิจัย 2 ประเภท คือ การวิจัยเชิงพรรณนาภาคตัดขวาง (cross-sectional descriptive survey) และการวิจัยเชิงวิเคราะห์ภาคตัดขวาง (cross-sectional analytical survey) ซึ่งทั้ง 2 แบบใช้ตอบคำถามที่ต่างกัน สถิติที่ใช้วิเคราะห์และการคำนวณขนาดตัวอย่างแต่ละวิธีแตกต่างกัน

การวิจัยเชิงสำรวจภาคตัดขวางที่เป็นการวิจัยเชิงพรรณนาจะศึกษาในประชากรกลุ่มเดียว เช่น กลุ่มที่เป็นโรค โดยมีวัตถุประสงค์เพื่อบรรยายลักษณะด้านต่าง ๆของประชากรที่ศึกษา ตัวอย่างเช่น ศึกษาจำนวน ลักษณะ และสาเหตุของความพิการหรือปัญหาในการเคลื่อนไหวร่างกายของคนพิการ เพื่อนำมาจัดทำแผนพื้นฟูช่วยเหลือคนพิการ การศึกษานี้เป็นการวิจัยเชิงพรรณนา ประชากรคือกลุ่มที่มีความพิการหรือผู้มีปัญหาในการเคลื่อนไหวร่างกายเพียงกลุ่มเดียว

คำถามงานวิจัยเชิงสำรวจเชิงพรรณนาอีกลักษณะหนึ่ง คือการวิจัยหาอัตราการเป็นโรค อัตราความครอบคลุมของบริการสาธารณสุขใน

ประชากรเป้าหมาย (กลุ่มเดียว) เช่น ต้องการหาอัตราความพิการของประชาชนในอำเภอแวงใหญ่ ซึ่งการพรรณนาลักษณะและอัตราความพิการอาจแยกตามกลุ่มอายุ เพศ หรือลักษณะต่างๆ

สำหรับการวิจัยเชิงวิเคราะห์ภาคตัดขวางจะมีลักษณะคำถามงานวิจัยเป็นการหาปัจจัยเสี่ยงหรือปัจจัยที่มีอิทธิพล ซึ่งจำเป็นต้องมีประชากร 2 กลุ่มเพื่อใช้เปรียบเทียบหาปัจจัยเสี่ยง ตัวอย่างเช่น การศึกษาปัจจัยเสี่ยงของเด็กน้ำหนักเกิน ในการศึกษาจะสุ่มเก็บข้อมูลพื้นฐานและปัจจัยต่างๆจากประชากร 2 กลุ่ม คือ ประชากรกลุ่มเด็กน้ำหนักเกินและประชากรกลุ่มเด็กน้ำหนักปกติ นำมาวิเคราะห์หาปัจจัยเสี่ยง

จากการที่งานวิจัยเชิงสำรวจภาคตัดขวางอาจเป็นได้ทั้งการวิจัยเชิงพรรณนาและการวิจัยเชิงวิเคราะห์ ดังนั้นสถิติที่ใช้และการคำนวณขนาดตัวอย่างจะขึ้นอยู่กับวัตถุประสงค์ของการสำรวจ

ถ้าวัตถุประสงค์ของการวิจัยเชิงสำรวจเป็นการพรรณนาลักษณะที่สนใจของประชากรกลุ่มเดียว เช่น กลุ่มคนที่เป็นโรค หรือการหาอัตราต่างๆของประชากรกลุ่มเป้าหมาย สถิติที่ใช้ในการวิเคราะห์จะใช้สถิติ Z หรือ t ประมาณค่า 95% ช่วงเชื่อมั่นของสัดส่วนหรือค่าเฉลี่ย การคำนวณขนาดตัวอย่างจะใช้สูตรคำนวณขนาดตัวอย่างเพื่อประมาณค่าสัดส่วนหรือค่าเฉลี่ยของประชากร

ถ้าวัตถุประสงค์เพื่อหาปัจจัยเสี่ยงหรือปัจจัยที่มีอิทธิพล จะวิเคราะห์ด้วยสถิติที่ใช้เปรียบเทียบค่าสัดส่วนหรือค่าเฉลี่ย เช่น OR (Odds Ratio), χ^2, t-test หรือ ANOVA ขึ้นอยู่กับประเภทของตัวแปรผลที่จะนำมาคำนวณค่าสถิติ การคำนวณขนาดตัวอย่างจะใช้สูตรคำนวณขนาดตัวอย่างเพื่อการเปรียบเทียบสัดส่วนหรือค่าเฉลี่ยระหว่างสองประชากร

ถ้าคำถามหลักต้องการหาปัจจัยเสี่ยงหลายตัวที่มีผลต่อตัวแปรตามที่เป็นตัวแปรกลุ่มแบบมี 2 ค่า (dichotomous variable) จะใช้การ

Using Statistics in Reseach 1_1 .indd 51 dummy 4 30/6/2557 11:29:29

วิเคราะห์เชิงพหุเพื่อดูว่ามีปัจจัยใดบ้างที่ร่วมกันทำให้เกิดเป็นโรค การวิเคราะห์จะใช้ multiple logistic regression คำนวณค่า OR การคำนวณขนาดตัวอย่างจะคำนวณจากสูตรการวิเคราะห์ logistic regression* ถ้าตัวแปรตามเป็นตัวแปรต่อเนื่อง เช่น คะแนนคุณภาพชีวิต และตัวแปรอิสระก็เป็นตัวแปรต่อเนื่องด้วย จะใช้ multiple regression ในการวิเคราะห์ การคำนวณขนาดตัวอย่างจะคำนวณจากสูตรการวิเคราะห์ multiple regression**

ในกรณีที่การศึกษาเชิงสำรวจภาคตัดขวางมีหลายวัตถุประสงค์ เช่น วัตถุประสงค์หลักเป็นการหาอัตราความครอบคลุมของการตรวจมะเร็งปากมดลูก จากข้อมูลที่เก็บมาศึกษาอัตราความครอบคลุม นักวิจัยสามารถแบ่งข้อมูลที่ได้ออกเป็น 2 กลุ่ม คือ กลุ่มที่มารับการตรวจและกลุ่มที่ไม่มารับการตรวจ จึงสามารถคำนวณหาปัจจัยที่มีอิทธิพลต่อการมารับการตรวจมะเร็งปากมดลูกจากข้อมูลที่เก็บมาศึกษาเรื่องความครอบคลุม นักวิจัยจึงสามารถตั้งวัตถุประสงค์รองเป็นการหาปัจจัยที่มีผลต่อความครอบคลุมของการตรวจมะเร็งปากมดลูก การคำนวณขนาดตัวอย่างจะคำนวณเพื่อตอบคำถามหลักเรื่องการหาอัตราความครอบคลุม ดังนั้นขนาดตัวอย่างที่ได้อาจไม่พอสำหรับการตอบคำถามรอง นักวิจัยควรมีความระมัดระวังในการแปลและสรุปผลการวิเคราะห์เรื่องปัจจัยที่มีอิทธิพลต่อการมารับการตรวจมะเร็งปากมดลูก

---

*, ** F.Y. Hsieh, Daniel A. Bloch, and Michael D. Larsen, “A Simple Method of Sample Size Calculation for Linear and Logistic Regression.” **Statistics in Medicine.** 1998; 17: 1623–34.

Using Statistics in Reseach 1_1 .indd 52 dummy 4 27/8/2557 11:49:57

สรุป

ในการศึกษาเชิงสำรวจภาคตัดขวางสามารถใช้ได้กับแบบงานวิจัยเชิงพรรณนาและแบบงานวิจัยเชิงวิเคราะห์ สถิติที่ใช้ในการวิเคราะห์และการคำนวณขนาดตัวอย่างจะกำหนดตามวัตถุประสงค์หลักของงานวิจัย ดังนั้นนักวิจัยควรทำความเข้าใจกับปัญหาให้ชัดเจนว่าในการศึกษานี้ต้องการคำตอบใดเป็นคำตอบหลัก จะได้นำมาตั้งวัตถุประสงค์หลักให้สอดคล้องกับคำตอบที่ต้องการรู้มากที่สุด

ความไม่เหมาะสมที่พบได้บ่อยคือนักวิจัยมีความเข้าใจว่าการวิจัยเชิงสำรวจเป็นการวิจัยเชิงพรรณนาอย่างเดียว และใช้วิธีการคำนวณขนาดตัวอย่างเพื่อการประมาณค่าสัดส่วนกับการสำรวจทุกเรื่อง โดยไม่ดูวัตถุประสงค์การศึกษาและประเภทข้อมูลของตัวแปรผล

Using Statistics in Reseach 1_1 .indd 53 dummy 4 30/6/2557 11:29:29

12

การคำนวณขนาดตัวอย่างในการศึกษาความชุกและปัจจัยเสี่ยงพร้อมกัน

มีคำถามว่าการทำวิจัยเรื่องการหาความชุกและปัจจัยเสี่ยงพร้อมกันควรคำนวณขนาดตัวอย่างอย่างไร เพราะการหาความชุกใช้แบบวิจัยเชิงพรรณนา ส่วนการหาปัจจัยเสี่ยงใช้แบบวิจัยเชิงวิเคราะห์

ในการออกแบบงานวิจัยและคำนวณขนาดตัวอย่างจะพิจารณาจากคำถามหรือวัตถุประสงค์หลัก ถ้าวัตถุประสงค์หลักต้องการหาความชุกจะใช้แบบวิจัยเชิงพรรณนาเพื่อประมาณค่าความชุก การคำนวณขนาดตัวอย่างจะใช้สูตรคำนวณขนาดตัวอย่างเพื่อประมาณค่าสัดส่วน (ความชุก)

ข้อมูลที่ได้จากการศึกษาความชุกของการเป็นโรคสามารถแบ่งได้ออกเป็น 2 กลุ่ม คือ กลุ่มคนเป็นโรคและกลุ่มคนไม่เป็นโรค จึงสามารถนำข้อมูลนี้มาคำนวณค่าสถิติที่ใช้วิเคราะห์ปัจจัยเสี่ยงได้ด้วย ดังนั้นในงานวิจัยที่ต้องการศึกษาความชุกจึงมักจะมีคำถามรองเป็นการหาปัจจัยเสี่ยง

ถ้ามีคำถามรองเป็นการหาปัจจัยเสี่ยง นักวิจัยต้องกำหนดปัจจัยที่ต้องการพิสูจน์ลงในแบบเก็บข้อมูลด้วย การหาปัจจัยเสี่ยงจากข้อมูลศึกษาความชุก ค่าสถิติที่ได้จากการวิเคราะห์ปัจจัยเสี่ยงมักจะมีความเที่ยงต่ำ (ค่า 95% CI ของ OR กว้าง) หรือปัจจัยเสี่ยงที่เคยพบจากการศึกษาอื่นแต่กลับไม่พบในการศึกษานี้ ที่เป็นเช่นนี้เพราะไม่มีการ

คำนวณขนาดตัวอย่างเพื่อหาปัจจัยเสี่ยง ทำให้ขนาดตัวอย่างของคนที่เป็นโรคมีไม่มากพอที่จะระบุปัจจัยเสี่ยงได้ครอบคลุมทุกปัจจัย โดยเฉพาะโรคที่มีความชุกต่ำ อัตราส่วนของคนเป็นโรคมีขนาดแตกต่างจากกลุ่มที่ไม่เป็นโรคมาก เช่น การศึกษาโรคที่มีความชุกร้อยละ 8 โดยใช้ขนาดตัวอย่างในการศึกษา 350 คน ในการคำนวณปัจจัยเสี่ยงจะมีคนเป็นโรคเพียง 28 คน เทียบปัจจัยเสี่ยงกับคนไม่เป็นโรค 322 คน

ดังนั้นการศึกษาปัจจัยเสี่ยงจากข้อมูลที่ศึกษาความชุกจึงมีประโยชน์เพียงเพื่อเป็นการดูเบื้องต้น (explore) ว่ามีปัจจัยใดเป็นปัจจัยเสี่ยง หรือปัจจัยใดมีแนวโน้มว่าจะเป็นปัจจัยเสี่ยง จะได้ใช้ความรู้ที่ได้ไปใช้ออกแบบในการศึกษาหาปัจจัยเสี่ยงต่อไป

อาจมีคำถามว่า สามารถคำนวณขนาดตัวอย่างหาความชุกเพื่อตอบวัตถุประสงค์หลัก และคำนวณขนาดตัวอย่างหาปัจจัยเสี่ยงเพื่อตอบวัตถุประสงค์รองด้วยจะได้หรือไม่

คำตอบคือ **ได้** เพราะการที่นักวิจัยต้องการศึกษาหาความชุกและปัจจัยเสี่ยงไปพร้อมกันก็สามารถทำได้ เพราะตัวอย่างที่ใช้เป็นกลุ่มเดียวกัน ในการหาปัจจัยเสี่ยงผู้วิจัยจะต้องเลือกว่าจะเป็นแบบวิจัยเชิงวิเคราะห์แบบใด ในการศึกษาปัจจัยเสี่ยงพร้อมกับความชุกจะเลือกได้ 2 แบบ คือ แบบวิจัยเชิงวิเคราะห์ภาคตัดขวาง หรือแบบวิจัยเชิงวิเคราะห์ case-control

ถ้าเลือกแบบวิจัยเชิงวิเคราะห์ภาคตัดขวาง การคำนวณขนาดตัวอย่างจะทราบว่าต้องใช้จำนวนคนเป็นโรค (cases) และคนไม่เป็นโรค (control) กลุ่มละกี่คน

จากขนาดตัวอย่างที่ใช้หาความชุก นักวิจัยสามารถประมาณจำนวนคนเป็นโรคที่คาดว่าจะพบมีจำนวนเท่าไร ถ้าจำนวนคนเป็นโรคที่คาดว่าจะพบน้อยกว่าจำนวนที่คำนวณได้จากปัจจัยเสี่ยง นักวิจัยต้องเพิ่มขนาดตัวอย่างในการหาความชุกเพื่อให้ได้คนเป็นโรคครบตามจำนวนที่

Using Statistics in Reseach 1_1 .indd 55 dummy 4 30/6/2557 11:29:29

คำนวณได้จากปัจจัยเสี่ยง ส่วนกลุ่มไม่เป็นโรคมักจะมีจำนวนมากกว่าที่คำนวณได้จากปัจจัยเสี่ยง โดยเฉพาะโรคที่มีความชุกต่ำจะมีกลุ่มคนไม่เป็นโรคมากกว่าที่คำนวณได้จากปัจจัยเสี่ยงหลายเท่า ถ้าเลือกแบบวิจัยนี้ นักวิจัยจะต้องนำกลุ่มไม่เป็นโรคทุกคนเข้าในการวิเคราะห์ จะสุ่มมาบางส่วนไม่ได้ การที่มีกลุ่มไม่เป็นโรคจำนวนมากกว่ากลุ่มเป็นโรคหลายเท่าย่อมมีผลต่อความถูกต้องของการระบุปัจจัยเสี่ยง

ถ้านักวิจัยอยากจะลดจำนวนกลุ่มควบคุมลง ให้เลือกแบบงานวิจัยเชิงวิเคราะห์ case-control การคำนวณขนาดตัวอย่างจะทราบว่าต้องใช้คนเป็นโรคและคนไม่เป็นโรคกลุ่มละกี่คน ถ้าจำนวนคนเป็นโรคไม่พอนักวิจัยต้องเพิ่มขนาดตัวอย่างในการหาความชุกเพื่อให้ได้จำนวนคนเป็นโรคเพียงพอที่จะตอบปัจจัยเสี่ยงเช่นกัน

แบบวิจัย case-control แตกต่างจากแบบวิจัยเชิงวิเคราะห์ภาคตัดขวางตรงที่นักวิจัยสามารถกำหนดวิธีการสุ่มคนไม่เป็นโรคให้มีคุณลักษณะและจำนวนเท่ากับที่ต้องการ การเก็บข้อมูลเกี่ยวกับปัจจัยเสี่ยงของคนไม่เป็นโรคก็เก็บเฉพาะกลุ่มที่ถูกสุ่มมาเท่านั้น ไม่ต้องเก็บจากคนไม่เป็นโรคทั้งหมด ทำให้ประหยัดเวลาและค่าใช้จ่าย ทั้งยังสามารถระบุปัจจัยเสี่ยงได้ถูกต้องเพิ่มขึ้น

การศึกษาปัจจัยเสี่ยงแบบ case-control ยังสามารถออกแบบให้ลดผลกระทบจากตัวแปรกวน เช่น กำหนดหลักเกณฑ์ในการสุ่มเลือกกลุ่มคนไม่เป็นโรค การจับคู่ (matching) แบบตัวต่อตัว หรือจับคู่แบบกลุ่ม เป็นต้น ทำให้ระบุปัจจัยเสี่ยงได้ถูกต้องเพิ่มขึ้น

Using Statistics in Reseach 1_1 .indd 56 dummy 4 27/8/2557 11:49:57

สรุป

การคำนวณขนาดตัวอย่างสำหรับการศึกษาความชุกและปัจจัยเสี่ยง ถ้านักวิจัยต้องการตอบเรื่องความชุกเป็นหลัก และการหาปัจจัยเสี่ยงเป็นการศึกษาเบื้องต้น เพื่อดูว่าปัจจัยใดเป็นปัจจัยเสี่ยงหรือมีแนวโน้มเป็นปัจจัยเสี่ยง การคำนวณขนาดตัวอย่างจะคำนวณหาความชุกเพื่อตอบวัตถุประสงค์หลักเพียงอย่างเดียว

แต่ถ้าต้องการหาปัจจัยเสี่ยงที่เชื่อถือได้จะต้องเลือกแบบวิจัยเชิงวิเคราะห์ที่เหมาะกับความชุกของโรค และต้องคำนวณขนาดตัวอย่างสำหรับการศึกษาปัจจัยเสี่ยงที่เป็นคำถามรองด้วย

Using Statistics in Reseach 1_1 .indd 57 dummy 4 30/6/2557 11:29:29

13

ความเหมาะสมของการใช้การสุ่มตัวอย่างแบบกลุ่มในการศึกษาปัจจัยเสี่ยง

มีคำถามว่าจะใช้การสุ่มตัวอย่างแบบกลุ่ม (cluster sampling) ในการศึกษาปัจจัยเสี่ยงได้หรือไม่ ในการทำความเข้าใจคำตอบนักวิจัยต้องเข้าใจหลักที่ใช้ในการศึกษาปัจจัยเสี่ยงและการสุ่มตัวอย่าง

การสุ่มตัวอย่างแบบกลุ่มใช้กันมากในการศึกษาขนาดปัญหาสาธารณสุขในพื้นที่ เช่น การศึกษาอัตราการป่วยด้วยโรคความดันโลหิตสูง หรือหาอัตราภาวะทุโภชนาการในเด็กก่อนวัยเรียน เป็นต้น เหตุผลที่ต้องมีการศึกษาเพราะไม่สามารถทราบจำนวนผู้ป่วยทั้งหมดจากระบบรายงาน หรือมีผู้ป่วยบางส่วนมีอาการป่วยแต่ไม่ไปรับการรักษาในสถานบริการ ดังนั้นการศึกษาส่วนใหญ่เป็นการหาจำนวนผู้ป่วยในพื้นที่เพื่อนำมาคำนวณอัตราการป่วยในพื้นที่ให้บริการของสถานบริการ

ในการศึกษาจะมีการกำหนดขอบเขตพื้นที่ทางภูมิศาสตร์ตามขอบเขตของระบบบริการที่ต้องการข้อมูลอัตราการป่วยไปใช้ในการวางแผนหรือประเมินผลงาน ประชากรเป้าหมายคือประชาชนทุกคนที่พักอาศัยกระจายอยู่ตามหมู่บ้านหรือชุมชนในพื้นที่ วิธีการสุ่มที่เหมาะสมคือการสุ่มแบบกลุ่ม จะใช้แบบขั้นตอนเดียว (single stage sampling) หรือหลายขั้นตอน (multi-stage sampling) ขึ้นอยู่กับขนาดของพื้นที่

วัตถุประสงค์หลักของการศึกษา คือ เพื่อหาค่าสัดส่วนของความครอบคลุมหรือความชุกที่เป็นตัวแทนของพื้นที่ ดังนั้นการคำนวณขนาดตัวอย่างและการวิเคราะห์จึงใช้วิธีการประมาณค่าสัดส่วนที่ใช้วิธีสุ่มตัวอย่างแบบกลุ่ม

ในบางการศึกษาอาจมีวัตถุประสงค์รองเพื่อหาปัจจัยเสี่ยงของโรคดังกล่าวด้วยก็สามารถทำได้ เพราะผลการศึกษาความชุกสามารถแยกตัวอย่างที่ศึกษาออกเป็น 2 กลุ่ม คือ กลุ่มเป็นโรคและกลุ่มไม่เป็นโรค นักวิจัยสามารถนำตัวแปรปัจจัยเสี่ยงของทั้ง 2 กลุ่มมาคำนวณหาปัจจัยเสี่ยง ค่าความเสี่ยงที่ได้จากข้อมูลการศึกษา การแปลผลเรื่องความเสี่ยงจากการศึกษาลักษณะนี้ควรทำด้วยความระมัดระวัง เพราะการศึกษาประเภทนี้ส่วนใหญ่จะมีจำนวนผู้ป่วยไม่เพียงพอจะตอบคำถามเรื่องปัจจัยเสี่ยง ตัวอย่างเช่น ในการศึกษาความชุกของผู้ป่วยเบาหวาน ขนาดตัวอย่างที่ใช้ศึกษา 1,000 คน พบว่าอัตราความชุกร้อยละ 1.6 จำนวนผู้ป่วยเบาหวานที่พบเพียง 16 คนอาจไม่เพียงพอสำหรับการตอบคำถามเรื่องปัจจัยเสี่ยง

ส่วนการศึกษาเรื่องปัจจัยเสี่ยงจะมีประชากร 2 กลุ่ม คือ กลุ่มเป็นโรคและกลุ่มไม่เป็นโรค ปัจจุบันอัตราการเป็นโรคต่างๆค่อนข้างต่ำ วิธีการศึกษาหาปัจจัยเสี่ยงจะใช้ case-control ซึ่งเริ่มจากคนเป็นโรคที่ได้จากผู้ป่วยที่มารับบริการ ส่วนคนไม่เป็นโรค อาจจะเป็นญาติ หรือผู้ป่วยโรคอื่นที่มารับบริการที่โรงพยาบาล (hospital control) หรืออาจไปสุ่มเลือกกลุ่มคนไม่เป็นโรคในชุมชน (community control) จะเห็นได้ว่าคุณสมบัติของประชากรที่ศึกษาจะเป็นเรื่องของการเป็นโรคไม่ใช่ตัวแทนของพื้นที่ วิธีการศึกษาแบบ case-control จึงไม่ต้องใช้วิธีการสุ่มผู้ป่วยแบบกลุ่ม

ในกรณีที่ศึกษาปัจจัยเสี่ยงโดยการสุ่มตรวจคนในพื้นที่เพื่อหาคนเป็นและไม่เป็นโรคเช่นเดียวกับการศึกษาเรื่องความชุกก็ทำได้ แต่การ

Using Statistics in Reseach 1_1 .indd 59 dummy 4 30/6/2557 11:29:29

คำนวณขนาดตัวอย่างจะต้องใช้วิธีการคำนวณเพื่อหาปัจจัยเสี่ยง โดยขนาดตัวอย่างจะเป็นจำนวนคนเป็นโรคและจำนวนคนไม่เป็นโรค ในการตอบคำถามหลักเรื่องปัจจัยเสี่ยงเป็นเรื่องของตัวโรค ดังนั้นคนเป็นโรค/ไม่เป็นโรคจึงไม่จำเป็นต้องสุ่มให้เป็นตัวแทนของพื้นที่ ทั้งนี้เพราะปัจจัยเสี่ยงเป็นปัจจัยที่เกี่ยวข้องกับตัวโรคไม่ใช่ตัวแทนพื้นที่ มีบางโรคที่ปัจจัยเสี่ยงเป็นสถานที่พักอาศัย (residential area) ซึ่งสถานที่พักอาศัยมักจะมีความสัมพันธ์ทางอ้อมกับการเกิดโรค เช่น ที่พักในเขตเมืองมีเสียงดังกว่าที่พักในเขตชนบทซึ่งมีผลต่อการได้ยิน ถ้านักวิจัยเก็บข้อมูลความดังของเสียงในที่พักอาศัยซึ่งเป็นปัจจัยที่มีความสัมพันธ์โดยตรง ตัวแปรพักอาศัยจะไม่ใช่ปัจจัยที่ต้องศึกษา

การคำนวณเพื่อให้ทราบว่าจะต้องตรวจคนกี่คนเพื่อให้ได้คนเป็นโรคจำนวนที่ต้องการ จะนำจำนวนของคนเป็นโรคที่ต้องใช้ในการศึกษาและความชุกมาคำนวณหาจำนวนคนที่จะตรวจ เช่น จากการคำนวณขนาดตัวอย่างเพื่อหาปัจจัยเสี่ยงต้องการคนเป็นโรค 150 คน ถ้าโรคดังกล่าวมีอัตราความชุกร้อยละ 1.6 จะต้องตรวจคนทั้งหมดอย่างน้อย 9,375 คนจึงจะได้คนเป็นโรค 150 คน ขนาดตัวอย่างที่นำมาตรวจทั้งหมดจะเท่ากับ 9,375 คน ส่วนกลุ่มควบคุมก็สุ่มเอามาจากผู้ที่ตรวจแล้วไม่เป็นโรค จะเห็นได้ว่าประชากรที่ศึกษาคือคนเป็นและไม่เป็นโรค ดังนั้นการที่จะใช้การสุ่มแบบกลุ่มเพื่อให้ได้จำนวนผู้เป็นโรคตามจำนวนที่คำนวณก็ได้ หรือจะทำการตรวจทีละหมู่บ้านเรียงต่อกันไปเพื่อให้ได้คนเป็นโรค 150 คนก็ทำได้โดยไม่มีอคติในการเลือกตัวอย่าง

ในการศึกษาปัจจัยเสี่ยงโดยการตรวจประชาชนในพื้นที่เพื่อหาคนเป็นโรค การให้การวินิจฉัยว่าเป็นโรคจะต้องมีเกณฑ์มาตรฐานในการวินิจฉัย ตัวอย่างเกณฑ์การวินิจฉัยโรคเบาหวาน คือ พบระดับน้ำตาล FBS มากกว่าหรือเท่ากับ 126 มิลลิกรัมต่อเดซิลิตร อย่างน้อย 2 ครั้ง โดยตรวจซ้ำในวันถัดไป หรือพบระดับน้ำตาลมากกว่าหรือเท่ากับ 126

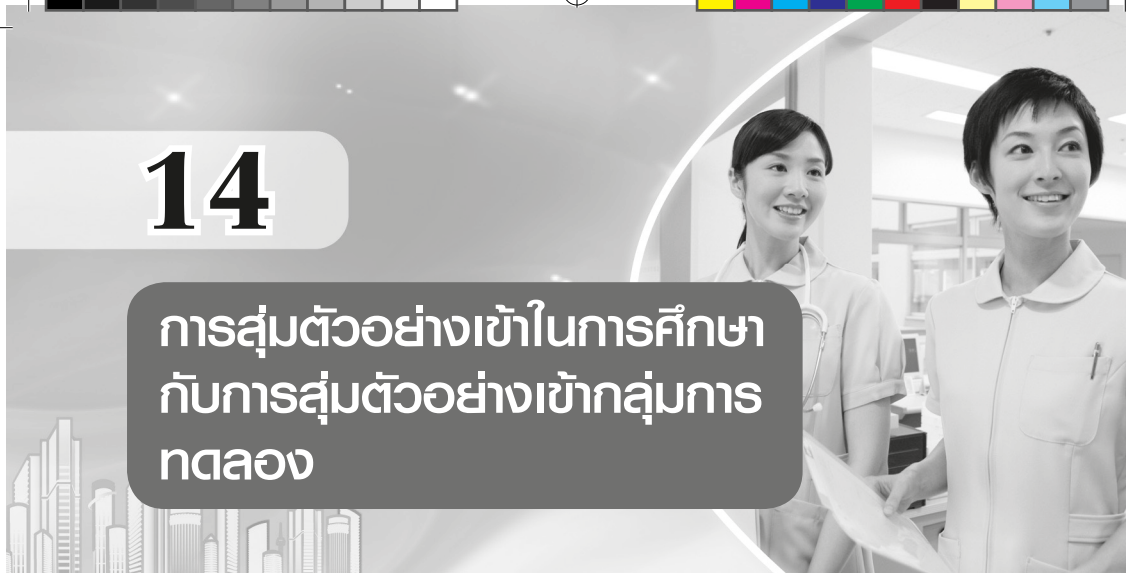
Using Statistics in Reseach 1_1 .indd 60 dummy 4 30/6/2557 11:29:29

มิลลิกรัมต่อเดซิลิตร และมีอาการที่เข้าได้กับโรคเบาหวาน ได้แก่ ปัสสาวะมาก ดื่มน้ำมาก น้ำหนักตัวลดลงโดยไม่ทราบสาเหตุ ดังนั้นการตรวจระดับน้ำตาลเพียงครั้งเดียวและไม่ได้รับการวินิจฉัยจากแพทย์อาจไม่ใช่ผู้ป่วยโรคเบาหวาน การหาปัจจัยเสี่ยงที่ศึกษาจากคนกลุ่มนี้จะไม่ถูกต้อง

สรุป

การศึกษาปัจจัยเสี่ยงนักวิจัยต้องเข้าใจเรื่องโรคที่ศึกษาและแบบงานวิจัยที่ใช้หาปัจจัยเสี่ยง วิธีการสุ่มแบบกลุ่มเหมาะสำหรับการหาความชุกเพื่อประเมินขนาดปัญหาในพื้นที่ สำหรับการตอบคำถามหลักเรื่องปัจจัยเสี่ยงเป็นเรื่องของตัวโรค คนเป็นโรค/ไม่เป็นโรคจึงไม่จำเป็นต้องสุ่มให้เป็นตัวแทนของพื้นที่ การสุ่มแบบกลุ่มจึงไม่มีผลต่อคุณภาพของการศึกษาที่ใช้หาปัจจัยเสี่ยง ถ้าใช้อาจทำให้การเก็บข้อมูลมีค่าใช้จ่ายมากขึ้น

Using Statistics in Reseach 1_1 .indd 61 dummy 4 30/6/2557 11:29:29



14

การสุ่มตัวอย่างเข้าในการศึกษา กับการสุ่มตัวอย่างเข้ากลุ่มการทดลอง

วิธีการสุ่มที่ใช้กับตัวอย่างในงานวิจัยเชิงทดลองทำใน 2 ขั้นตอน คือ การสุ่มตัวอย่างเข้าในการศึกษา (random sampling) กับการสุ่มตัวอย่างเข้ากลุ่มการทดลอง (random allocation หรือ random assignment หรือ randomization) นักวิจัยบางส่วนมีความเข้าใจสับสนหรือมีการใช้สลับกัน

การสุ่มอาสาสมัคร (ผู้เข้ารับการทดลอง) เช่น ผู้ป่วยเข้ามาเป็นตัวอย่างในการศึกษา เป็นการสุ่มตัวอย่างเพื่อให้ได้ตัวอย่างที่เป็นตัวแทนของประชากรศึกษา โดยมีจุดมุ่งหมายเพื่อป้องกันอคติในการเลือกตัวอย่าง และให้ได้ตัวอย่างเป็นตัวแทนที่ดีสำหรับอนุมานผลการศึกษาไปสู่ประชากรได้อย่างถูกต้อง (good external validity) การเลือกใช้วิธีการสุ่มแบบใดขึ้นอยู่กับลักษณะของการกระจายของตัวแปรผลหลักในประชากร อาจใช้วิธีการสุ่มแบบง่าย การสุ่มแบบชั้นภูมิ การสุ่มแบบระบบ การสุ่มแบบกลุ่ม เป็นต้น

การสุ่มตัวอย่างเข้ากลุ่มการทดลองที่มีหลายกลุ่ม การสุ่มนี้มีเฉพาะแบบงานวิจัยเชิงทดลองเท่านั้น เป็นการใช้วิธีสุ่มจัดสรรกลุ่มทดลองให้แก่อาสาสมัครแต่ละคน (หรือจัดสรรอาสาสมัครให้แก่กลุ่ม

ทดลองต่าง ๆ) เพื่อให้กลุ่มทดลองแต่ละกลุ่มมีอาสาสมัครที่มีลักษณะต่าง ๆเหมือนกัน ต่างกันเฉพาะสิ่งทดลอง (treatment) ที่ได้รับ ซึ่งจะช่วยให้สามารถแปลผลของสาเหตุ-ผลลัพธ์ (cause-effect) ของผลการทดลองได้อย่างถูกต้อง (good internal validity)

การใช้การสุ่มแบบง่ายจัดสรรกลุ่มทดลองให้แก่อาสาสมัครจะมีผลทำให้ไม่มีอคติในการจัดสรร แต่ไม่สามารถรับประกันว่ากลุ่มทดลองแต่ละกลุ่มจะมีอาสาสมัครที่มีลักษณะต่าง ๆเหมือนกัน ดังนั้นในกรณีการทดลองที่ต้องการให้ปัจจัยกวนที่สำคัญกระจายอยู่ในทุกกลุ่มเท่ากัน นักวิจัยต้องเลือกวิธีการสุ่มแบบชั้นภูมิ (stratified random sampling) การจะเลือกวิธีการสุ่มแบบใดจัดสรรกลุ่มทดลองให้แก่อาสาสมัครจะขึ้นอยู่กับลักษณะการกระจายของตัวแปรกวนของอาสาสมัครในประชากรและแบบงานวิจัย

นอกจากนี้ ถ้านักวิจัยต้องการให้มีอาสาสมัครกระจายอยู่ในแต่ละกลุ่มในทุก ๆช่วงเวลาในจำนวนที่ใกล้เคียงกัน จะต้องใช้วิธีสุ่มทีละชุด (block randomization) สุ่มกลุ่มทดลองให้แก่อาสาสมัคร

การสุ่มที่ใช้จัดสรรกลุ่มทดลองให้แก่อาสาสมัครเพื่อช่วยเพิ่มความถูกต้องของสรุปสาเหตุ-ผลลัพธ์แล้ว ยังช่วยลดอคติในการวัดผลโดยที่มีผู้รับผิดชอบจัดสรรกลุ่มทดลองให้แก่อาสาสมัครทำการสุ่มแบบปกปิดโดยไม่ให้ผู้ดำเนินการทดลองและผู้ประเมินผลลัพธ์รู้ว่าอาสาสมัครแต่ละคนได้รับสิ่งทดลองอะไร จะช่วยลดอคติในการวัดผลและเพิ่มความน่าเชื่อถือของผลวิจัย

Using Statistics in Reseach 1_1 .indd 63 dummy 4 27/8/2557 11:49:57

สรุป

งานวิจัยเชิงพรรณนา หรือแบบงานวิจัยเชิงวิเคราะห์ หรืองานวิจัยเชิงทดลองที่มีกลุ่มทดลองเพียงกลุ่มเดียว จะใช้วิธีการสุ่มอาสาสมัครเข้ามาเป็นตัวอย่างในการศึกษา ส่วนงานวิจัยเชิงทดลองที่มีกลุ่มทดลองมากกว่า 1 กลุ่ม นอกจากใช้วิธีการสุ่มอาสาสมัครเข้ามาเป็นตัวอย่างในการศึกษาแล้ว ยังต้องใช้วิธีสุ่มจัดสรรกลุ่มการทดลองให้แก่อาสาสมัครด้วย

Using Statistics in Reseach 1_1 .indd 64 dummy 4 30/6/2557 11:29:30

15

การปรับขนาดตัวอย่างเพิ่มสำหรับผู้ไม่ตอบแบบสำรวจ

การศึกษาเชิงพรรณนา/การสำรวจเพื่อประมาณค่าอุบัติการณ์ ความชุก ความครอบคลุม หรืออัตราต่างๆ ถ้าใช้วิธีสุ่มตัวอย่างแบบง่าย จะใช้สูตรการคำนวณขนาดตัวอย่างเพื่อประมาณค่าสัดส่วนของประชากร $n = \frac{Z^2p(1-p)}{e^2}$ ขนาดตัวอย่างที่คำนวณได้จะเป็นขนาดตัวอย่างที่น้อยที่สุดที่ใช้ในการศึกษา แต่ในทางปฏิบัติจะมีตัวอย่างที่สุ่มมาได้แต่เก็บข้อมูลไม่ได้ (non-response) อาจมีสาเหตุจากคนที่ถูกสุ่มเป็นตัวอย่างไม่อยู่ขณะเก็บข้อมูลหรือปฏิเสธจะตอบแบบสำรวจ ทำให้เก็บข้อมูลได้ไม่ครบตามจำนวนที่คำนวณไว้ ซึ่งอาจมีผลกระทบต่อความน่าเชื่อถือของผลการศึกษา

เพื่อแก้ไขปัญหาดังกล่าว นักวิจัยจำเป็นต้องมีการคำนวณขนาดตัวอย่างเพิ่มเผื่อไว้ให้มีจำนวนตัวอย่างที่เก็บได้เท่ากับหรือมากกว่าที่คำนวณได้ เช่น ถ้าคาดหมายว่าจะเก็บข้อมูลไม่ได้ร้อยละ 20 นักวิจัยต้องคำนวณขนาดตัวอย่างเพิ่มเผื่อไว้ เพื่อให้เวลาเก็บจริงได้เมื่อหักที่เก็บไม่ได้ร้อยละ 20 แล้วยังมีจำนวนตัวอย่างที่เก็บได้ครบตามจำนวนที่ต้องการ

อัตราการเก็บข้อมูลไม่ได้ควรกำหนดไว้มากน้อยเท่าไร อัตราดังกล่าวจะขึ้นอยู่กับวิธีที่ใช้ในการเก็บข้อมูล ความน่าสนใจของเรื่องที่ศึกษา และความยาวของแบบสอบถาม ถ้าเป็นการสำรวจทางไปรษณีย์หรือแบบสอบถาม อัตราการเก็บข้อมูลไม่ได้อาจจะอยู่ร้อยละ 35–60

กรณีเก็บข้อมูลโดยการสัมภาษณ์ อัตราการเก็บข้อมูลไม่ได้ควรจะไม่เกินร้อยละ 5 แต่ถ้าเลือกเวลาเก็บข้อมูลไม่เหมาะสม เช่น เก็บช่วงเวลาทำงาน คนตอบไม่อยู่บ้าน ถึงแม้จะเข้าไปสัมภาษณ์หลายรอบก็จะได้ไม่ครบ หรือกรณีในกรุงเทพฯ ถึงแม้จะไปเก็บวันหยุด แต่ผู้ให้สัมภาษณ์ต้องการพักผ่อนหรือไม่ว่างก็จะไม่ให้สัมภาษณ์ ผู้วิจัยควรจะกำหนดอัตราเก็บข้อมูลไม่ได้ไว้อย่างน้อยร้อยละ 10

นักวิจัยควรทบทวนงานวิจัยที่มีลักษณะคล้ายกัน ทำในกลุ่มเป้าหมายเดียวกัน ทำในช่วงระยะเวลาที่ใกล้กับช่วงเวลาที่นักวิจัยจะเก็บข้อมูลมากที่สุด ว่ามีอัตราการตอบแบบสัมภาษณ์หรืออัตราการตอบกลับที่ส่งคืนมาร้อยละเท่าไร เพื่อจะได้มีข้อมูลที่ใกล้เคียงที่สุดที่ใช้ในการปรับเพิ่มขนาดตัวอย่าง

วิธีการคำนวณเพื่อปรับขนาดตัวอย่างเพิ่มขึ้นที่พบมากคือ ผู้วิจัยคาดว่าจะมีผู้ที่เก็บข้อมูลไม่ได้ร้อยละ 20 เมื่อคำนวณขนาดตัวอย่างมาได้ 100 คน จะปรับเพิ่มอีกร้อยละ 20 เป็น 120 คน ($n_{adj} = n + n \times 20/100$) ซึ่งไม่ถูกต้อง ทั้งนี้เพราะถ้าวางแผนที่จะเก็บ 120 คน เก็บไม่ได้ร้อยละ 20 จำนวนที่จะเก็บได้จริงจะได้เพียง 96 คน [$(120 \times 80)/100 = 96$] ซึ่งน้อยกว่าที่คำนวณได้ 100 คน การปรับเพิ่มเพื่อทดแทนการเก็บข้อมูลไม่ได้จะต้องคำนวณโดยสูตร $n_{adj} = n/(1 - r)$ โดย r คือสัดส่วนการเก็บข้อมูลไม่ได้ เช่น นักวิจัยคาดว่าจะเก็บข้อมูลไม่ได้ร้อยละ 20 ค่าสัดส่วนคือ 0.2 จากตัวอย่างข้างต้น $n_{adj} = 100/(1 - 0.2) = 125$ คน ขนาดตัวอย่างดังกล่าวเมื่อหักจำนวนคนที่เก็บข้อมูลไม่ได้แล้วจะเหลือขนาดตัวอย่างเท่าที่ต้องการ 100 คน

Using Statistics in Reseach 1_1 .indd 66 dummy 4 8/9/2557 11:47:47

การคำนวณขนาดตัวอย่างเพิ่มเผื่อไว้มากจะดีหรือไม่ การเพิ่มเผื่อไว้มากนอกจากจะสิ้นเปลืองค่าใช้จ่ายแล้ว จะมีผลกระทบต่อคุณภาพของข้อมูลและความเป็นตัวแทนของประชากร เพราะคนที่เก็บไม่ได้อาจมีลักษณะแตกต่างจากคนที่เก็บได้ ถ้ามีมากจะทำให้ผลสรุปที่ได้มีอคติ (non-response bias) ซึ่งอาจต้องใช้วิธีการทางสถิติวิเคราะห์เพื่อลดอิทธิพลของอคติที่เกิดขึ้น

การลดจำนวนผู้เก็บข้อมูลไม่ได้สามารถแก้ไขได้โดยใช้แบบสอบถามสั้น (ควรใช้เวลาถาม-ตอบไม่เกิน 20 นาที) ข้อคำถามเข้าใจง่าย ตอบง่าย เลือกเวลาเก็บข้อมูลที่เหมาะสม และอาจมีของที่ระลึกขอบคุณที่ให้ข้อมูล

สรุป

ในการคำนวณขนาดตัวอย่างเพื่อการสำรวจ จำนวนตัวอย่างที่คำนวณได้จะเป็นขนาดตัวอย่างเล็กที่สุดที่ใช้ในการตอบคำถามหลักของงานวิจัย ในการเก็บข้อมูลจริงนักวิจัยต้องวางแผนวิธีเก็บข้อมูลให้มีอัตราการเก็บข้อมูลไม่ได้น้อยที่สุด และต้องคำนวณขนาดตัวอย่างเพิ่มเผื่อการเก็บข้อมูลไม่ได้

Using Statistics in Reseach 1_1 .indd 67 dummy 4 30/6/2557 11:29:30

16

การปรับขนาดตัวอย่างเพิ่มสำหรับการไม่ปฏิบัติตามแผนการทดลอง

ปัญหาหนึ่งที่พบได้เสมอในการวิจัยเชิงทดลอง ไม่ว่าจะเป็นการทดลองทางคลินิก การทดลองภาคสนาม หรือการทดลองวิธีการสอนที่ใช้แบบงานวิจัยแบบ clinical trial, true experiment หรือ quasi-experiment คือการมีอาสาสมัคร (ผู้เข้ารับการทดลอง) ออกระหว่างการทดลอง และมักจะพบว่านักวิจัยแก้ปัญหาโดยการตัดอาสาสมัครกลุ่มดังกล่าวออกจากตัวอย่างที่ศึกษา หรือกำหนดในเกณฑ์คัดออกว่าถ้าไม่สามารถเข้าร่วมการทดลองได้ครบจะไม่ใช่ประชากรในการศึกษา การแก้ปัญหาทั้ง 2 วิธีนี้เป็นการแก้ไขปัญหาที่ไม่ถูกต้อง

ในการทดลอง อาสาสมัครที่ออกระหว่างการทดลองโดยทั่วไปจะมีสาเหตุเกี่ยวข้องกับสิ่งทดลอง (intervention) เช่น มีผลข้างเคียงของยา รู้สึกไม่สบาย ทนไม่ได้กับอาการดังกล่าว หรือวิธีการสอนใช้เวลานาน มีจำนวนหลายครั้ง รู้สึกเบื่อและเสียเวลา เป็นต้น สาเหตุของการออกจากการทดลองส่วนใหญ่จะมีสาเหตุจากข้อเสียของสิ่งทดลอง ซึ่งควรนับเป็นส่วนหนึ่งของความล้มเหลวของสิ่งทดลองนั้น ดังนั้นจึงไม่สามารถตัดอาสาสมัครดังกล่าวออกจากการทดลอง หรือกำหนดเกณฑ์คัดออกจากประชากรศึกษา

การแก้ปัญหาอาสาสมัครออกระหว่างการทดลองต้องดำเนินการแก้ไขใน 2 ขั้นตอน คือ ในขั้นการดำเนินการทดลองและในขั้นการวิเคราะห์ ในขั้นการดำเนินการทดลองนักวิจัยต้องให้ข้อมูลต่างๆที่เกี่ยวข้องกับวิธีการทดลอง วิธีการปฏิบัติตัวของอาสาสมัครในแต่ละขั้นของการทดลอง เวลาที่ใช้ และผลข้างเคียงที่อาจเกิด ข้อมูลเหล่านี้จะช่วยทำให้อาสาสมัครเข้าใจบทบาทของตนเองในการทดลอง และเป็นข้อมูลประกอบการตัดสินใจว่ามีประโยชน์ หรือมีเวลาพอที่จะเข้าร่วมในการทดลองหรือไม่

ถึงแม้จะมีการให้ข้อมูลต่างๆอย่างชัดเจน และอาสาสมัครตัดสินใจเข้าร่วมด้วยความเต็มใจ แต่เมื่อทำการทดลองจริงยังคงพบว่ามีอาสาสมัครจำนวนหนึ่งออกระหว่างการทดลอง เมื่อมีอาสาสมัครออกจากการทดลองทำให้ขนาดตัวอย่างลดลง ขนาดตัวอย่างศึกษาที่มีขนาดเล็กลงจะทำให้อำนาจการทดสอบต่ำกว่าที่กำหนดไว้ ซึ่งมีผลทำให้การทดลองอาจมีจำนวนตัวอย่างไม่พอที่จะระบุความต่างระหว่างกลุ่ม นักวิจัยจึงต้องคำนวณขนาดตัวอย่างเพิ่มเผื่อไว้ทดแทนอาสาสมัครที่คาดว่าจะออกระหว่างการทดลอง เนื่องจากอาสาสมัครมีโอกาสออกจากการทดลองได้ทั้งกลุ่มทดลองและกลุ่มควบคุม จำต้องเพิ่มเผื่อไว้ทั้ง 2 กลุ่มในจำนวนที่เท่าๆกัน

สิ่งที่ต้องคำนึงถึงในการปรับขนาดตัวอย่างเพื่อทดแทนอาสาสมัครที่คาดว่าจะออกระหว่างการทดลอง นอกจากมีขนาดตัวอย่างน้อยลงแล้ว ยังทำให้ขนาดตัวอย่างแต่ละกลุ่มแตกต่างกัน Lachin* ได้นำเสนอสูตรง่ายๆสำหรับใช้ในการปรับขนาดตัวอย่างไว้ดังนี้

---

* John M. Lachin, "Introduction to Sample Size Determination and Power Analysis for Clinical Trials." **Controlled Clinical Trials.** 1981; 2: 93–113.

Using Statistics in Reseach 1_1 .indd 69 dummy 4 30/6/2557 11:29:30

$$n_{adj} = \frac{n_0}{(1 - r)^2}$$

โดยที่ n_{adj} = ขนาดตัวอย่างที่ปรับแล้ว
n_0 = ขนาดตัวอย่างก่อนปรับ
r = สัดส่วนที่คาดว่าจะออกระหว่างการทดลอง

ตัวอย่าง ในการทดลองเพื่อเปรียบเทียบวิธีการสอน 2 วิธี ขนาดตัวอย่างที่คำนวณได้คือ 35 คน/กลุ่ม นักวิจัยคาดว่าจะมีผู้ที่ออกระหว่างการทดลองร้อยละ 20 (r = 0.2) คำนวณขนาดตัวอย่างที่ปรับค่าได้ดังนี้

$$n_{adj} = \frac{35}{(1 - 0.2)^2} = 55 \text{ คน/กลุ่ม}$$

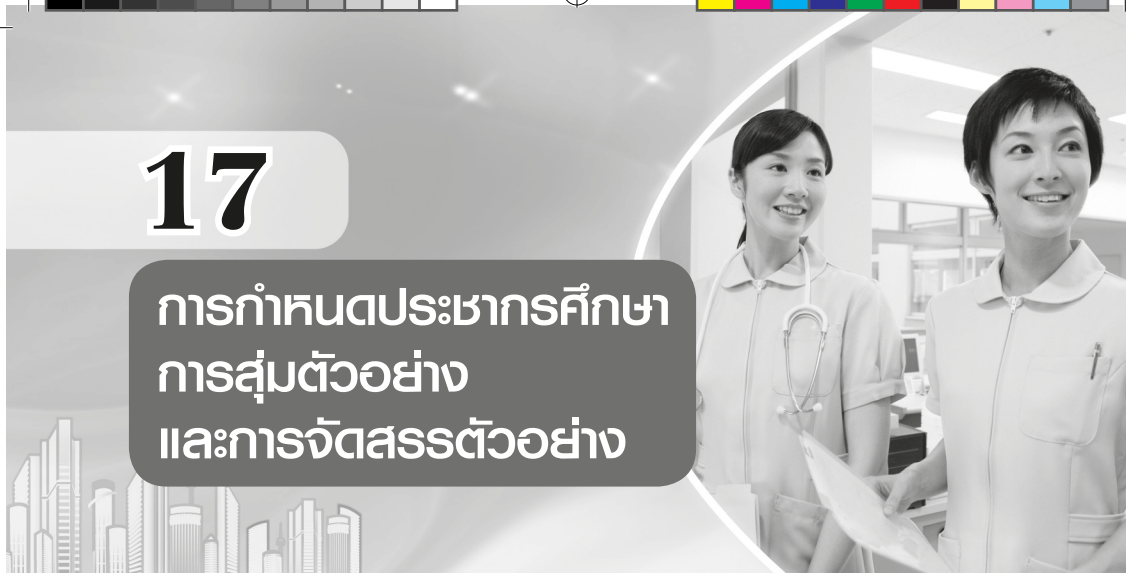
ในการทดลองถ้าพบว่ามีอาสาสมัครออกระหว่างการทดลอง นักวิจัยจะตัดทิ้งไม่นำมาวิเคราะห์ไม่ได้ อาสาสมัครทั้งหมดที่นำเข้ามาศึกษาตั้งแต่เริ่มต้นจะต้องนำเข้าร่วมในการวิเคราะห์ ถึงแม้มีบางคนอยู่ไม่ครบหรือไม่สามารถวัดผลหลังการทดลอง (post test) วิธีการวิเคราะห์ผลการทดลองในกรณีที่มีอาสาสมัครออกระหว่างการทดลองมี 2 วิธี คือ วิเคราะห์แบบ intention to treat analysis (ITT) และการวิเคราะห์แบบ per protocol analysis (PP) สำหรับวิธีการวิเคราะห์ดังกล่าวจะได้อธิบายในเรื่องที่ 33 *การวิเคราะห์ผลการทดลองในกรณีมีผู้ออกระหว่างการทดลอง*

Using Statistics in Reseach 1_1 .indd 70 dummy 4 27/8/2557 11:51:33

สรุป

ในงานวิจัยเชิงทดลองเมื่อคาดว่าจะมีอาสาสมัครออกระหว่างการทดลอง นักวิจัยต้องปรับขนาดตัวอย่างเพิ่มเผื่อไว้ทดแทนอาสาสมัครที่คาดว่าจะออกระหว่างการทดลอง ถ้ามีอาสาสมัครออกระหว่างการทดลองจะตัดอาสาสมัครกลุ่มนี้ออกจากการทดลองไม่ได้ ต้องนำอาสาสมัครทั้งหมดที่เข้าร่วมในการศึกษาตั้งแต่ต้นมาร่วมในการวิเคราะห์สรุปผล

Using Statistics in Reseach 1_1 .indd 71 dummy 4 30/6/2557 11:29:30



17

การกำหนดประชากรศึกษา การสุ่มตัวอย่าง และการจัดสรรตัวอย่าง

ในการนำเสนอผลงานวิจัยแบบกึ่งทดลองและบทความวิจัยที่ตีพิมพ์ในภาษาไทยพบว่า การกำหนดเกณฑ์คัดเข้า เกณฑ์คัดออก การสุ่มตัวอย่าง และการจัดสรรตัวอย่างให้กลุ่มการทดลอง มีการใช้สับสนและเข้าใจความหมายไม่ถูกต้อง

การกำหนดเกณฑ์คัดเข้า เกณฑ์คัดออก เป็นวิธีการกำหนดประชากรศึกษา (study population) หรือการกำหนดคุณสมบัติของประชากรที่จะถูกเลือกมาศึกษา ไม่ใช่เป็นเกณฑ์หรือวิธีที่ใช้เลือกตัวอย่าง (sampling method)

การเลือกตัวอย่างหรือการสุ่มตัวอย่าง (sampling) เป็นวิธีการที่ใช้เลือกอาสาสมัคร (เข้ารับการทดลอง) บางส่วนจากประชากรศึกษามาเป็นตัวอย่างในการศึกษา การสุ่มตัวอย่างแบ่งได้เป็น 2 ประเภท คือ

1) **การสุ่มตัวอย่างแบบใช้ความน่าจะเป็น** เช่น การสุ่มแบบง่าย การสุ่มแบบระบบ การสุ่มแบบชั้นภูมิ และการสุ่มแบบกลุ่ม เป็นต้น

2) **การสุ่มตัวอย่างแบบไม่ใช้ความน่าจะเป็น** เช่น การสุ่มแบบบังเอิญ การสุ่มแบบโควตา การสุ่มแบบเจาะจง เป็นต้น ซึ่งการสุ่มแบบนี้จะไม่สามารถคำนวณความน่าจะเป็นของการเกิดผลลัพธ์ จึงใช้สถิติในการทดสอบสมมติฐานและประมาณค่าผลลัพธ์ของประชากรไม่ได้

การจัดสรรตัวอย่างให้กลุ่มทดลอง หรือในการทดลองทางคลินิกนิยมเรียกว่า **การจัดสรรกลุ่มทดลองให้แก่อาสาสมัคร** ในการวิจัยเชิงทดลอง เมื่อสุ่มอาสาสมัครที่จะนำเข้ามาเป็นตัวอย่างได้ครบแล้ว การที่จะจัดสรรว่าอาสาสมัครแต่ละรายจะได้รับสิ่งทดลอง (treatment) แบบใด หรืออยู่ในกลุ่มทดลองใดทำได้ 2 วิธี คือ การจัดสรรแบบสุ่ม (randomization) ที่ใช้วิธีการสุ่มกลุ่มทดลองให้แก่อาสาสมัครแต่ละราย และการจัดสรรแบบไม่สุ่ม (non-randomization) ที่กำหนดเป็นเกณฑ์ในการจัดสรร เช่น ถ้าผู้ป่วยเลือกวิธีใดก็ให้อยู่ในกลุ่มนั้น หรือการจัดสรรที่กำหนดให้อาสาสมัครที่รับเข้ามาในช่วงแรกอยู่ในกลุ่มควบคุมจนครบแล้วจึงให้อาสาสมัครที่รับเข้าต่อๆมาเป็นของกลุ่มทดลอง เป็นต้น

ความผิดพลาดที่พบบ่อยคือ การที่นักวิจัยเข้าใจผิดว่าเกณฑ์คัดเข้า เกณฑ์คัดออก เป็นเกณฑ์ที่ใช้คัดเลือกตัวอย่าง จึงมักระบุว่าวิธีที่ใช้เลือกตัวอย่างเป็นแบบเจาะจง ในความเป็นจริงถ้าผู้วิจัยนำอาสาสมัครตามเกณฑ์ทุกรายที่เข้ามารับการรักษาในโรงพยาบาล ณ ช่วงเวลาที่กำหนดมาเป็นตัวอย่างในการศึกษา จะถือว่าเป็นการเลือกตัวอย่างแบบสุ่ม ในกรณีที่อาสาสมัครที่มีคุณสมบัติตามเกณฑ์ในช่วงเวลาเก็บข้อมูลมีจำนวนมาก ไม่สามารถนำเข้าศึกษาได้ทุกราย นักวิจัยต้องกำหนดวิธีการสุ่มตัวอย่างอาสาสมัครบางส่วนเข้ามาเป็นตัวอย่างในการศึกษา

Using Statistics in Reseach 1_1 .indd 73 dummy 4 30/6/2557 11:29:31

ตัวอย่าง ในการศึกษาเปรียบเทียบวิธีการสอนการออกกำลังกายให้แก่ผู้สูงอายุที่เข้ามารับการรักษาในโรงพยาบาลมอดินแดง ประชากรศึกษาเขียนได้ 2 แบบ ดังนี้

แบบที่ 1 เขียนเป็นแบบร้อยแก้ว *"ประชากรศึกษาคือผู้ป่วยนอกของโรงพยาบาลมอดินแดง ที่มีอายุ 60 ปีขึ้นไป ที่สามารถสื่อสารได้ และเข้ามารับการรักษาที่โรงพยาบาลในเดือนมกราคม พ.ศ. 2555"*

แบบที่ 2 เขียนเป็นเกณฑ์คัดเลือกประชากรศึกษา ดังนี้

เกณฑ์คัดเข้า

1) เป็นผู้ป่วยนอกของโรงพยาบาลมอดินแดง
2) เข้ามารับการรักษาในเดือนมกราคม พ.ศ. 2555
3) อายุ 60 ปีขึ้นไป

เกณฑ์คัดออก

1) ไม่ยินยอมเข้าร่วมในงานวิจัย
2) ไม่สามารถสื่อสารได้

จากประชากรศึกษาที่กำหนด ถ้าผู้วิจัยพบว่าในแต่ละวันมีผู้สูงอายุที่มีคุณสมบัติตามเกณฑ์ประมาณ 50 คน แต่นักวิจัยมีเวลาเก็บและทำการสอนได้วันละ 5 คน ในกรณีนี้นักวิจัยต้องกำหนดวิธีการสุ่มอาสาสมัครว่าจะสุ่ม 5 คนจาก 50 คนด้วยวิธีใด เช่น ใช้การสุ่มแบบง่าย หรือการสุ่มแบบระบบ

Using Statistics in Reseach 1_1 .indd 74 dummy 4 30/6/2557 11:29:31

สำหรับการจัดสรรกลุ่มทดลองให้แก่หน่วยตัวอย่าง วิธีการจัดสรรที่ดีจะช่วยลดอคติในการทดลอง เช่น เมื่อสุ่มแล้วอาสาสมัครแต่ละกลุ่มมีจำนวนอาสาสมัครที่มีอาการหนักเบาของโรค หรือปัจจัยเสี่ยงต่าง ๆที่มีอิทธิพลต่อผลลัพธ์การทดลองกระจายอยู่ในแต่ละกลุ่มไม่เท่ากันจะมีผลต่อการเกิดอคติในการสรุปผลการทดลอง ถ้านักวิจัยต้องการให้ปัจจัยกวนที่สำคัญกระจายอยู่ในแต่ละกลุ่มเท่า ๆกันจะต้องใช้วิธีการสุ่มตัวอย่างแบบชั้นภูมิ (stratified random sampling) ในกรณีที่ต้องการให้ทุก ๆช่วงเวลาการศึกษามีจำนวนอาสาสมัครในแต่ละกลุ่มมีจำนวนพอ ๆกันจะใช้วิธีสุ่มเป็นชุด (block randomization)

สรุป

การกำหนดเกณฑ์นำเข้า เกณฑ์คัดออก เป็นการกำหนดลักษณะประชากรศึกษา ไม่ใช่เกณฑ์เลือกตัวอย่าง ศัพท์คำว่า **"สุ่ม"** มี 2 ความหมาย ความหมายแรกหมายถึงการเลือก (การสุ่ม) ตัวอย่าง (sampling) และความหมายที่สองหมายถึงแบบสุ่ม (random) ดังนั้นการสุ่มตัวอย่างเป็นวิธีการที่นักวิจัยใช้เลือกอาสาสมัครจากประชากรศึกษามาเป็นตัวอย่าง ซึ่งทำได้ทั้งแบบใช้ความน่าจะเป็น (random sampling) และไม่ใช้ความน่าจะเป็น

การจัดสรรกลุ่มทดลองให้แก่อาสาสมัครแต่ละรายจะลดอคติในการทดลอง มีความสำคัญต่อความถูกต้องของผลการศึกษา นักวิจัยต้องพิจารณาปัจจัยที่มีอิทธิพลต่อผลลัพธ์ของการทดลอง แล้วเลือกวิธีการสุ่มที่เหมาะสมในการจัดสรรกลุ่มทดลองให้แก่อาสาสมัคร

Using Statistics in Reseach 1_1 .indd 75 dummy 4 30/6/2557 11:29:31

18

ความหมายของ univariate & multivariate และการเขียนแผนการวิเคราะห์

การใช้คำว่า univariate ในการเรียกวิธีการวิเคราะห์ข้อมูลในรายงานวิจัยมี 2 ลักษณะ ลักษณะแรกใช้บอกวิธีการวิเคราะห์เพื่อพรรณนาลักษณะของตัวอย่างที่ศึกษาโดยการวิเคราะห์ทีละตัวแปร และลักษณะที่สองในการวิเคราะห์ความสัมพันธ์พหุตัวแปร โดยเริ่มต้นดูความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระแต่ละตัวก็เรียกว่า univariate เช่นกันทั้งที่มีจำนวนตัวแปรในการวิเคราะห์ 2 ตัว

คำว่า univariate analysis เป็นคำเรียกวิธีการวิเคราะห์ว่าจะวิเคราะห์ข้อมูลทีละตัวแปร ดังนั้นในกรณีแรกที่ระบุว่าจะใช้วิธีการวิเคราะห์เพื่อพรรณนาลักษณะของตัวอย่างที่ศึกษา โดยจะดำเนินการวิเคราะห์ข้อมูลครั้งละ 1 ตัวแปรแต่ไม่รู้ว่าจะใช้สถิติอะไรวิเคราะห์

ส่วนในกรณีที่สอง ในขั้นตอนแรกของการวิเคราะห์ความสัมพันธ์พหุตัวแปรที่วิเคราะห์ดูความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระทีละตัวจะเรียกว่า univariate analysis เช่นกัน โดยหมายถึงดำเนินการวิเคราะห์ความสัมพันธ์ของตัวแปรอิสระทีละตัวแปรแต่ไม่รู้ว่าจะใช้สถิติอะไรวิเคราะห์เช่นกัน

ดังนั้นทั้ง 2 กรณีใช้ univariate analysis เรียกวิธีการวิเคราะห์ว่าจะวิเคราะห์ทีละตัวแปร (ไม่ใช่ครั้งละ 1 ตัวแปร) โดยกรณีแรกแต่ละครั้งมีตัวแปร 1 ตัว ส่วนกรณีที่สองแต่ละครั้งจะมีตัวแปรอิสระ 1 ตัว

ส่วน bivariate มาจาก bivariate data เดิมหมายถึงตัวแปรต่อเนื่อง 2 ตัวที่ค่าของตัวแปรทั้งสองแปรตามกัน เช่น น้ำหนัก ส่วนสูง คำนี้สร้างขึ้นในสมัยเริ่มพัฒนาสถิติในการวัดความสัมพันธ์ คำว่า bivariate analysis จึงหมายถึงการวัดความสัมพันธ์ของ bivariate data ปัจจุบันมีการใช้เรียกวิธีการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปร 2 ตัว

การเขียนอธิบายวิธีการวิเคราะห์ในบทความวิจัยว่าจะดำเนินการวิเคราะห์ข้อมูลทีละตัว (univariate) หาความสัมพันธ์ระหว่าง 2 ตัวแปร (bivariate) หรือทีละพหุตัวแปร (multivariate) ผู้อ่านหรือผู้พิจารณาบทความไม่สามารถใช้พิจารณาความเหมาะสมของวิธีการวิเคราะห์ทางสถิติในบทความดังกล่าวได้ เพราะไม่ได้บอกว่าจะใช้สถิติอะไรวิเคราะห์

การเขียนอธิบายวิธีการวิเคราะห์ในรายงานวิจัยจะต้องระบุวิธีการทางสถิติที่ใช้ในการวิเคราะห์ตัวแปรที่สอดคล้องกับวัตถุประสงค์ของการศึกษานั้น ไม่ใช้อธิบายหลักการกว้างๆ ตัวอย่างเช่น

ไม่เหมาะสม: *"การพรรณนาลักษณะตัวอย่างใช้ univariate analysis วิเคราะห์เพศ อายุ BMI และภาวะ metabolic syndrome (MS) ด้วยสถิติพรรณนา"*

เหมาะสม: *"การพรรณนาเพศและภาวะ MS ใช้การแจกแจงความถี่และร้อยละ ส่วนอายุและ BMI คำนวณค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และนำข้อมูลมาจัดกลุ่มตามเกณฑ์ที่กำหนด แล้วนำเสนอด้วยตารางความถี่และร้อยละ"*

ไม่เหมาะสม: *"ใช้ bivariate analysis วิเคราะห์ความสัมพันธ์ระหว่างเพศกับ BMI เพศกับภาวะ MS และกลุ่มอายุกับภาวะ MS"*

Using Statistics in Reseach 1_1 .indd 77 dummy 4 30/6/2557 11:29:31

เหมาะสม: *"ใช้สถิติ chi-square วิเคราะห์ความสัมพันธ์ระหว่างภาวะ MS กับเพศและกลุ่มอายุ ในการเปรียบเทียบค่าเฉลี่ยของ BMI ระหว่างเพศด้วย t-test ในกรณีที่ผลการตรวจสอบพบว่าการแจกแจงของ BMI ไม่ปกติจะทำการแปลงข้อมูลให้เป็นแบบปกติก่อนดำเนินการทดสอบ"*

ไม่เหมาะสม: *"วิเคราะห์ความเสี่ยงของการเกิดภาวะ MS ด้วย multivariate analysis"*

เหมาะสม: *"วิเคราะห์ความเสี่ยงของการเกิดภาวะ MS โดยคำนวณค่า OR ของปัจจัยแต่ละตัว แล้วคำนวณค่า adjusted OR โดยใช้สถิติ multiple logistic regression"*

จากตัวอย่างจะเห็นได้ว่าการระบุว่าวิเคราะห์ตัวแปรทีละตัว ทีละ 2 ตัว หรือทีละหลายตัว ไม่เหมาะสมที่จะนำเสนอในโครงร่างและในรายงานวิจัย เพราะไม่มีรายละเอียดของวิธีการทางสถิติที่ใช้ในการวิเคราะห์

การเขียนอธิบายวิธีการวิเคราะห์ข้อมูลในรายงานวิจัยจะต้องระบุวิธีการทางสถิติที่ใช้ในการสรุปผล หรือเปรียบเทียบผลสำหรับวัตถุประสงค์หลักและวัตถุประสงค์รอง รวมทั้งวิธีการทางสถิติอื่นๆที่ใช้ เช่น การวิเคราะห์กลุ่มย่อย การควบคุมตัวแปรกวน การจัดการกับข้อมูลสูญหาย หรือวิธีการแก้ไขกรณีที่ข้อมูลไม่เป็นไปตามข้อตกลงเบื้องต้น ฯลฯ

ในงานทดลองทางคลินิก การวิเคราะห์ผลการศึกษาตามวัตถุประสงค์ต้องใช้วิธีการทางสถิติที่ระบุอยู่ในแผนการวิเคราะห์ในโครงร่างวิจัย (ที่ได้ลงทะเบียนไว้) ถ้าใช้สถิติอื่นมาใช้วิเคราะห์แทนจะไม่ได้รับการยอมรับและไม่ให้ลงตีพิมพ์ในวารสาร ทั้งนี้ เพื่อป้องกันอคติที่เกิดจากนักวิจัยพยายามหาวิธีการวิเคราะห์เพื่อให้ได้ผลการวิเคราะห์ที่ต้องการ

Using Statistics in Reseach 1_1 .indd 78 dummy 4 27/8/2557 11:51:33

สรุป

ความหมายของการดำเนินการวิเคราะห์ข้อมูลทีละตัว ทีละ 2 ตัว หรือพหุตัว มีความเข้าใจสับสน และไม่เหมาะสมที่จะใช้คำดังกล่าวเขียนวิธีการวิเคราะห์ในโครงร่างหรือบทความวิจัย เพราะไม่สามารถใช้พิจารณาความเหมาะสมของวิธีการวิเคราะห์ การเขียนวิเคราะห์ในโครงร่างวิจัยและรายงานวิจัยไม่ควรอธิบายหลักการกว้างๆ ควรระบุวิธีการทางสถิติและสถิติที่ใช้ในการวิเคราะห์จริงในบทความเท่านั้น

Using Statistics in Reseach 1_1 .indd 79 dummy 4 27/8/2557 11:51:33

19

การตั้งวัตถุประสงค์งานวิจัย กับการวิเคราะห์ข้อมูล

การตั้งวัตถุประสงค์งานวิจัยที่ดีจะช่วยให้สามารถวิเคราะห์สรุปผลได้ตรงกับการนำไปใช้งาน ปัญหาที่พบบ่อยๆคือ เวลานักวิจัยเขียนวัตถุประสงค์งานวิจัยโดยยังไม่ได้คิดว่าจะสรุปคำตอบไปใช้งานอย่างไร ทำให้การวิเคราะห์ได้คำตอบไม่สอดคล้องกับที่จะนำผลสรุปจากงานวิจัยไปใช้แก้ปัญหา

ตัวอย่าง การศึกษาคุณภาพชีวิตของพ่อ/แม่ที่เป็นผู้ดูแลลูกที่เป็นอัมพาตจากอุบัติเหตุจากรถจักรยานยนต์และรถยนต์ ว่ามีจำนวนผู้ที่มีคุณภาพชีวิตไม่ดีมากเท่าไร ข้อมูลที่ได้จะใช้ในการพิจารณาจัดตั้งเป็นคลินิกให้คำปรึกษาเรื่องการปรับตัวและการดูแลลูกที่เป็นอัมพาต โดยมีวัตถุประสงค์การศึกษาดังนี้ *“เพื่อศึกษาคุณภาพชีวิตของพ่อ/แม่ที่ดูแลลูกที่เป็นอัมพาตจากอุบัติเหตุรถจักรยานยนต์/รถยนต์”*

การประเมินคุณภาพชีวิตนักวิจัยใช้แบบประเมิน WHOQOL-BREF ฉบับภาษาไทย ซึ่งประเมินคุณภาพชีวิตออกเป็น 4 ด้าน คือ ด้านร่างกาย ด้านจิตใจ ด้าน

สัมพันธภาพทางสังคม และด้านสิ่งแวดล้อม ผลรวมคะแนนแต่ละด้านและคะแนนรวมทุกด้านมีเกณฑ์แบ่งคะแนนคุณภาพชีวิตเป็น 3 ระดับ คือ คุณภาพชีวิตที่ไม่ดี คุณภาพชีวิตกลางๆ และคุณภาพชีวิตที่ดี

ในการวิเคราะห์ข้อมูลเพื่อตอบวัตถุประสงค์ ถ้านักวิจัยนำคะแนนคุณภาพชีวิตจากตัวอย่างทั้งหมดมาคำนวณหาค่าเฉลี่ย เพื่อแสดงว่าคุณภาพชีวิตพ่อ/แม่ที่เป็นผู้ดูแลลูกที่เป็นอัมพาตโดยเฉลี่ยมีคุณภาพชีวิตอยู่ในระดับใด จากแผนการวิเคราะห์ดังกล่าว เมื่อนำข้อมูลจากตัวอย่างทั้งหมดมาคำนวณค่าเฉลี่ยแล้วนำไปเทียบกับเกณฑ์พบว่าอยู่ในระดับกลางๆ นักวิจัยจึงสรุปว่าพ่อ/แม่ที่ดูแลลูกที่เป็นอัมพาตมีคุณภาพชีวิตกลางๆ การสรุปดังกล่าวไม่ถูกต้อง ทั้งนี้เพราะค่าเฉลี่ยที่ได้อาจมาจากพ่อ/แม่บางส่วนที่มีคุณภาพชีวิตไม่ดี และบางส่วนมีคุณภาพชีวิตดี ค่าเฉลี่ยจึงอยู่กลางๆ การใช้ค่าเฉลี่ยสรุปผลจึงไม่สามารถระบุว่าพ่อ/แม่ที่ดูแลลูกที่เป็นอัมพาตที่มีคุณภาพชีวิตไม่ดีมีอยู่มากน้อยเท่าใด

ในการตั้งวัตถุประสงค์ให้สามารถสรุปผลที่จะนำไปใช้ทำข้อเสนอแนะสำหรับการแก้ไขปัญหาได้ ควรคิดตั้งแต่แรกว่า เมื่อได้ผลจากงานวิจัยแล้วจะให้ข้อเสนอแนะอย่างไร ในกรณีตัวอย่างนี้ควรตั้งวัตถุประสงค์ว่า *“เพื่ออธิบายจำนวนพ่อ/แม่ที่ดูแลลูกที่เป็นอัมพาตจากอุบัติเหตุรถจักรยานยนต์/รถยนต์ที่มีคุณภาพชีวิตไม่ดี”* สำหรับการวิเคราะห์ควรนำข้อมูลคุณภาพชีวิตของพ่อ/แม่แต่ละคนมาจัดกลุ่มตามเกณฑ์ก่อนนำไปวิเคราะห์ จึงจะสามารถวิเคราะห์ออกมาได้ว่ามีพ่อ/แม่ที่มีคุณภาพชีวิตไม่ดีเท่าไร ในกรณีนักวิจัยคิดว่าการแก้ปัญหาแต่ละด้านซึ่งดำเนินการไม่เหมือนกัน ควรมี

Using Statistics in Reseach 1_1 .indd 81 dummy 4 30/6/2557 11:29:31

แผนการวิเคราะห์กลุ่มที่มีคุณภาพชีวิตที่ไม่ดีย่อยเป็นรายด้าน จากตัวอย่างจะเห็นได้ว่าการตั้งวัตถุประสงค์และกำหนดแผนการวิเคราะห์ที่ดีนำไปใช้แก้ปัญหาได้ นักวิจัยต้องเข้าใจปัญหาและลักษณะคำตอบที่ต้องการนำไปใช้แก้ปัญหา

มีคำถามว่าเมื่อเก็บข้อมูลคุณภาพชีวิตมาแล้วพบว่าคำนวณค่าเฉลี่ยแล้วตอบคำถามไม่ได้ก็เปลี่ยนวิธีวิเคราะห์โดยนำมาจัดกลุ่มตามเกณฑ์ก่อนนำไปวิเคราะห์ไม่ได้หรือ ในการประมาณค่าสัดส่วนของพ่อ/แม่ที่คุณภาพชีวิตไม่ดีจะใช้ขนาดตัวอย่างจำนวนมากกว่าการประมาณค่าเฉลี่ยในการวิเคราะห์ที่มีการคำนวณตัวอย่างเพื่อประมาณค่าเฉลี่ย แต่ในการวิเคราะห์เปลี่ยนเป็นการประมาณค่าสัดส่วนจะทำให้มีขนาดตัวอย่างไม่มากพอ มีผลทำให้ผลสรุปที่ได้มีความเที่ยงของการประมาณค่าลดลง

งานวิจัยที่ต้องการคำตอบไปใช้แก้ปัญหาส่วนใหญ่ต้องมีการแบ่งข้อมูลออกเป็นกลุ่มเพื่อให้สามารถนำไปวิเคราะห์หาปริมาณกลุ่มเป้าหมายได้ และเมื่อนำไปหาความสัมพันธ์กับปัจจัยก็สามารถระบุได้ชัดเจนว่าปัจจัยใดที่มีผลต่อระดับคุณภาพชีวิตที่ไม่ดี

สำหรับค่าเฉลี่ยให้ภาพกว้างของทั้งกลุ่มโดยไม่สามารถแยกแยะให้เห็นกลุ่มที่มีปัญหา ข้อมูลที่ได้จากการหาความสัมพันธ์จะระบุได้แต่เพียงว่าปัจจัยใดมีความสัมพันธ์กับคุณภาพชีวิตของพ่อแม่ ข้อค้นพบใช้ยืนยันว่าได้ผลเป็นไปตามทฤษฎี ไม่สามารถนำไปใช้แก้ปัญหาได้โดยตรง

Using Statistics in Reseach 1_1 .indd 82 dummy 4 27/8/2557 11:51:33

สรุป

การตั้งวัตถุประสงค์งานวิจัยเพื่อให้สามารถสรุปผลได้สอดคล้องกับการใช้งาน นักวิจัยต้องมีวิธีดำเนินการเป็นลำดับดังนี้

1) เมื่อได้ผลการวิจัยแล้วจะให้ข้อเสนอในการแก้ปัญหาอย่างไร
2) จะวิเคราะห์อย่างไรจึงจะได้ข้อสรุปสอดคล้องกับการนำไปใช้งานตามข้อ 1) และ
3) เลือกใช้สถิติและวิธีการคำนวณขนาดตัวอย่างให้สอดคล้องกับวิธีการวิเคราะห์

Using Statistics in Reseach 1_1 .indd 83 dummy 4 30/6/2557 11:29:32

20

การเขียนวิธีการวิเคราะห์ทางสถิติในโครงร่างวิจัย

ในการพิจารณาคุณภาพของโครงร่างวิจัย วิธีการวิเคราะห์ทางสถิติ (statistical analysis) จะเป็นหัวข้อหนึ่งที่ต้องนำมาพิจารณา จากแผนการวิเคราะห์ข้อมูล ผู้พิจารณาอนุมัติทุนและกรรมการจริยธรรมการวิจัยจะพิจารณาความเหมาะสมของการใช้สถิติ เพื่อประเมินว่านักวิจัยมีความสามารถทางด้านสถิติเพียงพอที่จะวิเคราะห์ข้อมูลในโครงการวิจัยที่ขอมาได้อย่างถูกต้องหรือไม่

การเขียนแผนการวิเคราะห์ทางสถิติไม่เหมาะสมที่พบมากมักจะเขียนว่า *"วิเคราะห์ข้อมูลเชิงปริมาณด้วยค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ค่ามัธยฐาน ทดสอบด้วยสถิติ t-test ข้อมูลกลุ่มวิเคราะห์ด้วยร้อยละ ทดสอบด้วยสถิติ χ^2 และ Fisher exact test"* จากตัวอย่างการเขียนดังกล่าวเป็นการอธิบายหลักการทางสถิติกว้างๆ ไม่เฉพาะสำหรับโครงการนี้ การเขียนแบบนี้จะไม่สามารถนำมาประเมินได้ว่านักวิจัยจะวิเคราะห์สรุปคำตอบของแต่ละวัตถุประสงค์อย่างไร และถ้าเกิดปัญหาในการวิจัย เช่น เก็บข้อมูลได้ไม่ครบ จะแก้ไขในขั้นตอนการวิเคราะห์อย่างไร ถ้าข้อมูลจากตัวอย่างมีการแจกแจงไม่เป็นไปตามข้อตกลงเบื้องต้นในการคำนวณค่าสถิติ นักวิจัยจะแก้ไขปัญหาอย่างไร

หลักในการเขียนวิธีวิเคราะห์ทางสถิติที่เหมาะสม มีดังนี้

1) ควรนำเสนอเฉพาะสถิติที่มีแผนที่จะใช้เท่านั้น
2) ต้องระบุสถิติที่ใช้วิเคราะห์คำตอบ (วัตถุประสงค์) หลักให้ชัดเจน โดยเฉพาะการทดลองทางคลินิก ถ้าใช้วิธีสถิติวิเคราะห์คำตอบหลักในรายงานไม่ตรงกับที่มีอยู่ในแผนการวิเคราะห์ของโครงร่างวิจัยจะไม่ได้รับการตีพิมพ์
3) เขียนชื่อสถิติที่ใช้ให้ชัดเจน เช่น t-test หรือ paired t-test
4) ควรอธิบายว่าจะตรวจสอบข้อตกลงเบื้องต้นสำหรับการทดสอบสมมติฐานและการประมาณค่าอย่างไร
5) ในกรณีของการเปรียบเทียบ ถ้าพบความไม่สมดุลของปัจจัยกวนระหว่างกลุ่มทดลองและกลุ่มควบคุมจะแก้ไขในการวิเคราะห์อย่างไร
6) ในกรณีมีข้อมูลสูญหายจะมีวิธีการวิเคราะห์อย่างไร
7) ถ้ามีแผนการวิเคราะห์ก่อนการวิจัยสิ้นสุด (interim analysis) หรือการวิเคราะห์แยกกลุ่มย่อย (sub-group analysis) จะต้องระบุอยู่ในแผนการวิเคราะห์ด้วย

ตัวอย่างที่ 1 การเขียนการวิเคราะห์ทางสถิติของงานวิจัยเชิงพรรณนา

"การพรรณนาลักษณะตัวอย่างที่เป็นตัวแปรต่อเนื่อง เช่น อายุ รายได้ จะสรุปด้วยค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน มัธยฐาน และค่าสูงสุด ต่ำสุด ส่วนตัวแปรกลุ่ม เช่น เพศ การสูบบุหรี่ จะใช้ความถี่และร้อยละ สรุปอัตราความครอบคลุมของผู้รับบริการด้วยร้อยละและ 95% ช่วงเชื่อมั่น"

Using Statistics in Reseach 1_1 .indd 85 dummy 4 30/6/2557 11:29:32

ตัวอย่างที่ 2 เป็นการเขียนวิธีการวิเคราะห์ในการศึกษาปัจจัยเสี่ยงของการเกิดโรคอุจจาระร่วงที่ใช้แบบวิจัยเชิงวิเคราะห์ภาคตัดขวาง

"การพรรณนาลักษณะตัวอย่างที่เป็นตัวแปรต่อเนื่อง เช่น อายุ รายได้ จะสรุปด้วยค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน มัธยฐาน และค่าสูงสุด ต่ำสุด ส่วนตัวแปรกลุ่ม เช่น เพศ วิธีล้างขวดนม จะใช้ความถี่และร้อยละ การวิเคราะห์ความสัมพันธ์ระหว่างการเกิดโรคอุจจาระร่วงกับปัจจัยเสี่ยงแต่ละตัวใช้สถิติ Odds ratio และ 95% ช่วงเชื่อมั่นของ OR ใช้ multiple logistic regression คำนวณค่า adjusted OR และ 95% ช่วงเชื่อมั่นของ adjusted OR ของปัจจัยเสี่ยงแต่ละตัว"

ตัวอย่างที่ 3 เป็นการเขียนวิธีการวิเคราะห์ทางสถิติในกรณีงานวิจัยเชิงทดลองหรือกึ่งทดลอง ตัวอย่างการศึกษาผลการสอนสุขศึกษา 2 วิธี คะแนนความรู้เป็นตัวแปรผล

"พรรณนาลักษณะข้อมูลพื้นฐานของตัวอย่างแต่ละกลุ่มที่เป็นข้อมูลต่อเนื่อง เช่น อายุ รายได้ คะแนนความรู้ด้วยค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน มัธยฐาน ค่าสูงสุด และข้อมูลพื้นฐานที่เป็นข้อมูลกลุ่ม เช่น เพศ กลุ่มอายุ ระดับการศึกษา ด้วยความถี่และร้อยละ

การเปรียบเทียบความรู้หลังการทดลองระหว่างกลุ่มด้วย t-test และ 95% ช่วงเชื่อมั่นของผลต่างของค่าเฉลี่ยระหว่างกลุ่ม ก่อนทำการวิเคราะห์ด้วยสถิติ t จะตรวจสอบว่าข้อมูลตัวแปรผลของทั้ง 2 กลุ่มมีการแจกแจงแบบปกติ

Using Statistics in Reseach 1_1 .indd 86 dummy 4 27/8/2557 11:51:33

หรือไม่ ถ้าพบว่ามีการแจกแจงไม่ปกติจะทำการแปลงข้อมูลให้ข้อมูลมีการแจกแจงปกติก่อน ในกรณีที่ข้อมูลแปลงยังมีการแจกแจงไม่ปกติจะเปลี่ยนไปใช้สถิติ Wilcoxon signed-rank test ทดสอบความแตกต่างของคะแนนระหว่างกลุ่ม

ในกรณีที่พบว่าคะแนนความรู้ก่อนการทดลองของ 2 กลุ่มแตกต่างกันมากซึ่งอาจกระทบต่อผลการทดลอง จะวิเคราะห์โดยใช้ covariate adjustment เพื่อปรับอิทธิพลของคะแนนความรู้ก่อนการทดลอง"

สรุป

การเขียนวิธีการวิเคราะห์ทางสถิติตามที่ได้ยกตัวอย่างนอกจากจะช่วยให้ผู้พิจารณาโครงร่างงานวิจัยพิจารณาความถูกต้องเหมาะสมของแผนการวิเคราะห์ข้อมูลแล้ว ยังเป็นการป้องกันอคติที่เกิดจากการที่นักวิจัยพยายามจะหาวิธีการวิเคราะห์กับข้อมูลจากตัวอย่างเพื่อให้ได้ข้อสรุปที่ตนเองต้องการ

Using Statistics in Reseach 1_1 .indd 87 dummy 4 27/8/2557 11:51:33

21

ความถูกต้องในการถามเหตุการณ์ย้อนหลัง

การเก็บข้อมูลงานวิจัยที่เกี่ยวข้องกับสุขภาพการเจ็บป่วย ส่วนใหญ่จะเป็นการสอบถามข้อมูลย้อนหลัง ผู้ตอบจะต้องย้อนระลึก (recall) ความถูกต้องของข้อมูลย้อนระลึกจะขึ้นอยู่กับความจำของผู้ตอบ ความสำคัญของเหตุการณ์ และช่วงห่างของเวลาที่ย้อนระลึก ปัจจัยทั้งสามมีผลต่อความถูกต้องของคำตอบแตกต่างกันเป็นกรณีๆไป เหตุการณ์ที่สำคัญ เช่น การเกิดอุบัติเหตุจนต้องเข้ามารักษาตัวในโรงพยาบาลเกิดนานเท่าไรก็จำได้ว่าเคยเกิด แต่ถ้าถามว่าสัปดาห์ที่ผ่านมากินอาหารอะไรบ้าง ผู้ตอบอาจจะจำได้ไม่ถูกต้อง

ความเอนเอียง (bias) ของข้อมูลที่ได้จากการถามเหตุการณ์ย้อนหลังไม่สามารถปรับแก้ด้วยวิธีการทางสถิติ จะต้องแก้ไขที่ต้นเหตุ คือวิธีการถาม

การเพิ่มความถูกต้องของคำตอบจากการถามเหตุการณ์ย้อนหลังทำได้ด้วยการกำหนดช่วงระยะเวลาการย้อนระลึกที่เหมาะสมกับเรื่องที่จะถาม และการถามกระตุ้นความจำด้วยเหตุการณ์ หรือการถามนำ (tracer)

ในการถามเรื่องการเจ็บป่วยย้อนหลังจากการศึกษาของ Kroeger* พบว่าระยะเวลาการย้อนระลึก 2 สัปดาห์ มีจำนวนครั้งรวมของการเจ็บป่วยผิดพลาดน้อยที่สุด เมื่อเทียบกับการถามความเจ็บป่วยในรอบ 1 เดือนหรือรอบ 3 เดือนที่ผ่านมา การเจ็บป่วยที่จำผิดพลาดหรือจำไม่ได้มากที่สุด ได้แก่ ปวดหลัง เวียนหัว เป็นหวัด ปวดแน่นท้อง เป็นไข้ ปวดกล้ามเนื้อ ท้องเสีย* ซึ่งมีทั้งที่ตอบน้อยกว่าและมากกว่าความจริง แต่ถ้ารวมเป็นของกลุ่มแล้วมีจำนวนครั้งและลักษณะอาการหรือโรคใกล้กับการป่วยจริงมากที่สุด ดังนั้นจึงนิยมใช้ระยะเวลา 2 สัปดาห์เป็นช่วงเวลาสำหรับถามการเจ็บป่วยย้อนหลัง

การกำหนดระยะเวลาว่าจะสอบถามย้อนหลังนานเท่าไรจึงจะเหมาะสม จะขึ้นอยู่กับความสำคัญของเหตุการณ์ เช่น การถามเรื่องอาหารที่กินเพื่อประเมินการได้สารอาหารพอเพียง พบว่าการถาม 24 ชั่วโมงย้อนหลังให้คำตอบที่ถูกต้องมาก เมื่อเทียบกับย้อนหลัง 3 วัน หรือ 5 วัน** ดังนั้นการถามเรื่องอาหารที่กินเพื่อประเมินการได้สารอาหารพอเพียงจึงนิยมถาม 24 ชั่วโมงย้อนหลัง

จะเห็นได้ว่าในแต่ละเหตุการณ์ใช้ระยะเวลาไม่เท่ากัน โดยเฉพาะเหตุการณ์ที่มีความสำคัญน้อยหรือเกิดขึ้นเป็นประจำ ช่วงเวลาการถามย้อนหลังควรสั้น ถ้านักวิจัยต้องใช้ตัวแปรที่ได้จากการถามย้อนหลังตอบคำถามหลัก ควรมีการทบทวนวรรณกรรมว่ามีวิธีการกำหนดระยะเวลาการถามย้อนหลังที่เหมาะสมเป็นเท่าไร

---

* Axel Kroeger, "Health Interview Surveys in Developing Countries: A Review of the Methods and Results." **International Journal of Epidemiology.** 1983; 12(4): 465–481.

** Miriam Chavez and Ruth Huenemann, "Measuring Impact by Assessing Dietary Intake and Food Consumption." in D.E. Sahn et al, **Method for the Evaluation of the Impact of Food and Nutrition Programs.** Tokyo: The United Nations University, 1984.

Using Statistics in Reseach 1_1 .indd 89 dummy 4 27/8/2557 11:51:33

การเพิ่มความถูกต้องของการถามเหตุการณ์ย้อนหลังอีกวิธีหนึ่ง คือการกระตุ้นความจำโดยใช้เหตุการณ์อื่นที่ทราบระยะเวลาเป็นจุดให้ย้อนระลึก เช่น ถามว่าท่านเริ่มสูบบุหรี่ตั้งแต่อายุเท่าไรอาจจะประมาณระยะเวลาได้ไม่ตรง การถามโดยใช้เหตุการณ์อื่นที่ทราบระยะเวลาเป็นตัวกระตุ้นความจำ เช่น ถามว่าตอนที่เริ่มสูบบุหรี่เรียนจบชั้นประถมศึกษาแล้วหรือยัง ก่อนหรือหลังบวช ก่อนหรือหลังเกณฑ์ทหาร ก่อนหรือหลังแต่งงาน ฯลฯ แล้วจึงถามต่อ หาว่าห่างจากเหตุการณ์นั้นกี่ปี จะช่วยให้ย้อนระลึกได้ถูกต้องมากขึ้น

ในกรณีการถามเรื่องการเจ็บป่วยย้อนหลัง 2 สัปดาห์ ถ้าต้องการความถูกต้องเพิ่มขึ้นจะใช้การถามนำเป็นตัวกระตุ้น เช่น ในช่วง 2 สัปดาห์ที่ผ่านมาท่านมีอาการปวดหัว เวียนหัวหรือไม่ ตาแดง ตาเจ็บหรือไม่ ถามไล่ไปตั้งแต่ศีรษะจรดเท้า จะช่วยให้จำนวนการเจ็บป่วยเพิ่มขึ้นใกล้เคียงกับจำนวนการเจ็บป่วยที่เป็นจริง*

สรุป

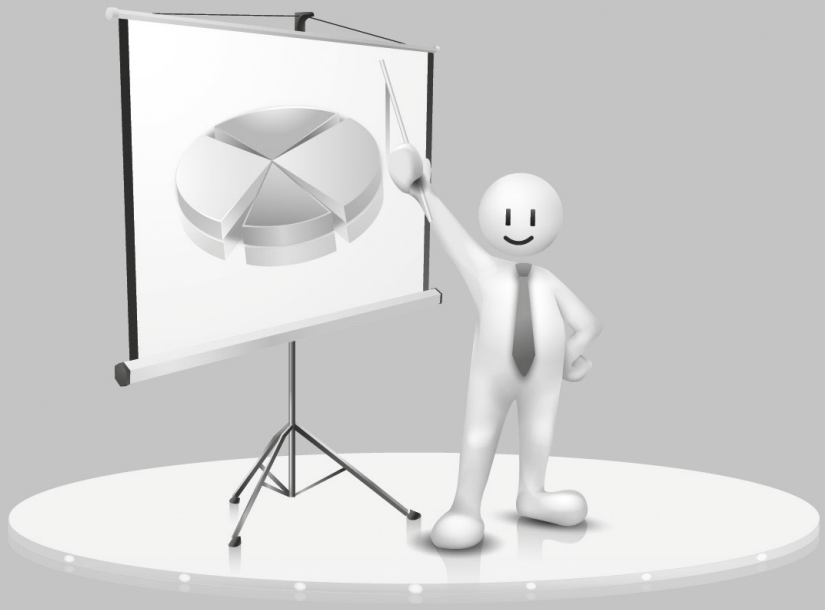
การถามย้อนหลังระยะเวลาการเกิดเหตุการณ์ในอดีตควรใช้วิธีการถามกระตุ้นความจำด้วยเหตุการณ์อื่นที่ทราบระยะเวลาเพื่อช่วยในการย้อนระลึก ถ้าเป็นพฤติกรรมในชีวิตประจำวัน หรือการเจ็บป่วยควรกำหนดระยะเวลาการย้อนระลึกให้เหมาะสมกับสิ่งที่จะถาม หรือกระตุ้นความจำด้วยการถามนำ ระยะเวลาที่เหมาะสมของการถามย้อนหลังในแต่ละเรื่องสามารถศึกษาได้จากงานวิจัยที่เกี่ยวข้องกับความถูกต้องของการสอบถามย้อนหลังในเรื่องนั้นๆ

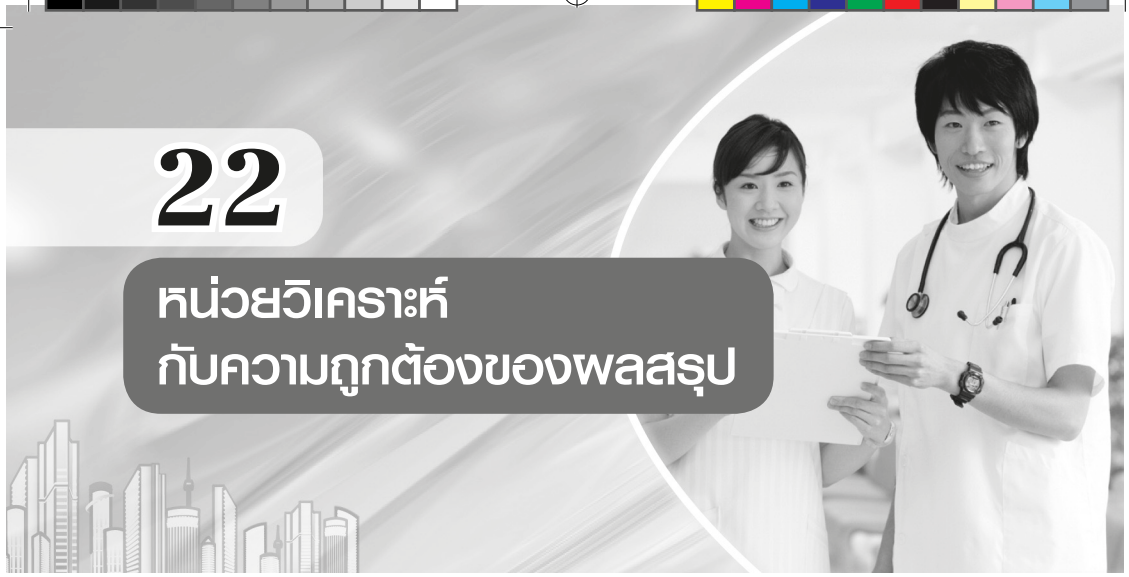
---

* David A. Ross. **Health Interview Surveys in Developing Countries.** London: Evaluation and Planning Centre for Health Care, London School of Hygiene and Tropical Medicine, 1984.

Using Statistics in Reseach 1_1 .indd 90 dummy 4 27/8/2557 11:51:34

สถิติในการวิเคราะห์ข้อมูลวิจัย





22

หน่วยวิเคราะห์ กับความถูกต้องของผลสรุป

การทำวิจัยที่มีการสุ่มตัวอย่างจะต้องกำหนดว่าจะใช้อะไรเป็นหน่วยสุ่ม (sampling unit) ในการศึกษาทางด้านการแพทย์และสาธารณสุข หน่วยสุ่มที่พบส่วนใหญ่จะเป็นบุคคลหรือผู้ป่วย แต่ในบางครั้งหน่วยสุ่มอาจเป็นหมู่บ้าน โรงงาน โรงพยาบาล หรือหอผู้ป่วย สำหรับหน่วยวิเคราะห์ซึ่งหมายถึงหน่วยที่นำข้อมูลมาคำนวณค่าสถิติต่างๆ ส่วนใหญ่จะเป็นบุคคลหรือผู้ป่วย แต่ในบางการศึกษาอาจหมายถึงส่วนต่างๆของร่างกาย เช่น ตา หู แขน ฟัน เป็นต้น วิธีการวิเคราะห์ที่ไม่คำนึงถึงหน่วยสุ่มและหน่วยวิเคราะห์จะทำให้ได้ผลสรุปที่ไม่ถูกต้อง

1) **หน่วยสุ่มและหน่วยวิเคราะห์เป็นหน่วยเดียวกัน** ความผิดพลาดที่พบแบ่งได้ 2 กรณี คือ

   1.1) **ความผิดพลาดที่เกิดจากการกำหนดหน่วยวิเคราะห์ผิด** ทำให้มีการนับจำนวนหน่วยวิเคราะห์ซ้ำ เช่น การศึกษาเปรียบเทียบระดับการได้ยินระหว่างคนงานในโรงงานกับคนงานในไร่ โดยสุ่มตัวอย่างมาศึกษากลุ่มละ 40 คน วัดระดับการได้ยินแยกเป็นหูซ้ายและขวา ในการวิเคราะห์ถ้านำข้อมูลระดับการได้ยินของหูทั้ง 2 ข้าง

มารวมกัน จะทำให้หน่วยวิเคราะห์ของแต่ละกลุ่มเพิ่มขึ้นจาก 40 เป็น 80 โดยแต่ละคนจะมีค่าระดับการได้ยิน 2 ค่า การที่มีจำนวนตัวอย่างในการวิเคราะห์เพิ่มขึ้นอีก 1 เท่าจะส่งผลทำให้ค่า P value มีค่าน้อยกว่าความเป็นจริง และข้อมูลการได้ยินของหูแต่ละข้างของบุคคลเดียวจะไม่เป็นอิสระต่อกัน แต่เมื่อนำมาวิเคราะห์รวมกันเสมือนข้อมูลการได้ยินของหูแต่ละข้างเป็นอิสระต่อกัน นำไปสู่การสรุปผลที่ผิด

การวิเคราะห์ที่ถูกต้อง ถ้าคิดว่าหู 2 ข้างได้รับผลกระทบต่อสิ่งที่ทำให้สูญเสียการได้ยินเท่ากันจะใช้ค่าเฉลี่ยระดับการได้ยินของหูซ้ายและขวามาเป็นตัวแทนการได้ยินของแต่ละคน ทำให้มีหน่วยวิเคราะห์เท่ากับจำนวนคนตัวอย่าง หรือถ้าคิดว่าหู 2 ข้างได้รับผลกระทบต่อสิ่งที่ทำให้สูญเสียการได้ยินไม่เท่ากันจะใช้ค่าระดับการได้ยินของหูข้างที่มีค่าสูง (หูข้างที่ได้ยินไม่ดี) เป็นค่าที่ใช้ในการวิเคราะห์

1.2) **ความผิดพลาดที่เกิดจากการที่ไม่คำนึงถึงผลที่มีการเปลี่ยนแปลงตามช่วงเวลา** ในการศึกษาที่มีการวัดผลซ้ำหลายครั้งในหน่วยศึกษาเดียวกัน เช่น การเปรียบเทียบการฝึกอบรม 2 วิธี มีการวัดความรู้ก่อนการฝึกอบรม สิ้นสุดการฝึกอบรม หลังการฝึกอบรม 3 เดือนและ 6 เดือน โดยจะดูว่าวิธีการฝึกอบรมแบบใดให้ผลดีกว่ากัน การวิเคราะห์ที่ไม่เหมาะสมคือการเปรียบเทียบความรู้แต่ละช่วงทีละคู่

Using Statistics in Reseach 1_1 .indd 93 dummy 4 27/8/2557 11:51:34

การออกแบบงานวิจัยที่มีการวัดผลหลายครั้ง ส่วนใหญ่จะมีจุดมุ่งหมายที่จะดูผลในภาพรวม และการเปลี่ยนแปลงตามช่วงเวลา ในการวิเคราะห์ให้ได้ภาพรวมที่ถูกต้องจะต้องกำหนดหน่วยวิเคราะห์เป็นบุคคล และต้องเลือกวิธีการวิเคราะห์ที่คำนึงถึงผลจากการวัดซ้ำในระยะเวลาต่าง ๆด้วย ในกรณีนี้จะต้องใช้การวิเคราะห์แบบ repeated measure ANOVA ซึ่งสามารถเปรียบเทียบภาพรวมของผลการฝึกอบรม และเปรียบเทียบผลจากการเปลี่ยนแปลงความรู้ตามระยะเวลาของการฝึกอบรมอย่างถูกต้อง

2) **หน่วยสุ่มและหน่วยวิเคราะห์ไม่ได้เป็นหน่วยเดียวกัน** ความผิดพลาดที่พบมีดังนี้

2.1) **ความผิดพลาดที่เกิดจากการเลือกใช้สถิติวิเคราะห์ไม่ถูกต้อง** เช่น การศึกษาเรื่องสารเคลือบร่องฟัน 2 ชนิด ว่าชนิดใดจะติดทนนานกว่ากัน สุ่มเด็กเข้ามาในการศึกษา 30 คน แต่ละคนเคลือบร่องฟันด้วยสารเคลือบร่องฟัน 2 ชนิดบนฟันกรามด้านซ้ายและขวาข้างละ 1 ชนิด ในการวิเคราะห์จะกำหนดให้หน่วยวิเคราะห์เป็นฟันกรามแต่ละซีก (ไม่ใช่คน) แต่เนื่องจากสารเคลือบทั้ง 2 ชนิดเคลือบบนฟันคนละซีกในคนคนเดียวกัน การเปรียบเทียบจะต้องใช้สถิติในกรณีที่ 2 กลุ่มไม่เป็นอิสระต่อกัน

2.2) **ความผิดพลาดที่เกิดจากการไม่รวมความแปรปรวนระหว่างหน่วยสุ่มในการเปรียบเทียบ** ในกรณีที่หน่วยสุ่มเป็นกลุ่ม (หมู่บ้าน) แต่หน่วยวิเคราะห์เป็นคน เช่น

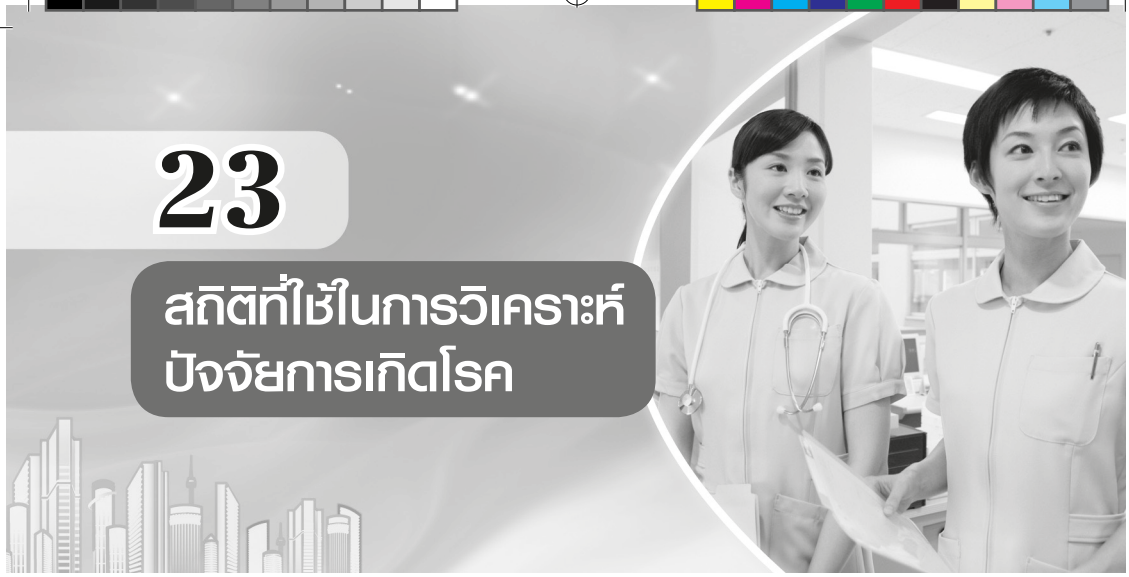
Using Statistics in Reseach 1_1 .indd 94 dummy 4 30/6/2557 11:29:33

การเปรียบเทียบอัตราการเป็นโรคอ้วนของเด็กที่มีอายุต่ำกว่า 5 ปีในเขตเมืองกับเขตชนบท โดยกำหนดให้ชุมชน (หรือหมู่บ้าน) เป็นหน่วยสุ่ม การสุ่มแบบนี้เรียกว่าการสุ่มแบบกลุ่ม (cluster sampling) ความผิดพลาดที่พบเสมอในการวิเคราะห์ คือ การนำเอาผลการศึกษาจากแต่ละคนในชุมชน/หมู่บ้านตัวอย่างมารวมกันวิเคราะห์โดยตรงโดยไม่มีการปรับอิทธิพลของการสุ่มแบบกลุ่ม การวิเคราะห์เพื่อสรุปผลการเปรียบเทียบที่ถูกต้องจะต้องใช้ความแปรปรวนที่ได้มาจากความต่างระหว่างบุคคล และความแตกต่างระหว่างหน่วยสุ่มมาใช้คำนวณค่าสถิติที่ใช้ในการประมาณค่าหรือทดสอบสมมติฐาน

สรุป

ในการศึกษาควรพิจารณากำหนดหน่วยสุ่มและหน่วยวิเคราะห์ให้เหมาะสมกับคำถามงานวิจัย หน่วยสุ่มและหน่วยวิเคราะห์ไม่เป็นหน่วยเดียวกันก็ได้ แต่จะต้องเลือกใช้สถิติในการวิเคราะห์ให้เหมาะสม ความผิดพลาดในการกำหนดหน่วยวิเคราะห์และวิธีการวิเคราะห์ที่ไม่เหมาะสมจะมีผลกระทบต่อความถูกต้องของผลสรุปของงานวิจัย

Using Statistics in Reseach 1_1 .indd 95 dummy 4 30/6/2557 11:29:33



23 สถิติที่ใช้ในการวิเคราะห์ปัจจัยการเกิดโรค

การศึกษาปัจจัยที่มีความสัมพันธ์กับการเกิดโรคเป็นขั้นตอนเบื้องต้นของการหาสาเหตุของการเกิดโรค ในการวิเคราะห์ถ้าพบว่าเมื่อมีการเปลี่ยนค่าข้อมูลของตัวแปรอิสระ (ปัจจัย) แล้วมีผลทำให้ค่าข้อมูลของตัวแปรตาม (การเกิดโรค) เปลี่ยนแปลง แสดงว่าตัวแปรทั้งสองมีความสัมพันธ์กัน

ตาราง 1 การเกิดโรคเส้นโลหิตในสมองแตกกับโรคความดันโลหิตสูง

	D+	D–
ความดันโลหิตสูง	60 (5.9%)	950 (94.1%)
ความดันโลหิตปกติ	6 (0.4%)	1,500 (99.6%)

จากข้อมูลตาราง 1 แสดงการเกิดโรคเส้นโลหิตในสมองแตกกับโรคความดันโลหิตสูง ถ้าเป็นโรคความดันโลหิตสูงแล้วทำให้มีการเกิดโรคเส้นโลหิตในสมองแตก (5.9%) มากกว่าคนที่มีความดันโลหิตปกติ (0.4%) แสดงว่าการเป็นโรคความดันโลหิตสูงเป็นปัจจัยทำให้เกิดโรค

Using Statistics in Reseach 1_1 .indd 96 dummy 4 30/6/2557 11:29:33

เส้นโลหิตในสมองแตกเพิ่มขึ้น จะเรียกปัจจัยแบบนี้ว่าปัจจัยเสี่ยง (risk factor)

ตาราง 2 การเกิดโรคหัวใจและหลอดเลือดกับการออกกำลังกาย

	D+	D–
ออกกำลังกายสม่ำเสมอ	78 (13.5%)	500 (86.5%)
ไม่ออกกำลังกายสม่ำเสมอ	112 (20.7%)	430 (79.3%)

จากข้อมูลตาราง 2 แสดงการเกิดโรคหัวใจและหลอดเลือด พบว่าถ้าออกกำลังกายสม่ำเสมอทำให้เกิดโรคหัวใจและหลอดเลือด (13.5%) น้อยกว่าคนที่ไม่ได้ออกกำลังกายสม่ำเสมอ (20.7%) แสดงว่าปัจจัยการออกกำลังกายทำให้เกิดโรคหัวใจและหลอดเลือดลดลง จะเรียกปัจจัยแบบนี้ว่าปัจจัยป้องกัน (protective factor)

ตาราง 3 การขับขี่รถผิดกฎจราจรจำแนกตามระดับการศึกษา

	D+	D–
ปริญญาตรีและสูงกว่า	78 (19.6%)	320 (80.4%)
ต่ำกว่าระดับปริญญาตรี	108 (20.1%)	430 (79.9%)

จากข้อมูลตาราง 3 แสดงการขับขี่รถผิดกฎจราจร โดยพบว่าไม่ว่าเป็นผู้มีการศึกษาระดับใดต่างก็มีอัตราการขับขี่รถผิดกฎจราจรพอๆกัน แสดงว่าระดับการศึกษาไม่มีอิทธิพลหรือไม่มีความสัมพันธ์ต่อการขับขี่รถผิดกฎจราจร

Using Statistics in Reseach 1_1 .indd 97 dummy 4 30/6/2557 11:29:33

สำหรับการศึกษาเพื่อหาปัจจัยเสี่ยงจะต้องออกแบบงานวิจัยเป็นแบบการศึกษาเชิงวิเคราะห์ (analytical study) ซึ่งแบ่งย่อยได้ 3 แบบ คือ การศึกษาแบบ cross-sectional การศึกษาแบบ case-control และการศึกษาแบบ cohort แบบงานวิจัยทั้ง 3 แบบใช้กับในสถานการณ์ที่แตกต่างกันตามสภาพปัญหาและข้อจำกัดของการศึกษา รายละเอียดหาอ่านเพิ่มเติมจากหนังสือวิธีวิทยาการวิจัยทางด้านการแพทย์และสาธารณสุข หรือหลักการศึกษาด้านระบาดวิทยา

ประเภทตัวแปรของปัจจัยและการเกิดโรค การเกิดโรคเป็นตัวแปรกลุ่มแบบมี 2 ค่า (dichotomous variable) เช่น เป็นโรค/ไม่เป็นโรค ปัจจัยส่วนใหญ่จะเป็นตัวแปรกลุ่มแบบมี 2 ค่าด้วย เช่น สวมหมวกหรือไม่สวมหมวกนิรภัย สัมผัสหรือไม่ได้สัมผัสอาหารติดเชื้อ หรือปริมาณโลหะหนักเกินหรือไม่เกินระดับมาตรฐาน

สถิติที่ใช้แสดงความสัมพันธ์ระหว่างตัวแปรการเกิดโรคและตัวแปรปัจจัยที่เป็นตัวแปรกลุ่มแบบมี 2 ค่า คือ

1) **สถิติไคสแควร์ (Chi-square)** ใช้แสดงว่าปัจจัยมีความสัมพันธ์กับการเกิดโรคหรือไม่ โดยไม่สามารถระบุขนาดว่ามีความสัมพันธ์มากน้อยเท่าไร และไม่สามารถระบุทิศทางความสัมพันธ์ว่าเป็นปัจจัยเสี่ยงหรือปัจจัยป้องกัน
2) **สถิติ Relative Risk (RR) หรือ Odds Ratio (OR)** ใช้แสดงความสัมพันธ์ระหว่างปัจจัยกับการเกิดโรค นอกจากจะระบุว่ามีความสัมพันธ์หรือไม่แล้ว ยังสามารถบอกขนาดและทิศทางความสัมพันธ์ได้ด้วย การจะใช้สถิติ RR หรือ OR ขึ้นอยู่กับว่าแบบงานวิจัยนั้นสามารถคำนวณอัตราการเกิดโรคที่แท้จริงได้หรือไม่ การศึกษาเป็นแบบ cohort คำนวณอัตราการเกิดโรคที่แท้จริงได้จะใช้สถิติ RR และการศึกษาแบบ cross-sectional หรือ case-control จะใช้สถิติ OR

Using Statistics in Reseach 1_1 .indd 98 dummy 4 30/6/2557 11:29:33

การพิจารณาสรุปปัจจัยเสี่ยงจากค่า RR หรือ OR พิจารณาจากค่าช่วงเชื่อมั่นของ RR หรือ OR ถ้าช่วงเชื่อมั่นของ RR หรือ OR คร่อม 1 เช่น 95% ช่วงเชื่อมั่น OR (0.87–1.21) แสดงว่าตัวแปรดังกล่าวไม่มีความสัมพันธ์กับการเกิดโรค ในกรณีช่วงเชื่อมั่นของ RR หรือ OR มีค่าน้อยกว่า 1 เช่น 95% ช่วงเชื่อมั่น OR (.051–.072) แสดงว่าเป็นปัจจัยป้องกัน ถ้าช่วงเชื่อมั่นของ RR หรือ OR มีค่ามากกว่า 1 แสดงว่าเป็นปัจจัยเสี่ยง เช่น ค่า 95% ช่วงเชื่อมั่นของ OR (3.7–5.8)

การแปลความหมายค่า 95% ช่วงเชื่อมั่นของ RR ถ้าค่า 95% ช่วงเชื่อมั่นของ RR อยู่ระหว่าง 3.7–5.8 จะสรุปว่าผู้ที่ได้รับปัจจัยดังกล่าวมีโอกาสเกิดโรคมากกว่าผู้ที่ไม่ได้รับปัจจัยอย่างน้อยที่สุด 3.7 เท่า หรืออาจมากถึง 5.8 เท่า ขนาดของค่า RR และ OR ยิ่งมากยิ่งแสดงว่าปัจจัยมีความสัมพันธ์กับการเกิดโรคมาก

การคำนวณค่า OR นอกจากคำนวณจากสูตร ad/bc จากข้อมูลในตาราง 2 × 2 แล้ว ยังสามารถคำนวณค่า OR จากตัวแบบ logistic regression ค่า OR ที่ได้จากการคำนวณทั้ง 2 วิธีมีค่าเท่ากัน

ในกรณีที่ตัวแปรปัจจัยเป็นตัวแปรแบบ polytomous ภายในตัวแปรมีมากกว่า 2 กลุ่ม เช่น ความหนักเบาของโรคแบ่งเป็น น้อย ปานกลาง มาก หรือกรณีที่เป็นตัวแปรต่อเนื่อง จะต้องคำนวณค่า OR จากตัวแบบ logistic regression ค่า OR ที่ได้จะเป็นความเสี่ยงที่เพิ่มขึ้นทุก 1 หน่วยวัดของตัวแปรปัจจัย เช่น ความสัมพันธ์ของการเกิดโรคหัวใจกับอายุ ถ้าใช้อายุเป็นปีมาคำนวณ ค่า OR เท่ากับ 1.2 จะหมายความว่าอายุที่เพิ่มขึ้น 1 ปีจะมีโอกาสเสี่ยงเพิ่มขึ้น 1.2 เท่า ในทางปฏิบัติจะไม่นิยมนำตัวแปรปัจจัยที่เป็นตัวแปรต่อเนื่องมาคำนวณค่า OR โดยตรง เพราะความเสี่ยงไม่ได้เพิ่มขึ้นทุกๆหน่วยวัดที่เพิ่มขึ้น แต่ความเสี่ยงจะเกิดในบางช่วงของหน่วยวัด เช่น ความเสี่ยงของการเกิดโรคหัวใจจะแปรตามอายุเป็นช่วงๆ อายุน้อยกว่า 45 ปี, 46–60 ปี, 60 ปีขึ้นไป

Using Statistics in Reseach 1_1 .indd 99 dummy 4 27/8/2557 11:51:34

เป็นต้น จึงต้องแปลงข้อมูลให้เป็นกลุ่มตามระดับความเสี่ยงก่อนจึงนำมาคำนวณค่า OR

ข้อควรระวังในการแปลความหมายของความสัมพันธ์ คือ ความบังเอิญที่เกิดจากค่าสถิติที่คำนวณได้ระบุว่ามีความสัมพันธ์ แต่ไม่สามารถหาเหตุผลทางวิทยาศาสตร์มาอธิบายได้ว่ามีความสัมพันธ์กันอย่างไร เช่น การเป็นโรคความดันโลหิตสูงกับการชอบใส่เสื้อสีแดง เพื่อจะลดความผิดพลาดดังกล่าว ตัวแปรที่นำมาศึกษาปัจจัยเสี่ยงควรเป็นตัวแปรที่มีเหตุผลทางทฤษฎีสนับสนุนว่าน่าจะมีความสัมพันธ์กับการเกิดโรค

สรุป

การใช้สถิติสำหรับการวิเคราะห์หาปัจจัยเสี่ยงของการเกิดโรคที่สามารถให้ข้อมูลครบถ้วนทั้งขนาดและทิศทางของความสัมพันธ์ คือ ค่า RR และ OR จึงนิยมใช้ค่าสถิตินี้ในการวิเคราะห์ข้อมูลงานวิจัยเชิงวิเคราะห์ การแปลผลปัจจัยเสี่ยงให้แปลผลจากค่า 95% ช่วงเชื่อมั่นของ RR หรือ OR

ตัวแปรที่นำมาหาปัจจัยเสี่ยงควรมีเหตุผลทางวิทยาศาสตร์อธิบายได้ว่ามีความสัมพันธ์กับการเกิดโรคอย่างไร

Using Statistics in Reseach 1_1 .indd 100 dummy 4 27/8/2557 11:51:34

24

การใช้ adjusted OR วิเคราะห์ค่าปัจจัยเสี่ยง

รายงานวิจัยเชิงสำรวจเรื่องหาปัจจัยเสี่ยงที่ตีพิมพ์ในวารสารภายในประเทศส่วนใหญ่ใช้สถิติ OR ในการวิเคราะห์ค่าปัจจัยเสี่ยง มีบางรายงานวิจัยใช้สถิติ adjusted OR ในการวิเคราะห์ปัจจัยเสี่ยง มีคำถามว่าควรพิจารณาขนาดความเสี่ยงจากค่า OR หรือ adjusted OR

ค่า OR คำนวณจากข้อมูลของแต่ละปัจจัย จะบอกขนาดความเสี่ยงแบบหยาบ (curd) เพราะค่าที่ได้จะมีอิทธิพลของตัวแปรอื่นที่มีต่อตัวแปรตาม (เป็นโรค) ร่วมอยู่ด้วย ทำให้ค่า OR ไม่สามารถบอกอิทธิพลที่แท้จริงของตัวแปรปัจจัยนั้นได้

ตัวอย่าง ในการศึกษาความสัมพันธ์ของความรู้เรื่องความปลอดภัยในการขับขี่รถจักรยานยนต์กับการสวมใส่หมวกนิรภัยในขณะขับขี่ ได้ผลดังตาราง 1

ตาราง 1 แสดงความสัมพันธ์ระหว่างความรู้เรื่องความปลอดภัยฯกับการสวมใส่หมวกนิรภัยฯ

ความรู้เรื่องความปลอดภัยฯ	การสวมใส่หมวกนิรภัยฯ		OR (95% CI)
	สวม n (%)	ไม่สวม n (%)	
รู้	62 (28.7)	154 (71.3)	2.28 (1.44, 2.62)
ไม่รู้	36 (15.0)	204 (85.0)	

จากตาราง 1 ผู้ที่มีความรู้เรื่องความปลอดภัยในการขับขี่รถจักรยานยนต์มีการสวมใส่หมวกนิรภัยในขณะขับขี่เป็น 2.28 เท่า (OR = 2.28) ของผู้ที่ไม่มีความรู้ แต่ค่า OR ที่ได้ยังอาจมีอิทธิพลของเพศผู้ขับขี่ ในการวิเคราะห์เพื่อแยกอิทธิพลของเพศที่มีต่อความรู้เรื่องความปลอดภัยฯและการสวมใส่หมวกนิรภัยฯทำได้โดยการสร้างตารางแสดงความสัมพันธ์ระหว่างความรู้เรื่องความปลอดภัยฯกับการสวมใส่หมวกนิรภัยฯ แยกตามเพศ และคำนวณ adjusted OR ด้วยวิธีของ Mantel-Haenszel

Using Statistics in Reseach 1_1 .indd 102 dummy 4 30/6/2557 11:29:33

ตาราง 2 แสดงความสัมพันธ์ระหว่างความรู้เรื่องความปลอดภัยฯกับการสวมใส่หมวกนิรภัยฯ แยกตามเพศ

เพศ	ความรู้เรื่องความปลอดภัยฯ	การสวมใส่หมวกนิรภัยฯ		OR (95% CI)	OR_{MH}
		สวม n (%)	ไม่สวม n (%)		
ชาย	รู้	24 (25.5)	70 (74.5)	1.40 (0.74, 2.66)	2.31 (1.45, 3.67)
	ไม่รู้	24 (19.7)	49 (80.3)		
หญิง	รู้	38 (32.8)	84 (68.8)	3.994 (1.97, 8.12)	
	ไม่รู้	12 (10.2)	106 (89.8)		

ตาราง 2 เมื่อพิจารณาจากค่า 95% ช่วงเชื่อมั่นของ OR พบว่าช่วงเชื่อมั่นของเพศชายมีค่าคร่อม 1 แสดงว่าความรู้และการสวมใส่หมวกนิรภัยไม่สัมพันธ์กัน ผู้ชายมีความรู้ฯ หรือไม่มีความรู้ฯ มีการสวมใส่หมวกนิรภัยฯไม่ต่างกัน ส่วนค่า OR ของเพศหญิงมีค่าไม่คร่อม 1 แสดงว่าผู้หญิงที่รู้เรื่องความปลอดภัยฯจะมีการสวมใส่หมวกนิรภัยฯมากกว่าผู้ที่ไม่รู้ 4 เท่า (OR = 3.99) และค่า adjusted OR ของความรู้เมื่อปรับอิทธิพลของเพศแล้วมีค่า 2.31 ซึ่งมีค่ามากกว่า curd OR (2.28) หมายความว่าถ้าตัดอิทธิพลของเพศออกไปแล้ว พบว่าความรู้เรื่องความปลอดภัยในการขับขี่รถจักรยานยนต์จะทำให้มีการสวมใส่หมวกนิรภัยในขณะขับขี่เพิ่มอีก 2.3 เท่า ดังนั้นถ้านักวิจัยต้องการนำเสนอค่าความเสี่ยงจากความรู้ต้องนำเสนอด้วยค่า adjusted OR

Using Statistics in Reseach 1_1 .indd 103 dummy 4 8/9/2557 11:47:47

การคำนวณค่า OR ยังสามารถคำนวณได้จากตัวแบบถดถอยลอจิสติกพหุคูณ (multiple logistic regression) ซึ่งเป็นตัวแบบแสดงความสัมพันธ์เชิงเส้นตรงของค่า odds ของการเกิดโรคกับตัวแปรอิสระ (ปัจจัยเสี่ยง) ทั้งหลาย

$$\ln(\text{Odds}) = a + b_1x_1 + b_2x_2 + \ldots + b_kx_k$$

ถ้าตัวแบบมีตัวแปรอิสระเพียง 1 ตัว คือความรู้เรื่องความปลอดภัยฯ ค่า OR ที่คำนวณได้จากค่าสัมประสิทธิ์ของตัวแปรปัจจัยในตัวแบบถดถอยลอจิสติกจะมีค่าเท่ากับค่า OR หรือ curd OR ที่คำนวณโดยตรงจากตาราง ในทำนองเดียวกันถ้ามีตัวแปรอิสระ 2 ตัว คือ ความรู้เรื่องความปลอดภัยฯและเพศอยู่ในสมการ ค่า OR ของตัวแปรความรู้เรื่องความปลอดภัยฯที่คำนวณได้จากสมการเป็นค่า adjusted OR ซึ่งมีค่าเท่ากับค่า adjusted OR ที่คำนวณด้วยวิธีของ Mantel-Haenszel

เนื่องจากปัจจุบันมีการใช้คอมพิวเตอร์อย่างแพร่หลาย มีโปรแกรมใช้วิเคราะห์ตัวแบบถดถอยลอจิสติกพหุคูณ และสามารถคำนวณค่า OR ในกรณีที่มีปัจจัยหลายตัวได้ง่าย ดังนั้นปัจจุบันจึงนิยมใช้ตัวแบบถดถอยลอจิสติกพหุคูณคำนวณค่า adjusted OR

ค่า adjusted OR ของแต่ละตัวแปรจะปรับอิทธิพลจากตัวแปรที่นำมาวิเคราะห์เท่านั้น ถ้านักวิจัยไม่ทบทวนวรรณกรรมให้ครอบคลุมตัวแปรที่สำคัญไม่ได้เก็บข้อมูลมาร่วมวิเคราะห์ด้วย ก็จะทำให้ค่า adjusted OR ที่ได้ไม่ถูกต้องตามความเป็นจริง ดังนั้นในการศึกษานักวิจัยต้องแสดงให้เห็นว่าตัวแปรที่ศึกษาครอบคลุมตัวแปรที่สำคัญทั้งหมด และมีขนาดตัวอย่างที่เพียงพอที่จะใช้ในการวิเคราะห์ความสัมพันธ์ของปัจจัยทั้งหมดได้ถูกต้อง

Using Statistics in Reseach 1_1 .indd 104 dummy 4 27/8/2557 11:51:34

สรุป

ค่า adjusted OR ใช้แสดงขนาดความสัมพันธ์ระหว่างการเกิดโรคกับปัจจัยนั้นโดยได้ปรับอิทธิพลจากปัจจัยอื่นที่นำมาวิเคราะห์ทั้งหมด ดังนั้นการรายงานขนาดของความสัมพันธ์ที่แท้จริงต้องรายงานด้วย adjusted OR ส่วน OR หรือ curd OR ใช้วิเคราะห์ดูความสัมพันธ์เบื้องต้นเท่านั้น

Using Statistics in Reseach 1_1 .indd 105 dummy 4 30/6/2557 11:29:34

25

ผลที่เกิดจากการแบ่งกลุ่มตัวแปรต่อเนื่อง

ในการวิเคราะห์ข้อมูลที่เป็นตัวแปรต่อเนื่อง เช่น อายุ ความดันโลหิต ฯลฯ พบว่ามีการวิเคราะห์ใน 2 รูปแบบ คือ นำค่าข้อมูลต่อเนื่องมาวิเคราะห์โดยตรง เช่น คำนวณหาค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน คำนวณค่า 95% ช่วงเชื่อมั่นของค่าเฉลี่ย ฯลฯ การวิเคราะห์อีกแบบหนึ่งทำโดยการนำข้อมูลต่อเนื่องมาแบ่งให้เป็นกลุ่มก่อนแล้วจึงนำไปวิเคราะห์ เช่น อายุแบ่งเป็นกลุ่มอายุ แล้วนำไปหาความถี่ หรือนำไปหาความสัมพันธ์กับการเกิดโรคจำแนกตามกลุ่มอายุ

ทำไมจึงมีการแบ่งกลุ่มให้กับข้อมูลต่อเนื่อง โดยปกติเหตุผลของการจัดกลุ่มให้กับข้อมูลต่อเนื่องแบ่งได้เป็น 2 ประการ คือ

1) **เพื่อช่วยเพิ่มความเข้าใจในการนำเสนอข้อมูล** เช่น ถ้าต้องการนำเสนอเพื่อเน้นว่าปัญหาของเด็กที่เป็นไข้เลือดออกมีมากในบางกลุ่มอายุ ถ้านำเสนอโดยค่าเฉลี่ย เช่น เด็กที่ป่วยด้วยโรคไข้เลือดออกมีอายุเฉลี่ย 5.6 ปี และส่วนเบี่ยงเบนมาตรฐาน 4.9 ปี ผู้อ่านจะไม่ทราบว่าอายุใดเป็นมาก แต่ถ้านำอายุมาจัดกลุ่มแล้วนำเสนอความถี่ของการป่วยของกลุ่มอายุที่เป็นมาก

เช่น เด็กอายุต่ำกว่า 2 ปี ป่วยด้วยโรคไข้เลือดออก 54% จะช่วยทำให้เข้าใจมากกว่า นอกจากนี้การแบ่งกลุ่มยังช่วยในการนำเสนอผลให้สอดคล้องกับระบบการจัดบริการ เช่น แบ่งอายุไปตามกลุ่มเป้าหมายที่มารับบริการ ต่ำกว่า 1 ปี ก่อนวัยเรียน วัยเรียน วัยรุ่น เป็นต้น

2) **เพื่อใช้หาปัจจัยเสี่ยงหรือความสัมพันธ์** ในกรณีที่มีเหตุผลทางวิทยาศาสตร์อธิบายว่าความเสี่ยงเพิ่มขึ้นเป็นช่วง ๆของตัวแปรปัจจัย เช่น ความเสี่ยงของการเป็นโรคหัวใจไม่ได้เพิ่มขึ้นตามอายุที่เพิ่มขึ้นในแต่ละปี แต่จะมีความเสี่ยงเพิ่มขึ้นเป็นช่วงของอายุ เช่น < 40 ปี, 41-60 ปี และ > 60 ปี การแบ่งกลุ่มอายุตามระดับความเสี่ยงยังช่วยในการแปลผลค่าสถิติ Odds Ratio ได้สอดคล้องกับลักษณะทางคลินิก ทำให้นำผลไปใช้งานได้ถูกต้อง

การกำหนดจุดตัดในการแบ่งกลุ่มตัวแปรต่อเนื่อง ถ้าเป็นไปได้ต้องกำหนดจากเหตุผลทางคลินิกหรือทางวิทยาศาสตร์ เช่น จุดตัดที่ได้จากเหตุผลทางคลินิกของความดันโลหิตซีสโตลี (> 160 mmHg และ ≤ 160 mmHg) ที่นำไปหาความสัมพันธ์กับโรคเส้นโลหิตในสมองแตก หรือในการสรุปมาตรฐานของน้ำดื่ม ปริมาณสารตะกั่วที่มีอยู่ในน้ำดื่ม จุดตัดคือ 0.05 มก./ล.

ในกรณีที่ยังไม่มีข้อมูลทางคลินิกหรือทางวิทยาศาสตร์ในการกำหนดจุดตัด วิธีการทางระบาดวิทยาที่ง่ายที่สุดคือ การใช้ quartiles ในการกำหนดจุด ซึ่งจะแบ่งจำนวนตัวอย่างออกเป็น 4 กลุ่ม นอกจากนี้ยังสามารถกำหนดจุดตัดโดยใช้วิธีการทางสถิติคำนวณหาจุดตัด optimal ที่สามารถทำนายความสัมพันธ์ได้ดีที่สุด เช่น ORC curve หรือการที่นักวิจัยกำหนดจุดตัดเองหลาย ๆแบบเพื่อหาจุดตัดที่มีความสัมพันธ์มากที่สุด

Using Statistics in Reseach 1_1 .indd 107 dummy 4 27/8/2557 11:52:44

ปัญหาที่พบในการวิเคราะห์ข้อมูลกลุ่มที่ได้จากการแบ่งกลุ่มตัวแปรต่อเนื่อง มีดังนี้

1) **ข้อมูลแต่ละประเภทมีระดับความละเอียดของการวัดแตกต่างกัน** ข้อมูลต่อเนื่องมีระดับการวัดที่ละเอียด เมื่อแปลงเป็นข้อมูลอันดับที่มีความละเอียดต่ำกว่าจะทำให้ผลการวิเคราะห์สูญเสียอำนาจการทดสอบ (power) ในการระบุความต่างการทดสอบของการเปรียบเทียบและความกระชับ (precision) ในการประมาณค่าพารามิเตอร์ของประชากร เช่น การทดสอบความต่างของความปวดหลังการผ่าตัดช่องท้องระหว่างหญิงกับชาย ถ้านำคะแนนที่ได้มาใช้สถิติ t ทดสอบว่าหญิงและชายมีค่าเฉลี่ยความปวดต่างกันหรือไม่ จะมีโอกาสพบความแตกต่างมากกว่านำคะแนนความปวดมาจัดเป็นกลุ่มปวดน้อยและปวดมาก แล้วนำมาทดสอบว่าชายและหญิงปวดมากต่างกันหรือไม่
2) **การกำหนดจุดตัดแบ่งกลุ่มเอง** อาจมีผลทำให้ข้อมูลที่มาจากตัวอย่างชุดเดียวกันมีโอกาสพบทั้งความสัมพันธ์เชิงบวกและความสัมพันธ์เชิงลบ ทั้งนี้ ขึ้นอยู่กับจุดที่กำหนด ซึ่งอคติที่เกิดขึ้นนี้มีเหตุจากนักวิจัยเห็นลักษณะการแจกแจงก่อนแบ่งกลุ่มจึงเลือกแบ่งกลุ่มให้มีผลในทิศทางที่นักวิจัยต้องการ ในการลดอคตินี้ทำได้โดยนักวิจัยกำหนดจุดตัดไว้ล่วงหน้าในแผนการวิเคราะห์ของโครงร่างวิจัย
3) **ความไม่เหมาะสมในการแปลผล** เช่น ในการวัดทักษะชีวิตด้วยการทดสอบ 100 ข้อ โดยมีระบบการให้คะแนนข้อละ 1 คะแนน ผลการวัดที่ได้จากแต่ละบุคคลจะมีค่าตั้งแต่ 0 (ศูนย์) ถึง 100 คะแนน ซึ่งจะทำให้ข้อมูลที่ได้ต้องใช้วิธีการวิเคราะห์แบบข้อมูลต่อเนื่อง ถ้านักวิจัยแบ่งคะแนนที่ได้เป็น 5 กลุ่มโดยใช้ค่าพิสัยหารด้วย 5 เช่น ค่าต่ำสุด 20 สูงสุด 45 พิสัยจะ

Using Statistics in Reseach 1_1 .indd 108 dummy 4 30/6/2557 11:29:34

เท่ากับ 45 – 20 = 25 จะได้ช่วงคะแนนในแต่ละกลุ่มดังนี้ 20–24, 25–29, 30–34, 35–39 และ 40–45 โดยกำหนดการแปลผลทั้ง 5 กลุ่มเป็น น้อยมาก น้อย ปานกลาง ดี และดีที่สุด การกำหนดกลุ่มโดยใช้ค่าสถิติแบ่งตามลักษณะการกระจายของข้อมูลของตัวอย่าง ไม่ได้สะท้อนเกณฑ์การวัดทักษะชีวิตที่มาจากเหตุผลทางวิชาการ บางการศึกษาที่กำหนดจุดตัดจากข้อมูลอาจพบว่ากลุ่มดีที่สุดได้คะแนนไม่ถึงครึ่งหนึ่งของคะแนนรวมซึ่งทำให้การแปลผลไม่ตรงความจริง

ในการแบ่งกลุ่มเครื่องมือวัดที่เป็นมาตรฐาน เช่น การวัดคุณภาพชีวิต การกำหนดกลุ่มระดับคะแนนจะทำโดยใช้เหตุผลทางวิชาการเป็นเกณฑ์ในการแบ่งกลุ่ม เช่น เครื่องมือวัดคุณภาพชีวิต WHOQOL BREF คะแนน 26–60 = คุณภาพชีวิตไม่ดี คะแนน 61–95 = คุณภาพชีวิตกลางๆ และคะแนน 96–130 = คุณภาพชีวิตที่ดี การจัดกลุ่มจะทำโดยเทียบกับการวัดด้วยเครื่องมือที่ใช้เป็นเกณฑ์

4) **การกำหนดจำนวนกลุ่มไม่เหมาะสม** ในการแบ่งกลุ่ม นักวิจัยต้องพิจารณาให้มีจำนวนกลุ่มที่เหมาะสมกับเนื้อหาของเรื่องที่จะวิเคราะห์ โดยใช้ความรู้ทางคลินิกหรือทางวิทยาศาสตร์ในการกำหนดจำนวนกลุ่ม เช่น ในการศึกษาเพื่อประเมินว่าอาหารเสริมธาตุเหล็กจะช่วยลดการขาดธาตุเหล็กได้หรือไม่ ถ้านักวิจัยแบ่งระดับการขาดธาตุเหล็กเป็น 2 กลุ่ม คือ กลุ่มขาดธาตุเหล็กกับกลุ่มปกติ โดยไม่ดูเนื้อหาว่าอาหารเสริมธาตุเหล็กมีผลเฉพาะกลุ่มที่ขาดธาตุเหล็กมาก (severe) เท่านั้น การวิเคราะห์อาจสรุปว่าอาหารเสริมไม่ได้ช่วยลดการขาดธาตุเหล็ก ซึ่งเป็นข้อสรุปที่ไม่ตรงกับความจริง

Using Statistics in Reseach 1_1 .indd 109 dummy 4 27/8/2557 11:52:44

สรุป

การแบ่งกลุ่มข้อมูลตัวแปรต่อเนื่องช่วยให้การนำเสนอข้อมูลสามารถสื่อลักษณะได้ชัดเจนและเข้าใจได้ง่าย จุดตัดที่ใช้แบ่งกลุ่มจะต้องกำหนดจากเหตุผลทางคลินิกหรือทางวิทยาศาสตร์ การวิเคราะห์ปัจจัยเสี่ยงในกรณีที่ตัวแปรอิสระเป็นตัวแปรต่อเนื่อง การแบ่งกลุ่มตัวแปรอิสระตามระดับความเสี่ยงจะช่วยในการระบุความเสี่ยงได้ถูกต้อง

Using Statistics in Reseach 1_1 .indd 110 dummy 4 30/6/2557 11:29:34

26

การวิเคราะห์ข้อมูลจากแบบวัด

การวัดตัวแปรที่ไม่สามารถใช้คำถามเพียงคำถามเดียววัดได้ เช่น คุณภาพชีวิต ความพึงพอใจ ทัศนคติ ฯลฯ ในการวัดตัวแปรดังกล่าว นักวิจัยจะสร้างแบบวัดซึ่งเป็นชุดคำถามที่มีระบบการให้ค่าคะแนนของแต่ละข้อคำตอบ แล้วนำคะแนนแต่ละข้อมารวมเพื่อสรุปเป็นผลการวัดตัวแปรดังกล่าว

ค่าที่วัดได้จากแบบวัดเป็นคะแนนรวม ซึ่งมีการใช้สถิติในการวิเคราะห์สรุปผลหลายลักษณะ ตัวอย่างแบบวัดคุณภาพชีวิตของผู้ป่วยโรคหนึ่งมีคำถามทั้งหมด 20 ข้อ แต่ละข้อมีคำตอบ 5 ระดับ ถ้าตอบทุกข้อคะแนนสูงสุดคือ 100 คะแนน ต่ำสุดคือ 20 คะแนน

คำถามที่เกี่ยวข้องกับการใช้สถิติและการวิเคราะห์ที่พบบ่อยมีดังนี้

คำถาม: ข้อมูลที่ได้จากแบบวัดที่มีคำตอบตัวเลือกเป็น rating scale เป็นข้อมูลอันดับหรือข้อมูลอัตราส่วน

คำตอบ: คำตอบจากคำถามแต่ละข้อจะเป็นข้อมูลอันดับ ถ้านักวิจัยกำหนดให้อันดับต่ำสุดมีค่า 1 คะแนน และแต่ละ

อันดับห่างกัน 1 คะแนน ถ้ามี 5 ตัวเลือกอันดับสูงสุดจะมีค่า 5 คะแนน เมื่อนำคำตอบจากทุกข้อมาคิดเป็นคะแนนรวม คะแนนรวมที่ได้จะเป็นข้อมูลอัตราส่วน ซึ่งสามารถนำไปวิเคราะห์ได้เช่นเดียวกันกับข้อมูลต่อเนื่อง

คำถาม: นำคะแนนของทุกข้อมารวมกันแล้วนำไปหาค่าเฉลี่ยต่อข้อก่อนนำไปวิเคราะห์สรุปผล ทำได้หรือไม่

คำตอบ: การนำคะแนนทั้งชุดมารวมกันแล้วทำเป็นค่าเฉลี่ยต่อข้อก่อนนำไปวิเคราะห์ ไม่มีผลกระทบต่อการคำนวณและการสรุปผลทางสถิติ แต่อาจมีความผิดพลาดในการแปลผล เช่น จากตัวอย่างคะแนนรวมต่างกัน 6 คะแนน แต่ค่าเฉลี่ยต่อข้อจะต่างเพียง 0.3 การแปลความต่างเป็นจุดทศนิยมจะรู้สึกน้อย

คำถาม: ควรมีการนำคะแนนรวมของแต่ละคนที่ได้มาจัดเป็นกลุ่มก่อนนำไปวิเคราะห์หรือไม่ เช่น ข้อมูลคุณภาพชีวิตของแต่ละคนมาจัดเป็น 3 กลุ่ม คือ คุณภาพชีวิตดี คุณภาพชีวิตปกติ และคุณภาพชีวิตไม่ดี

คำตอบ: ถ้านักวิจัยมีเป้าหมายชัดเจนว่าคำตอบจากงานวิจัยจะนำไปแก้ปัญหาอะไร ในการวิเคราะห์โดยจัดกลุ่มจะช่วยให้สามารถระบุกลุ่มปัญหาที่จะแก้ได้ เช่น ต้องการจัดกิจกรรมช่วยเหลือผู้ที่มีคุณภาพชีวิตไม่ดี การจัดเป็นกลุ่มก่อนวิเคราะห์จะสามารถระบุปริมาณของผู้ที่มีคุณภาพชีวิตไม่ดีว่ามีเท่าไร

Using Statistics in Reseach 1_1 .indd 112 dummy 4 30/6/2557 11:29:34

คำถาม: มีวิธีจัดกลุ่มคะแนนคำตอบอย่างไร

คำตอบ: ถ้าเป็นแบบวัดมาตรฐานจะมีเกณฑ์การจัดกลุ่มคะแนนคำตอบที่มีการอ้างอิง (reference criteria) กับวิธีการวัดแบบอื่น เช่น แบบวัดคุณภาพชีวิต SF36 การแบ่งกลุ่มต้องแบ่งตามเกณฑ์ที่เครื่องมือนั้นกำหนด

ในกรณีที่ไม่มีเกณฑ์ สามารถใช้วิธีการทางสถิติแบ่งกลุ่มได้ 2 วิธี คือ วิธีอิงเกณฑ์และวิธีอิงกลุ่ม การใช้วิธีการทางสถิติแบ่งกลุ่มมีจุดอ่อนอย่างมากเรื่องการแปลผล เพราะแต่ละกลุ่มที่แบ่งอิงกลุ่มตามข้อมูลอาจไม่ตรงหรือไม่สะท้อนความหมายที่แท้จริง

การจัดกลุ่มโดยวิธีอิงเกณฑ์นักวิจัยต้องกำหนดเกณฑ์ ถ้าเครื่องมือสร้างจากข้อคำถามที่เป็น rating scale อาจนำเอาเกณฑ์การให้คะแนนแต่ละระดับมาเป็นเกณฑ์สรุปผล โดยนำคะแนนรวมมาหาค่าเฉลี่ยเป็นรายข้อ แล้วแปลผลตามระดับคะแนนคำตอบ (1–5) เช่น 4.5 ขึ้นไป = พอใจมาก, 3.5–4.5 = พอใจ, 2.5–3.5 = เฉยๆ, 1.5–2.5 ไม่พอใจ, ต่ำกว่า 1.5 = ไม่พอใจมาก แล้วจึงนำค่าเฉลี่ยของคะแนนรวมของแต่ละคนมาจัดกลุ่มตามเกณฑ์ที่กำหนด การจัดกลุ่มแบบนี้อาจกำหนดเกณฑ์ขึ้นเองหรือนำเกณฑ์ที่ใช้ในงานวิจัยอื่นมาใช้ก็ได้ ในกรณีต้องนำผลที่ได้ไปเปรียบเทียบกับงานวิจัยอื่นควรใช้เกณฑ์ที่เหมือนกัน

อีกวิธีคือวิธีอิงกลุ่ม ทำโดยนำค่าพิสัย (คะแนนสูงสุด – คะแนนต่ำสุด) มาแบ่งเป็นช่วงเท่ากันตามจำนวนกลุ่มที่ต้องการแปลผล หรือแบ่งให้แต่ละช่วงห่าง

Using Statistics in Reseach 1_1 .indd 113 dummy 4 30/6/2557 11:29:34

กันตามขนาดของส่วนเบี่ยงเบนมาตรฐานกับค่าเฉลี่ย

การกำหนดกลุ่มคะแนนคำตอบด้วยวิธีการทางสถิติต่าง ๆ เป็นการแบ่งเบื้องต้นตามลักษณะคะแนนคำตอบที่ได้เท่านั้น ซึ่งอาจไม่ตรงกับลักษณะปัญหาที่เป็นจริง การจัดกลุ่มคำตอบที่มีการเทียบกับ reference criteria จึงเป็นวิธีการที่ดีที่สุด

คำถาม: การวิเคราะห์สรุปผล ควรใช้ค่าเฉลี่ยคะแนนที่วัดได้ หรือเอาข้อมูลที่ได้จากการจัดกลุ่มคำตอบตามเกณฑ์ก่อน

คำตอบ: ในกรณีที่ต้องการดูภาพรวม ไม่มีเป้าหมายเฉพาะในการใช้ผลงานวิจัย จะวิเคราะห์ด้วยค่าเฉลี่ยอย่างเดียว ซึ่งนักวิจัยต้องมีความระมัดระวังในการแปลสรุปผล เช่น จากตัวอย่างคุณภาพชีวิตเฉลี่ย กลุ่มทดลอง 84.8 กลุ่มควบคุม 80.3 ผลการทดสอบสมมติฐานพบว่ามีความต่างกันอย่างมีนัยสำคัญ ถ้าสรุปว่าคุณภาพชีวิตกลุ่มทดลองดีกว่าอาจไม่ถูกต้อง เพราะถ้าพิจารณาระดับคะแนนเฉลี่ยของทั้ง 2 กลุ่มแล้วอยู่ในระดับดีถึงดีมากเหมือนกัน

ในกรณีที่มีการจัดกลุ่มคะแนนคำตอบต้องนำข้อมูลจัดกลุ่มมาวิเคราะห์สรุปผล จะช่วยให้การสรุปผลทำได้ชัดเจนมากยิ่งขึ้น เช่น ผู้ป่วยในกลุ่มทดลองมีคุณภาพชีวิตไม่ดีน้อยกว่ากลุ่มควบคุมร้อยละ 25 (95% CI 22–27.8) ทำให้เห็นว่าถ้านำวิธีใหม่ไปให้บริการจะช่วยให้ผู้ป่วยมีคุณภาพชีวิตดีเพิ่มขึ้นมากน้อยเท่าไร

ในกรณีที่นักวิจัยใช้ทั้งค่าเฉลี่ยและข้อมูลจากการจัดกลุ่มคะแนนคำตอบไปใช้สรุปผลทั้ง 2 อย่างก็

Using Statistics in Reseach 1_1 .indd 114 dummy 4 27/8/2557 11:52:44

ทำได้ แต่ต้องระมัดระวังในการแปลผลที่ได้ให้สอดคล้องกัน

สรุป

การวิเคราะห์ข้อมูลจากแบบวัดด้วยค่าเฉลี่ยต้องมีความระมัดระวังในการแปลผล ค่าเฉลี่ยให้ภาพสรุปรวมทำให้มีข้อจำกัดในการระบุกลุ่มปัญหา ในการจัดกลุ่มคะแนนคำตอบจะช่วยระบุขนาดของกลุ่มที่มีระดับ (เช่น คุณภาพชีวิต) ต่างๆ ทำให้ข้อสรุปจากงานวิจัยสามารถนำไปใช้พัฒนางานได้โดยตรง

Using Statistics in Reseach 1_1 .indd 115 dummy 4 30/6/2557 11:29:35

27

การใช้สถิติสัมประสิทธิ์สหสัมพันธ์อย่างไม่เหมาะสม

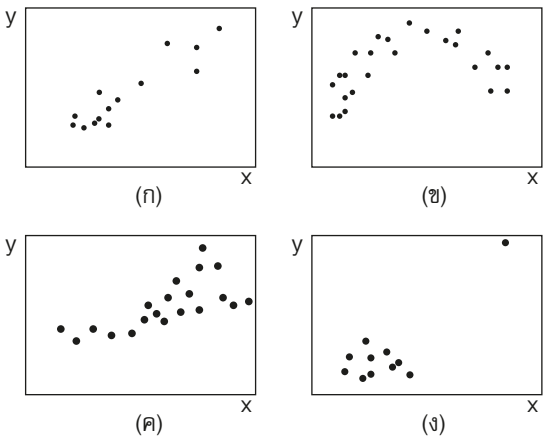
งานวิจัยทางด้านพฤติกรรมสุขภาพมีการใช้สถิติสัมประสิทธิ์สหสัมพันธ์วิเคราะห์หาความสัมพันธ์ระหว่างตัวแปร จากบทความวิจัยที่ตีพิมพ์ยังพบว่ามีการใช้สัมประสิทธิ์สหสัมพันธ์อย่างไม่เหมาะสมหลายกรณีที่มีผลต่อความถูกต้องของข้อสรุป

สัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรสุ่ม x และ y คำนวณจากสัดส่วนความแปรปรวนร่วมของ x และ y หารด้วยผลคูณของส่วนเบี่ยงเบนมาตรฐานของ x และส่วนเบี่ยงเบนมาตรฐานของ y

$$\text{correlation } (x, y) = \frac{\text{covariance } (x, y)}{\sqrt{\text{variance } (x) \times \text{variance } (y)}}$$

ค่าสัมประสิทธิ์สหสัมพันธ์ที่คำนวณจากสูตรดังกล่าวใช้แสดงความสัมพันธ์เชิงเส้นตรงระหว่าง 2 ตัวแปร โดยตัวแปร x และ y ที่เป็นข้อมูลต่อเนื่องมาจากหน่วยวัดเดียวกันและมีการแจกแจงปกติ (bivariate normal distribution) ค่าสัมประสิทธิ์สหสัมพันธ์ใช้วัดว่า 2 ตัวแปรมีความสัมพันธ์เชิงเส้นตรงในระดับใด

ความไม่เหมาะสมของการใช้สัมประสิทธิ์สหสัมพันธ์ (r) ที่พบมากที่สุดคือการนำเสนอค่า r และผลการทดสอบว่ามีนัยสำคัญโดยไม่แสดงแผนภาพการกระจาย (scatter plot) หรือไม่ระบุการตรวจสอบว่าลักษณะความสัมพันธ์เป็นเชิงเส้นตรงหรือไม่ โปรแกรมวิเคราะห์ทางสถิติคำนวณค่า r โดยไม่ได้ดูว่าตัวแปรทั้งสองมีลักษณะความสัมพันธ์เป็นเชิงเส้นตรงหรือไม่ ดังนั้นถ้านักวิจัยดูเฉพาะค่า r และ P value ที่คำนวณได้แล้วนำมาสรุปผลงานวิจัยโดยไม่พิจารณาแผนภาพการกระจายว่าตัวแปรทั้งสองมีลักษณะความสัมพันธ์เป็นแบบใดอาจทำให้สรุปความสัมพันธ์ผิด



ภาพ 1 แผนภาพการกระจายแสดงความสัมพันธ์ 2 ตัวแปร

จากภาพ 1 (ก) แผนภาพการกระจายแสดงว่าตัวแปรทั้งสองมีความสัมพันธ์เชิงเส้นตรง ใช้ค่า r สรุปความสัมพันธ์ได้

ส่วนภาพ 1 (ข) ตัวแปรทั้งสองไม่ได้สัมพันธ์เชิงเส้นตรงถึงแม้ได้ค่า r = 0.4, P value < 0.001 จะใช้ค่า r สรุปความสัมพันธ์ไม่ได้

Using Statistics in Reseach 1_1 .indd 117 dummy 4 27/8/2557 11:52:44

ความสัมพันธ์ของข้อมูลที่มีลักษณะต่างกันเป็นช่วงๆ ดังในภาพ 1 (ค) ในช่วงที่ตัวแปร x มีค่าน้อย จะไม่เหมือนกับช่วง x มีค่ามาก (heteroscedasticity) ความสัมพันธ์ในลักษณะนี้ไม่เหมาะที่จะใช้ r แสดงความสัมพันธ์ สถิติ r เหมาะสมกับข้อมูลที่ทุกช่วงมีลักษณะความสัมพันธ์เหมือนกัน (homoscedasticity) เช่น ภาพ 1 (ก)

ภาพ 1 (ง) ข้อมูลที่ศึกษาบางค่ามีค่าห่างผิดปกติจากกลุ่ม (outliers) รวมอยู่ด้วยจะมีผลกระทบต่อค่า r ที่คำนวณได้ ทั้งที่จะทำให้ค่า r มีค่ามากขึ้นหรือน้อยลง

จะเห็นได้ว่าลักษณะความสัมพันธ์มีผลต่อการแปลผลค่า r ดังนั้นนักวิจัยต้องสั่งให้โปรแกรมวิเคราะห์สร้างแผนภาพการกระจายเพื่อใช้พิจารณาประกอบกับค่า r และ P value จะทำให้ข้อสรุปที่ได้ถูกต้อง ในการนำเสนอผลการศึกษาที่ใช้ค่า r แสดงความสัมพันธ์นักวิจัยควรนำเสนอแผนภาพการกระจายประกอบด้วย

อีกกรณีหนึ่งของการใช้ไม่เหมาะสมคือการใช้สถิติ r แสดงค่าความสอดคล้อง (agreement) ของการวัดผล 2 วิธี เช่น การศึกษาความสอดคล้องของผลการวัดความดันโลหิตสูงด้วยเครื่องอัตโนมัติ (x) กับการวัดโดยเจ้าหน้าที่ (y) ค่า r ที่ได้จะแปลว่าผลการวัด 2 วิธีแปรไปในทิศทางเดียวกัน หากเครื่องอัตโนมัติวัดได้ค่าสูง เจ้าหน้าที่ก็วัดได้ค่าสูงด้วย ซึ่งไม่ได้แปลว่าค่าทั้งสองใกล้เคียงกัน

ความสอดคล้องของการวัดเป็นการดูว่าผลการวัดด้วยเครื่องอัตโนมัติสอดคล้อง (เท่ากับหรือใกล้เคียง) กับผลการวัดโดยเจ้าหน้าที่หรือไม่ สถิติที่ใช้วัดความสอดคล้องคือ intraclass correlation coefficient หรือ concordance correlation coefficient

การศึกษาความสัมพันธ์ของตัวแปรที่ได้จากการวัดซ้ำ (repeated measure) ซึ่งมีรูปแบบความสัมพันธ์แบบมีเงื่อนไข (condition) ดังนั้นจึงไม่เหมาะที่จะใช้ค่า r ในการวิเคราะห์ความสัมพันธ์ ต้องใช้ concordance

Using Statistics in Reseach 1_1 .indd 118 dummy 4 27/8/2557 11:52:44

correlation coefficient หรือ mixed effects model วิเคราะห์ความสัมพันธ์ที่ได้จากการวัดซ้ำ

การอธิบายผลความสัมพันธ์ที่วิเคราะห์ด้วยสถิติ r นอกจากระบุระดับความสัมพันธ์แล้ว นักวิจัยควรอธิบายว่าความสัมพันธ์ที่พบเป็นเชิงสาเหตุและผล หรือเป็นความสัมพันธ์ทางอ้อมที่ตัวแปรทั้งคู่ไปมีความสัมพันธ์กับตัวแปรอีกตัวหนึ่ง เช่น ในการศึกษาคุณภาพชีวิตของคนงาน พบว่าสมรรถนะการทำงานมีความสัมพันธ์ระดับสูงกับคุณภาพชีวิต นักวิจัยจึงเสนอให้พัฒนาสมรรถนะการทำงานเพื่อเพิ่มคุณภาพชีวิตให้ดีขึ้น ในความเป็นจริงสมรรถนะการทำงานมีความสัมพันธ์กับเงินเดือน และคุณภาพชีวิตก็มีความสัมพันธ์กับเงินเดือน จึงมีผลทำให้สมรรถนะการทำงานมีความสัมพันธ์กับคุณภาพชีวิต แต่การเสนอให้พัฒนาสมรรถนะการทำงานให้ดีขึ้นโดยไม่เพิ่มเงินเดือนตามสมรรถนะการทำงานที่เพิ่มขึ้น คุณภาพชีวิตจะไม่ดีขึ้นตามสมรรถนะการทำงาน

สรุป

ในการนำเสนอผลการวิเคราะห์ด้วยสถิติสัมประสิทธิ์สหสัมพันธ์ต้องนำเสนอคู่กับแผนภาพการกระจายเพื่อให้แปลผลได้ถูกต้อง การวัดความสอดคล้องและการวัดซ้ำไม่เหมาะที่จะใช้สถิติสัมประสิทธิ์สหสัมพันธ์ในการวิเคราะห์ การแปลผลสถิติสัมประสิทธิ์สหสัมพันธ์ต้องพิจารณาว่าเป็นความสัมพันธ์ในลักษณะใด จะช่วยให้แปลผลและให้ข้อเสนอแนะได้ถูกต้อง

Using Statistics in Reseach 1_1 .indd 119 dummy 4 30/6/2557 11:29:35

28

การวิเคราะห์ทางสถิติสำหรับข้อมูลที่ศึกษาทั้งประชากร

งานวิจัยเกือบทั้งหมดเป็นการศึกษาจากตัวอย่าง การวิเคราะห์ทางสถิติทำใน 2 ลักษณะ คือ ใช้สถิติพรรณนาบรรยายลักษณะตัวอย่าง และใช้สถิติอนุมานสรุปผลการศึกษาจากตัวอย่างไปเป็นข้อสรุปของประชากร แต่บางกรณีนักวิจัยสามารถเก็บข้อมูลจากทุกหน่วยศึกษาของประชากร จะใช้สถิติอะไรในการสรุปผลการศึกษา

ค่าสถิติที่คำนวณจากทุกหน่วยศึกษาในตัวอย่างเพื่อใช้สรุปลักษณะของตัวอย่าง เช่น ค่าเฉลี่ย ($\bar{x}$) ส่วนเบี่ยงเบนมาตรฐาน (s) ค่าสัดส่วน (p) เรียกว่า ค่าสถิติ (statistics) จะใช้อักษรโรมันเป็นสัญลักษณ์ ส่วนค่าทางสถิติที่ใช้สรุปลักษณะของประชากรที่คำนวณจากทุกหน่วยศึกษาในประชากร เรียกว่า ค่าพารามิเตอร์ (parameter) เช่น ค่าเฉลี่ย (μ) ส่วนเบี่ยงเบนมาตรฐาน (σ) ค่าสัดส่วน (π) ใช้อักษรกรีกเป็นสัญลักษณ์

งานวิจัยที่ศึกษาด้วยตัวอย่างจะนำข้อมูลจากตัวอย่างมาคำนวณค่าสถิติ เช่น ค่าเฉลี่ยของตัวอย่าง ($\bar{x}$) แล้วนำค่าเฉลี่ยของตัวอย่างที่ได้ไปประมาณ 95% ช่วงเชื่อมั่นของค่าเฉลี่ยของประชากร (μ) หรือจาก $\bar{x}$ ที่ได้นำไปทดสอบสมมติฐาน (H_0: $\mu = 60$, H_1: $\mu < 60$) เพื่อสรุป

ว่าค่าเฉลี่ยของประชากรมีค่าน้อยกว่า 60 อย่างมีนัยสำคัญที่ระดับความเชื่อมั่น 95% หรือไม่

งานวิจัยที่เก็บข้อมูลจากทุกหน่วยศึกษาในประชากร เมื่อใช้สถิติสรุปลักษณะข้อมูลของทุกหน่วยศึกษาของประชากร ค่าที่คำนวณได้จะเป็นค่าพารามิเตอร์ ดังนั้นการสรุปผล นักวิจัยไม่ต้องประมาณค่าหรือทดสอบสมมติฐานกับข้อมูลดังกล่าว ตัวอย่างการศึกษาเรื่องการสวมใส่หมวกนิรภัยในวิทยาลัยแห่งหนึ่ง ผู้วิจัยเก็บรวบรวมข้อมูลจากนักศึกษาทั้งวิทยาลัย การวิเคราะห์เพื่อตอบคำถามว่าสัดส่วนของนักศึกษาหญิงที่สวมใส่หมวกนิรภัยในขณะขับขี่จักรยานยนต์ต่างจากนักศึกษาชายหรือไม่ ถ้าผลการวิเคราะห์พบว่าสัดส่วนของการสวมใส่หมวกนิรภัยในขณะขับขี่จักรยานยนต์ของนักศึกษาหญิงเท่ากับ 0.32 ของนักศึกษาชายเท่ากับ 0.21 ผลต่างของสัดส่วน ($\pi_{ญ} - \pi_{ช}$) เท่ากับ 0.11 จะสรุปว่าสัดส่วนของนักศึกษาหญิงที่สวมใส่หมวกนิรภัยฯมากกว่านักศึกษาชายร้อยละ 11 จริงโดยไม่ต้องทดสอบสมมติฐาน หาค่า P value หรือหาค่า 95% ช่วงเชื่อมั่น

ในการวิเคราะห์หาความสัมพันธ์ระหว่างตัวแปรสามารถใช้สถิติที่วิเคราะห์หาความสัมพันธ์ที่ใช้กับการศึกษาจากตัวอย่าง เช่น สัมประสิทธิ์สหสัมพันธ์ สมการถดถอย Relative Risk (RR) สมการถดถอยลอจิสติก มาใช้กับการวิเคราะห์ข้อมูลทั้งประชากรและผลการวิเคราะห์ที่ได้สามารถบอกความสัมพันธ์ในประชากรได้โดยตรงโดยไม่ต้องทดสอบสมมติฐาน เช่น คำนวณค่าสหสัมพันธ์ $\rho = 0.8$ สามารถสรุปได้เลยว่าตัวแปรทั้งสองมีความสัมพันธ์ในระดับสูง

Using Statistics in Reseach 1_1 .indd 121 dummy 4 30/6/2557 11:29:35

ตาราง 1 สถิติที่ใช้ในการวิเคราะห์ของข้อมูลวิจัย

การวิเคราะห์	ศึกษาจากตัวอย่าง		ศึกษาทั้งประชากร
	ค่าสถิติ	ประมาณค่าพารามิเตอร์	ค่าพารามิเตอร์
ค่าเฉลี่ย	$\bar{x}$	95% CI of μ	μ
ค่าสัดส่วน	p	95% CI of π	π
เปรียบเทียบค่าเฉลี่ย	$\bar{x}_1 - \bar{x}_2$	test hypothesis (P value) หรือ 95% CI of $\mu_1 - \mu_2$	$\mu_1 - \mu_2$
เปรียบเทียบค่าสัดส่วน	$p_1 - p_2$	test hypothesis (P value) หรือ 95% CI of $\pi_1 - \pi_2$	$\pi_1 - \pi_2$
ความสัมพันธ์ตัวแปรต่อเนื่อง 2 ตัว	r	test hypothesis (P value) หรือ 95% CI of ρ	ρ
	$y = a + bx$	Confidence bands	$Y = \alpha + \beta x$
ความสัมพันธ์ตัวแปรกลุ่ม 2 ตัว	RR	95% CI of RR	RR
	logistic regression	test hypothesis (P value) หรือ 95% CI of RR	logistic regression

จากตาราง 1 แสดงสถิติที่ใช้ในการวิเคราะห์ของข้อมูลวิจัยที่ศึกษาจากตัวอย่างและศึกษาทั้งประชากร ในการหาความสัมพันธ์ของตัวแปรกลุ่ม 2 ตัวจะใช้ Odds Ratio (OR) ไม่ได้ เพราะ OR ใช้ในกรณีที่ไม่ทราบอัตราการเกิดโรคที่แท้จริง ในกรณีที่ศึกษาทั้งประชากรจะสามารถคำนวณอัตราการเกิดโรคที่แท้จริงได้ จึงใช้ RR ในการแสดงขนาดความสัมพันธ์

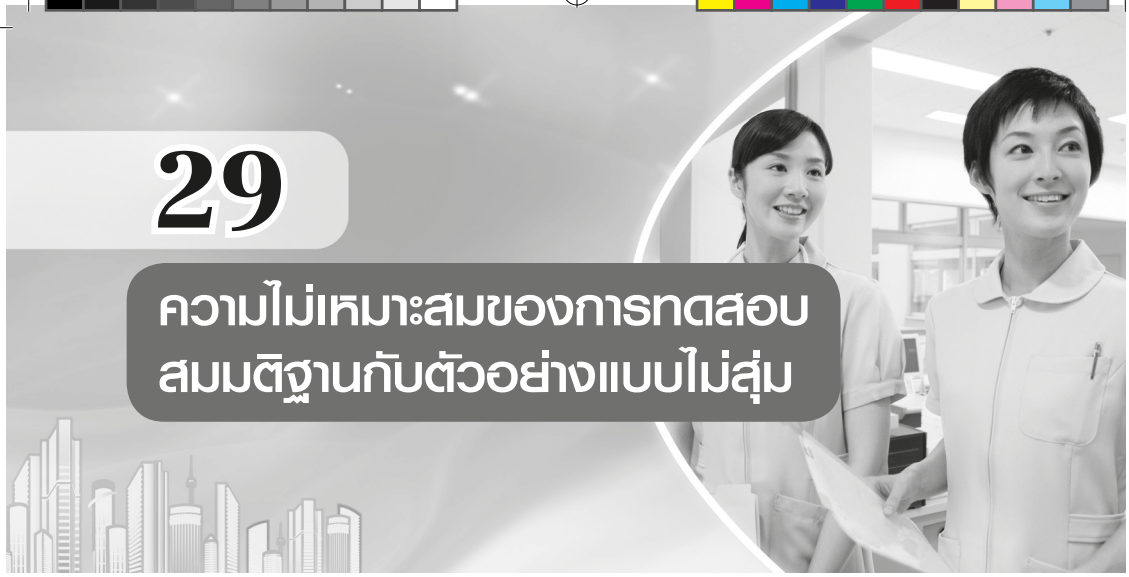
Using Statistics in Reseach 1_1 .indd 122 dummy 4 27/8/2557 11:52:44

โดยปกติในการศึกษาทั้งประชากร คำถามงานวิจัยจะเป็นการหาขนาดของปัญหา และมุ่งที่จะหากลุ่มย่อยที่เกิดปัญหา เพื่อจะได้หาวิธีการแก้ไขปัญหานั้น โดยไม่ต้องการดูความสัมพันธ์ เช่น การศึกษาสมรรถนะการทำงานของข้าราชการสังกัดกระทรวงสาธารณสุขภายในจังหวัด ผู้วิจัยส่งแบบสอบถามไปเก็บข้อมูลจากเจ้าหน้าที่ทุกคนที่ทำงานอยู่ในจังหวัด (ทั้งประชากร) การวิเคราะห์ข้อมูลจะทำให้ทราบว่ามีผู้ที่มีสมรรถนะการทำงานต่ำกว่าที่กำหนดมีมากน้อยเท่าไร และเป็นผู้ที่สังกัดหน่วยงานใด หรือมีลักษณะอย่างไร เพื่อจะได้จัดกิจกรรมในการแก้ไขปัญหาโดยตรง มากกว่าจะตั้งคำถามว่าสมรรถนะการทำงานมีความสัมพันธ์กับเพศหรือไม่

สรุป

ในการใช้สถิติสรุปลักษณะของข้อมูลที่ได้จากการศึกษาทั้งประชากร นักวิจัยสามารถคำนวณค่าพารามิเตอร์ต่างๆได้โดยตรง ไม่ต้องใช้การประมาณค่าหรือทดสอบสมมติฐานในการสรุปผลการศึกษาของประชากร

Using Statistics in Reseach 1_1 .indd 123 dummy 4 30/6/2557 11:29:35



29

ความไม่เหมาะสมของการทดสอบสมมติฐานกับตัวอย่างแบบไม่สุ่ม

การเลือกตัวอย่างสำหรับการวิจัยโดยปกติจะเลือกโดยวิธีสุ่มหรือแบบใช้ความน่าจะเป็น เช่น การสุ่มแบบง่าย การสุ่มแบบระบบ เป็นต้น มีบางงานวิจัยที่ใช้การเลือกตัวอย่างแบบไม่สุ่ม (non-random sampling) หรือแบบไม่ใช้ความน่าจะเป็น เช่น การเลือกแบบบังเอิญ การเลือกแบบเจาะจง หรือการเลือกแบบโควตา จึงมีคำถามว่าจะใช้การทดสอบสมมติฐานกับข้อมูลที่ได้จากตัวอย่างแบบไม่สุ่มเพื่อสรุปผลการศึกษาของประชากรได้หรือไม่

ทฤษฎีทางสถิติที่ใช้ทดสอบสมมติฐานค่าสถิติของตัวอย่างให้เป็นข้อสรุปของประชากรมีข้อตกลงพื้นฐานเบื้องต้นว่า ข้อมูลที่นำมาทดสอบสมมติฐานต้องมาจากตัวอย่างแบบสุ่ม ตัวอย่างแบบสุ่ม หมายถึงหน่วยตัวอย่างที่ใช้ในการศึกษาต้องสุ่มมาจากประชากร โดยที่แต่ละหน่วยศึกษามีโอกาสถูกเลือกเป็นตัวอย่างเท่ากัน หรือทราบความน่าจะเป็นของโอกาสที่ถูกเลือก ตัวแปรที่ได้จากตัวอย่างดังกล่าวเรียกว่า ตัวแปรสุ่ม การเลือกตัวอย่างแบบสุ่มจะทำให้ได้ตัวอย่างที่ไม่เอนเอียง ค่าเฉลี่ยที่คำนวณจากตัวอย่างแบบสุ่มจะเป็นค่าที่ไม่เอนเอียง (unbiased) ทำให้การทดสอบสมมติฐานได้ข้อสรุปของประชากรที่ตรงกับความจริง

การทดสอบสมมติฐาน ไม่ว่าจะเป็นแบบ parametric เช่น t-test, paired t-test, ANOVA หรือแบบ non-parametric เช่น Mann-Whitney test, Wilcoxon signed-rank test, Kruskal-Wallis test มีข้อตกลงเบื้องต้นว่าข้อมูลจากตัวอย่างที่นำมาทดสอบสมมติฐานต้องมาจากตัวอย่างแบบสุ่ม การจะเลือกใช้สถิติแบบ parametric หรือ non-parametric ขึ้นอยู่กับข้อตกลงเบื้องต้นของการแจกแจงของตัวแปรที่นำมาทดสอบสมมติฐาน ถ้าเป็น parametric ต้องทราบว่าหน่วยตัวอย่างถูกสุ่มมาจากประชากรที่มีการแจกแจงแบบใด หรือมีข้อตกลงเบื้องต้นว่าประชากรมีการแจกแจงแบบใด สำหรับ non-parametric ไม่มีข้อตกลงเบื้องต้นในเรื่องการแจกแจงของประชากร ดังนั้นการทดสอบสมมติฐานไม่ว่าเป็น parametric หรือ non-parametric ข้อมูลที่จะนำมาทดสอบจะต้องได้มาจากตัวอย่างแบบสุ่ม ต่างกันเพียงเรื่องข้อตกลงเบื้องต้นของการแจกแจงของประชากรเท่านั้น

การเลือกตัวอย่างแบบบังเอิญเป็นการเลือกตัวอย่างแบบไม่สุ่ม ทำให้ได้ตัวอย่างที่เอนเอียง (biased) เมื่อนำค่าสถิติที่คำนวณจากตัวอย่างที่เอนเอียงมาสรุปผลการศึกษาจะได้คำตอบที่เอนเอียงไม่ตรงกับความจริง ตัวอย่างเช่น ในการศึกษาเรื่องความเห็นของการทำแท้งเสรีของประชาชนที่อาศัยอยู่ในเขตเทศบาล นักวิจัยใช้วิธีการเลือกตัวอย่างแบบบังเอิญโดยเก็บข้อมูลจากผู้ที่มาซื้ออาหารที่ตลาดสดเทศบาลทั้ง 2 แห่ง ในวันจันทร์ พุธ และศุกร์ ระหว่างเวลา 6.00-12.00 น. เป็นเวลา 1 สัปดาห์

ถ้าพิจารณาลักษณะคนที่จะมาซื้ออาหารที่ตลาดสดในวันธรรมดาอาจเป็นกลุ่มที่ทำงานบ้านเป็นหลัก ผู้ที่ทำงานสำนักงานอาจมาซื้อในวันเสาร์หรืออาทิตย์ นอกจากนี้ บางคนไม่ได้ซื้ออาหารจากตลาดสด แต่ซื้อจากแหล่งอื่น เช่น ร้านค้าในชุมชน ตลาดเช้า หรือซูเปอร์มาร์เกต แผนการเลือกตัวอย่างนี้ยังทำให้ได้ตัวอย่างที่เป็นผู้ชาย ผู้สูงอายุ หรือ

Using Statistics in Reseach 1_1 .indd 125 dummy 4 30/6/2557 11:29:36

วัยรุ่นน้อย เพราะคนกลุ่มนี้จะไม่ค่อยมาตลาดในช่วงเวลาเก็บข้อมูล และผู้ที่ไม่ได้เป็นผู้ซื้ออาหารจะไม่มีโอกาสถูกเลือกเป็นตัวอย่าง

ดังนั้นจะเห็นได้ว่าตัวอย่างที่เลือกได้จากแผนการสุ่มดังกล่าวมีความเอนเอียง ไม่สามารถใช้เป็นตัวแทนที่ดีของประชาชนที่อาศัยอยู่ในเขตเทศบาล ถ้านักวิจัยนำข้อมูลที่เก็บได้ทดสอบสมมติฐานทางสถิติจะมีความไม่เหมาะสม 2 ประการ คือ

1) ตัวอย่างที่ได้เป็นตัวอย่างที่เอนเอียงจะทำให้ได้ข้อสรุปที่เอนเอียง
2) เมื่อตัวอย่างที่ศึกษาไม่ทราบความน่าจะเป็นของโอกาสที่ถูกเลือกเป็นตัวอย่าง ทำให้ไม่สามารถใช้ความน่าจะเป็นในการอนุมานผลการทดสอบของตัวอย่างให้เป็นข้อสรุปของประชากร

สรุป

นักวิจัยที่ใช้การเลือกตัวอย่างแบบบังเอิญหรือวิธีเลือกตัวอย่างแบบไม่ใช้ความน่าจะเป็นอื่นๆ ตัวอย่างที่ได้จะมีความเอนเอียง ทำให้ได้คำตอบงานวิจัยที่เอนเอียงไม่ตรงกับความจริง ตามทฤษฎีค่าสถิติที่คำนวณจากข้อมูลของตัวอย่างดังกล่าวจะไม่สามารถนำมาทดสอบสมมติฐานเพื่อใช้สรุปผลการศึกษาในประชากร

Using Statistics in Reseach 1_1 .indd 126 dummy 4 30/6/2557 11:29:36

30

การทดสอบความแตกต่างกับการสรุปความสัมพันธ์ระหว่างตัวแปร

ในการทดสอบความแตกต่างของค่าเฉลี่ยของตัวแปรต่อเนื่องว่าแตกต่างกันระหว่างกลุ่มของตัวแปรกลุ่มแบบมี 2 ค่า (dichotomous variable) เช่น ความเร็วของการขับขี่รถจักรยานยนต์มีความแตกต่างตามเพศ เมื่อพบว่าแตกต่างกันอย่างมีนัยสำคัญ นอกจากจะสรุปว่าความเร็วของการขับขี่รถจักรยานยนต์แตกต่างตามเพศแล้ว จะสรุปว่ามีความสัมพันธ์กันได้หรือไม่

การทดสอบสมมติฐานของค่าเฉลี่ยของความเร็วในการขับขี่รถจักรยานยนต์ (ตัวแปรต่อเนื่อง) มีความแตกต่างระหว่างเพศ (ตัวแปรกลุ่มแบบมี 2 ค่า) มีผลการศึกษาแสดงในตาราง 1

ตาราง 1 ช่วงเชื่อมั่นของผลต่างความเร็ว
ในการขับขี่รถจักรยานยนต์ระหว่างเพศชายและเพศหญิง

เพศ	n	$\bar{x}$ (SD) กม./ชม.	95% CI
ชาย	40	90 (7)	30.9–39.1
หญิง	40	55 (11)	

จากผลการทดสอบด้วยสถิติ t พบว่าค่าขีดจำกัดชั้นล่างของ 95% ช่วงเชื่อมั่นมีค่ามากกว่าศูนย์ (0) แสดงว่าความเร็วเฉลี่ยการขับรถของเพศชาย (90 กม./ชม.) มากกว่าเพศหญิง (55 กม./ชม.) อย่างมีนัยสำคัญ

จากข้อค้นพบดังกล่าว เมื่อเห็นผู้ขับขี่รถมาด้วยความเร็วสูงอาจใช้ผลสรุปจากการศึกษาทำนายว่าผู้ขับขี่จะต้องเป็นเพศชาย ในทางกลับกันถ้าผู้ขับขี่รถมาด้วยความเร็วต่ำก็จะทำนายว่าผู้ขับขี่เป็นเพศหญิง การที่ทราบค่าของตัวแปรหนึ่ง (ความเร็วในการขับขี่รถ) แล้วสามารถทำนายค่าของอีกตัวแปรหนึ่ง (เพศ) ได้ แสดงว่าเพศมีความสัมพันธ์กับความเร็วในการขับขี่รถจักรยานยนต์

ในกรณีที่ตัวแปรทั้งสองเป็นตัวแปรกลุ่มแบบมี 2 ค่า เช่น การเคยได้รับอุบัติเหตุจากการขับขี่รถจักรยานยนต์ กับการสวมใส่หมวกนิรภัยขณะขับขี่ มีผลการศึกษาแสดงในตาราง 2

ตาราง 2 แสดงร้อยละของการสวมใส่หมวกนิรภัย กับการเคยได้รับอุบัติเหตุ

การได้รับอุบัติเหตุ	การสวมใส่หมวกนิรภัย จำนวน (ร้อยละ)		95% CI	χ^2 P value
	สวม	ไม่สวม		
เคย	178 (79.1)	47 (20.9)	48.2–65.1	< 0.001
ไม่เคย	42 (22.5)	145 (77.5)		

จากตาราง 2 ผลการทดสอบความแตกต่างของสัดส่วนการสวมใส่หมวกนิรภัยระหว่างผู้ที่เคยและไม่เคยได้รับอุบัติเหตุด้วยสถิติ Z โดยพิจารณาค่าช่วงเชื่อมั่น พบว่าค่าขีดจำกัดชั้นล่างของ 95% ช่วงเชื่อมั่นมีค่ามากกว่าศูนย์ (0) แสดงว่าการสวมใส่หมวกนิรภัยขณะขับขี่ของผู้ที่

Using Statistics in Reseach 1_1 .indd 128 dummy 4 27/8/2557 11:52:44

เคยได้รับอุบัติเหตุ (79.1%) มากกว่าผู้ที่ไม่เคยได้รับอุบัติเหตุ (22.5%) อย่างมีนัยสำคัญ หรือสามารถสรุปได้ว่าการเคยได้รับอุบัติเหตุจากการขับขี่รถจักรยานยนต์จะมีความสัมพันธ์กับการสวมหมวกนิรภัยขณะขับขี่

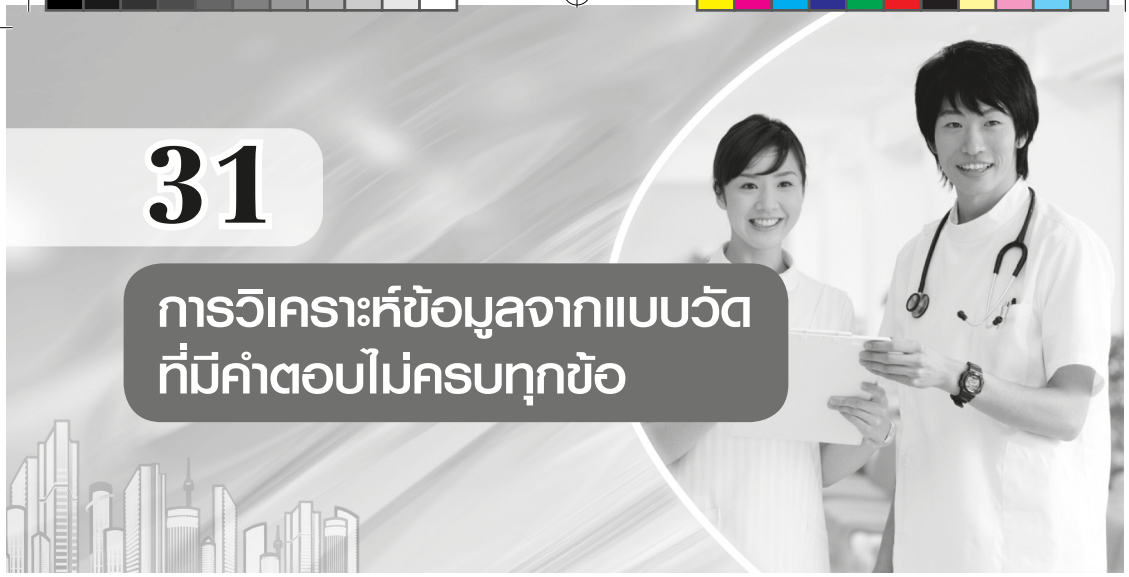
การทดสอบสมมติฐานกรณีที่ตัวแปรทั้งสองเป็นตัวแปรกลุ่มแบบมี 2 ค่า ยังสามารถใช้สถิติ χ^2 ทดสอบความสัมพันธ์ได้ด้วย ผลการวิเคราะห์ในตาราง 2 ค่า P value ของการทดสอบ χ^2 มีค่า < 0.0001 สามารถสรุปได้ว่าการสวมใส่หมวกนิรภัยขณะขับขี่มีความสัมพันธ์กับการเคยได้รับอุบัติเหตุจากการขับขี่รถจักรยานยนต์ การทดสอบสมมติฐานในกรณีนี้ทำได้ทั้งสถิติ Z และ χ^2 เพราะทั้ง 2 วิธีใช้หลักการวิเคราะห์ทางสถิติเดียวกัน ดังนั้นผลการวิเคราะห์ทดสอบความแตกต่างของสัดส่วนด้วยสถิติ Z หรือการวิเคราะห์หาความสัมพันธ์ด้วยสถิติ χ^2 จะให้ข้อสรุปที่ตรงกันทุกกรณี

แต่การวิเคราะห์ด้วยสถิติ Z นอกจากจะให้ผลการทดสอบสมมติฐานแล้ว ยังสามารถคำนวณค่าช่วงเชื่อมั่นของผลต่างระหว่างค่าสัดส่วน ทำให้สามารถระบุขนาดของความต่าง ซึ่งสถิติ χ^2 คำนวณช่วงเชื่อมั่นไม่ได้ ดังนั้นในการทดสอบความแตกต่างระหว่างค่าสัดส่วน หรือทดสอบความสัมพันธ์ในกรณีตัวแปรทั้งสองเป็นตัวแปรกลุ่มแบบมี 2 ค่า จึงควรใช้สถิติ Z ในการวิเคราะห์

สรุป

ในการแปลผลการทดสอบความแตกต่างของค่าเฉลี่ยหรือค่าสัดส่วนระหว่าง 2 ประชากร ถ้าพบว่ามีความแตกต่างอย่างมีนัยสำคัญ นอกจากจะสรุปว่าค่าเฉลี่ยหรือค่าสัดส่วนทั้งสองแตกต่างกันแล้ว ยังสามารถสรุปได้ว่าตัวแปรทั้งสองมีความสัมพันธ์กัน

Using Statistics in Reseach 1_1 .indd 129 dummy 4 30/6/2557 11:29:36



31

การวิเคราะห์ข้อมูลจากแบบวัดที่มีคำตอบไม่ครบทุกข้อ

แบบวัด (scales) เป็นชุดคำถามเพื่อใช้วัดค่าตัวแปรที่ไม่สามารถวัดค่าได้ด้วยคำถามเพียงข้อเดียว เช่น แบบวัดคุณภาพชีวิตขององค์การอนามัยโลกอย่างย่อมีคำถามทั้งหมด 26 ข้อ เพื่อใช้วัดคุณภาพชีวิตโดยรวม ในแบบวัดยังแบ่งออกเป็นองค์ประกอบต่างๆ 4 ด้าน เช่น องค์ประกอบด้านสุขภาพกาย มีคำถามทั้งหมด 7 ข้อ เป็นข้อคำถามเชิงลบ 2 ข้อ ข้อคำถามเชิงบวก 5 ข้อ ฯลฯ

เมื่อมีคำถามบางข้อที่ผู้ตอบไม่ได้ตอบทำให้ไม่สามารถให้ค่ากับคำตอบข้อนั้นๆ มีผลทำให้คะแนนรวมที่ได้มีคะแนนน้อยลง ถ้านำผลรวมเท่าที่ได้มาใช้เป็นค่าของตัวแปรที่วัดจะทำให้ไม่ถูกต้อง เพราะผู้ที่ไม่ตอบจะมีคะแนนรวมน้อยกว่าความจริง

ความไม่เหมาะสมที่พบบ่อยในการวิเคราะห์ข้อมูลแบบวัดที่มีคำตอบไม่ครบทุกข้อ คือ การนำค่าเฉลี่ยของข้อที่มีคำตอบในชุดเดียวกันมาแทนค่าข้อมูลที่สูญหาย โดยไม่พิจารณาเหตุของการที่ไม่ตอบ ทำให้ได้ผลรวมที่มีอคติ

ข้อมูลสูญหายที่เกิดจากการตอบคำถามในแบบวัดไม่ครบแบ่งออกเป็น 2 ประเภท คือ สูญหายเกิดอย่างสุ่มและสูญหายเกิดอย่างไม่สุ่ม

ข้อมูลสูญหายที่เกิดอย่างสุ่ม เกิดจากผู้ตอบไม่ได้ตั้งใจที่จะไม่ตอบข้อนั้น แต่อาจตอบข้ามบางข้อด้วยความสะเพร่าหรือด้วยความรีบร้อน หรือการวางรูปแบบการจัดพิมพ์ไม่ดี บรรทัดติดกันเกินไปทำให้ทำข้ามข้อโดยไม่ตั้งใจ

ข้อมูลสูญหายเกิดอย่างไม่สุ่ม เกิดจากผู้ตอบไม่ต้องการตอบคำถามบางข้อ เช่น แบบประเมินคุณภาพชีวิตขององค์การอนามัยโลกด้านความสัมพันธ์ทางสังคม มี 1 ข้อประเมินความพอใจในชีวิตทางเพศ คำถามข้อนี้ผู้หญิงจะลำบากใจในการตอบ โดยเฉพาะผู้ที่ไม่เคยมีเพศสัมพันธ์ จึงมีโอกาสไม่ตอบข้อนี้ ข้อมูลสูญหายเกิดอย่างไม่สุ่มจึงมีสาเหตุมาจากคำถามที่ผู้ตอบรู้สึกลำบากใจหรืออับอายที่จะตอบ ส่วนใหญ่จะเกี่ยวกับการถามเรื่องเพศ รูปลักษณ์ ความสามารถ และระดับสติปัญญา

อีกสาเหตุหนึ่งเกิดจากคำถามไม่ดี ผู้ตอบไม่เข้าใจ หรือคำถามไม่เหมาะกับบริบทของผู้ตอบทำให้ผู้ตอบไม่สามารถตอบได้ เช่น ถามผู้สูงอายุว่า *"ท่านมีปัญหาในการเดินขึ้นลงบันไดในการดำเนินชีวิตประจำวันหรือไม่"* ถ้าผู้สูงอายุที่อาศัยอยู่ในบ้านชั้นเดียว หรือการดำเนินชีวิตปกติอยู่แต่ชั้นล่าง จะตอบคำถามข้อนี้ไม่ได้

การวิเคราะห์ข้อมูลแบบวัดที่มีข้อมูลบางข้อสูญหาย อาจทำได้โดย

1) **ตัดชุดที่มีข้อมูลไม่ครบออกจากการวิเคราะห์** มีผลทำให้ตัวอย่างเล็กลงและข้อมูลหลายส่วนที่เก็บมาได้ถูกทิ้งไปโดยไม่ได้ใช้ประโยชน์
2) **ใช้วิธีการทางสถิติคำนวณค่าแทนข้อมูลที่สูญหาย** ใช้ได้ดีกับการสูญหายที่เกิดอย่างสุ่ม ถ้าข้อมูลสูญหายเกิดอย่างไม่สุ่ม ค่าที่

Using Statistics in Reseach 1_1 .indd 131 dummy 4 30/6/2557 11:29:36

คำนวณแทนอาจทำให้ได้ผลรวมที่มีอคติ ดังนั้นการใช้วิธีการทางสถิติคำนวณค่าแทนจึงควรมีการพิจารณาลักษณะข้อมูลที่สูญหายว่าเกิดอย่างสุ่มหรือไม่ ดังนี้

2.1) ลักษณะของกลุ่มผู้ที่ไม่ตอบข้อนี้มีลักษณะเหมือนกลุ่มผู้ตอบหรือไม่ เช่น การไม่ตอบคำถามความพอใจในชีวิตทางเพศ ถ้าพบว่ากลุ่มไม่ตอบเป็นหญิงโสดมากกว่ากลุ่มที่ตอบ การจะนำค่าเฉลี่ยของผู้ตอบที่ส่วนใหญ่เป็นคนที่มีประสบการณ์ทางเพศมาเป็นค่าแทนกลุ่มที่ไม่ตอบ ทำให้ผลรวมที่ได้มีอคติ

2.2) ข้อคำถามที่มีผู้ไม่ตอบมากกว่าข้ออื่น อาจเกิดจากเขียนคำถามไม่ดี มีผู้ตอบส่วนหนึ่งไม่เข้าใจเลยไม่ตอบ หรือรูปแบบการจัดพิมพ์ไม่ดีทำให้มีโอกาสตอบข้ามข้อนี้ไปโดยไม่ตั้งใจ

2.3) การไม่ตอบข้อคำถามอาจมีปัจจัยเฉพาะในเรื่องที่ศึกษา เช่น ในการศึกษาคุณภาพชีวิตของผู้ป่วยโดยผู้ให้การรักษาพยาบาล การไม่ตอบของผู้ป่วยแต่ละคนอาจมีเหตุผลที่ต่างกัน ดังนั้นการใช้ค่าเฉลี่ยของกลุ่มเป็นค่าแทนข้อมูลที่สูญหายจะมีโอกาสได้ผลรวมที่มีอคติ

ในกรณีคำถามที่ไม่มีคำตอบมีไม่เกินร้อยละ 1 ของคำถามทั้งหมด และนักวิจัยพิจารณาแล้วเห็นว่าการสูญหายน่าจะเกิดอย่างสุ่ม สามารถใช้ค่าเฉลี่ยของกลุ่มที่ตอบในข้อเดียวกันแทนได้ ไม่ควรใช้ค่าเฉลี่ยของคำตอบที่ได้จากแบบวัดทั้งชุด เพราะในแบบวัดแต่ละชุดจะมีองค์ประกอบหลายด้าน และข้อคำถามมีทั้งที่เป็นเชิงบวกและลบ การนำค่าเฉลี่ยของคำตอบทั้งชุดมาแทนค่าที่สูญหายย่อมทำให้ได้ผลรวมที่มีอคติ

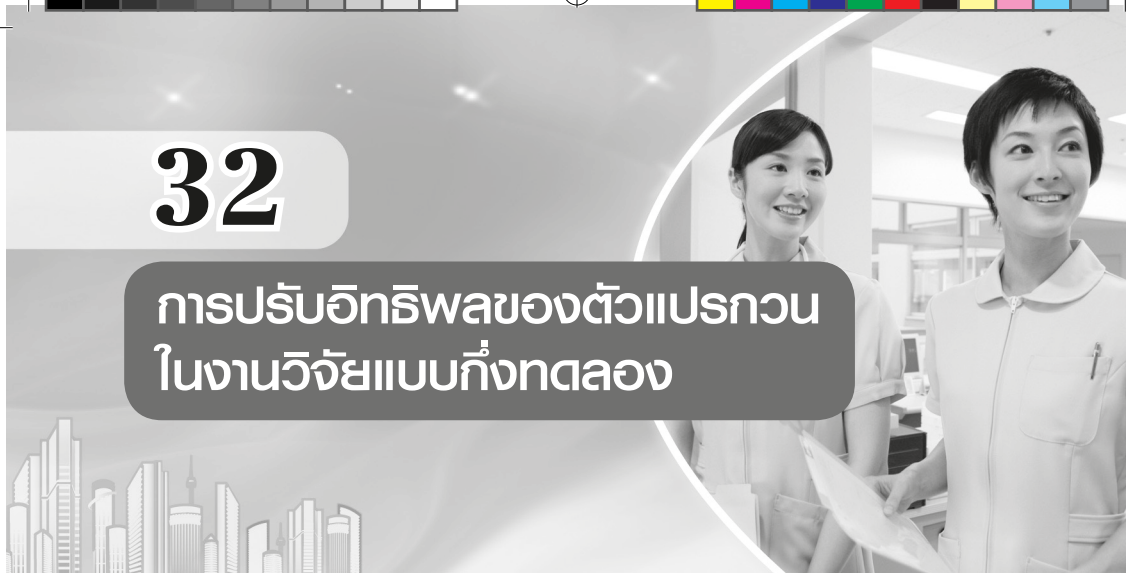
Using Statistics in Reseach 1_1 .indd 132 dummy 4 30/6/2557 11:29:36

ในการใช้วิธีการทางสถิติคำนวณค่าของข้อมูลสูญหายมีหลายวิธี ความเหมาะสมของแต่ละวิธีจะขึ้นอยู่กับลักษณะของการไม่ตอบ ความสัมพันธ์ของตัวแปรที่มีข้อมูลสูญหายกับตัวแปรอื่นที่เก็บมาศึกษา และจำนวนข้อที่สูญหาย ในการดำเนินการควรปรึกษานักสถิติในการเลือกวิธีการคำนวณค่าแทนข้อมูลสูญหาย

สรุป

การวัดผลโดยใช้แบบวัดเมื่อมีข้อมูลสูญหายบางข้อจากการไม่ตอบ เพื่อไม่ให้ได้ผลรวมที่มีอคติ นักวิจัยต้องพิจารณาว่าการไม่ตอบเกิดอย่างสุ่มหรือไม่ ก่อนที่จะเลือกใช้วิธีการทางสถิติที่เหมาะสมในการคำนวณค่าแทนข้อมูลสูญหาย

Using Statistics in Reseach 1_1 .indd 133 dummy 4 30/6/2557 11:29:36



32

การปรับอิทธิพลของตัวแปรกวนในงานวิจัยแบบกึ่งทดลอง

ในงานวิจัยกึ่งทดลองมักจะมีปัญหาเรื่องของอาสาสมัคร (ผู้ที่เข้าร่วมในการทดลอง) ในกลุ่มทดลองและกลุ่มควบคุมมีลักษณะต่างกัน ซึ่งอาจเป็นตัวกวนที่ส่งผลต่อผลการทดลอง ทำให้ความแตกต่างที่พบมีบางส่วนเกิดจากตัวแปรกวนที่กระจายอยู่ไม่เท่ากันทั้ง 2 กลุ่ม ไม่ใช่เกิดจากอิทธิพลของสิ่งทดลอง (intervention) อย่างเดียว

ในการแก้ไขปัญหาตัวแปรกวนกระจายอยู่ไม่เท่ากันในแต่ละกลุ่มทดลอง นักวิจัยต้องทราบก่อนว่าเรื่องที่ศึกษามีตัวแปรอะไรเป็นตัวแปรกวนบ้าง และอิทธิพลของตัวแปรกวนแต่ละตัวมีมากน้อยเท่าไร โดยการทบทวนวรรณกรรมจากงานวิจัยที่ศึกษาเรื่องปัจจัยความสัมพันธ์และปัจจัยเสี่ยงที่มีต่อตัวแปรผลที่ศึกษา การพิจารณาอิทธิพลของตัวแปรกวนแต่ละตัวจะพิจารณาจากค่า standardize coefficient ของตัวแปรกวนที่ได้จากงานวิจัยที่วิเคราะห์ความสัมพันธ์ด้วยตัวแบบ multiple regression หรือพิจารณาจากค่า adjusted OR จากงานวิจัยที่วิเคราะห์หาปัจจัยเสี่ยงด้วยตัวแบบ multiple logistic regression

การปรับอิทธิพลของตัวแปรกวนในงานวิจัยแบบกึ่งทดลองทำได้ใน 2 ขั้นตอน คือ ขั้นตอนแรกทำในช่วงการสุ่มอาสาสมัคร (ผู้เข้ารับการทดลอง) ให้หน่วยทดลอง หรือสุ่มกลุ่มทดลองให้อาสาสมัคร โดยใช้

วิธีสุ่มแบบชั้นภูมิ (stratified random sampling) อาสาสมัครจะถูกแบ่งเป็นกลุ่มๆ (ชั้นภูมิ) ของตัวแปรกวนที่สำคัญ เช่น ในการศึกษาเรื่องการสอนให้กินอาหารที่เหมาะสมในผู้ป่วยเบาหวาน จากงานวิจัยที่ผ่านมาพบว่าระดับความรุนแรงของโรคเป็นปัจจัยสำคัญมากต่อการเลือกกินอาหารที่เหมาะสม จึงได้แบ่งอาสาสมัครออกเป็น 3 กลุ่มตามระดับความรุนแรงน้อย ปานกลาง และมาก แล้วจึงสุ่มอาสาสมัครในแต่ละกลุ่มความรุนแรงให้เข้าอยู่ในกลุ่มทดลองหรือกลุ่มควบคุม การสุ่มด้วยวิธีนี้จะช่วยให้มีผู้ป่วยทุกระดับความรุนแรงกระจายอยู่ในกลุ่มทดลองและกลุ่มควบคุมเท่าๆกัน

วิธีการปรับอิทธิพลของตัวแปรกวนโดยแบ่งอาสาสมัครตามชั้นภูมิของตัวแปรกวนที่สำคัญก่อนที่จะสุ่มเข้ากลุ่มการทดลอง เป็นวิธีที่ดีที่สุดที่ใช้ปรับอิทธิพลของตัวแปรกวน แต่ในทางปฏิบัติการแบ่งอาสาสมัครตามชั้นภูมิทำได้เพียง 1–2 ตัวแปรกวนเท่านั้น เพราะถ้าแบ่งมากกว่า 2 ตัวแปรกวนในบางชั้นภูมิอาจมีอาสาสมัครเพียง 1 คน ทำให้ไม่สามารถสุ่มให้กระจายไปตามกลุ่มการทดลอง หรือใช้เวลานานในการหาอาสาสมัครในแต่ละชั้นภูมิให้มีจำนวนเพียงพอกับจำนวนกลุ่ม

การปรับอิทธิพลของตัวแปรกวนยังสามารถทำได้ในอีกขั้นตอนหนึ่งคือในขั้นการวิเคราะห์ โดยปกติในกรณีที่มีตัวแปรกวนที่สำคัญหลายตัว ตัวที่สำคัญที่สุดแก้ไขในขั้นตอนการสุ่มอาสาสมัครเข้ากลุ่มการทดลอง ส่วนอิทธิพลของตัวแปรกวนส่วนที่เหลือปรับในขั้นวิเคราะห์ หรือในกรณีที่มีข้อจำกัดของการออกแบบงานวิจัย เช่น ในการทดลองการสอนสุขศึกษา 2 วิธี ความรู้ก่อนการทดลองมีผลต่อการเรียนรู้เพิ่ม แต่ในทางปฏิบัติไม่สามารถจัดให้อาสาสมัครที่ได้คะแนนความรู้เท่ากันกระจายเท่ากันทั้ง 2 กลุ่ม ทำได้มากที่สุดคือแบ่งเป็นช่วงคะแนน แต่นักวิจัยก็ไม่มั่นใจว่าเมื่อแบ่งกลุ่มตามช่วงคะแนนที่กำหนดแล้ว คะแนนความรู้ก่อนการทดลองในกลุ่มเดียวกันจะมีอิทธิพลต่างกันหรือไม่

Using Statistics in Reseach 1_1 .indd 135 27/8/2557 11:52:45
dummy 4

วิธีการวิเคราะห์ที่ใช้ปรับอิทธิพลของตัวแปรกวนทำโดยการวิเคราะห์ covariate หรือเรียกว่า covariate adjustment โดยให้ตัวแปรผลเป็นตัวแปรตาม ตัวแปรกลุ่มทดลองและตัวแปรกวนทั้งหลายเป็นตัวแปรอิสระ ในการทำ covariate adjustment ตัวแบบที่ใช้ในการวิเคราะห์จะแตกต่างกันไปตามประเภทของตัวแปรตาม ถ้าตัวแปรตามเป็นตัวแปรต่อเนื่องจะใช้ตัวแบบ multiple regression แต่ถ้าตัวแปรตามเป็นตัวแปรกลุ่มแบบมี 2 ค่าจะใช้ตัวแบบ multiple logistic regression การปรับโดยวิธีนี้จะใช้ได้ผลดีถ้าข้อมูลที่นำมาคำนวณเป็นไปตามข้อตกลงเบื้องต้นและเงื่อนไขความสัมพันธ์ของข้อมูล

สรุป

การปรับอิทธิพลของตัวแปรกวนในงานวิจัยแบบกึ่งทดลองควรทำในขั้นการสุ่มอาสาสมัครให้กลุ่มทดลอง ถ้ามีตัวแปรกวนบางตัวที่ไม่สามารถปรับได้ในขั้นการสุ่มอาสาสมัคร หรือมีตัวแปรบางตัวกระจายอยู่ไม่เท่ากัน และนักวิจัยคาดว่าอาจมีผลกระทบต่อตัวแปรผลควรใช้วิธีการวิเคราะห์เพื่อปรับอิทธิพลของตัวแปรดังกล่าว

จากรายงานวิจัยเพื่อหาปัจจัยความสัมพันธ์และปัจจัยเสี่ยงช่วยนักวิจัยกำหนดได้ว่าตัวแปรใดเป็นตัวแปรกวน และการทบทวนวรรณกรรมที่ครอบคลุมจะช่วยในการกำหนดตัวแปรที่มีอิทธิพลได้ครบทุกตัว

Using Statistics in Reseach 1_1 .indd 136 dummy 4 30/6/2557 11:29:37

33

การวิเคราะห์ผลการทดลองในกรณีมีผู้ออกระหว่างการทดลอง

ในงานวิจัยเชิงทดลองหรือกึ่งทดลองเพื่อเปรียบเทียบวิธีการสอน ส่วนใหญ่มักจะมีอาสาสมัคร (ผู้เข้ารับการทดลอง) ออกระหว่างการทดลอง หรือเข้าร่วมการทดลองไม่ครบทุกกิจกรรมตามระยะเวลาที่กำหนดไว้ในการทดลอง การแก้ไขปัญหาที่นักวิจัยส่วนใหญ่ใช้คือการตัดอาสาสมัครเหล่านี้ออกจากการวิเคราะห์ ซึ่งไม่ถูกต้อง

อาสาสมัครที่ออกระหว่างการทดลองบางส่วนอาจมีธุรกิจด่วนทำให้ไม่สามารถเข้าร่วมงานวิจัยได้ตลอดตามที่รับปากไว้ กลุ่มที่ขาดด้วยสาเหตุนี้มีไม่มากและเป็นการออกอย่างสุ่ม ไม่เจาะจงว่ามีอยู่ในกลุ่มใดกลุ่มหนึ่งโดยเฉพาะ

อีกสาเหตุหนึ่งคือการที่อาสาสมัครเบื่อหรือไม่เห็นประโยชน์ที่ได้รับจากการเข้ารับการสอน ทั้งนี้ อาจเนื่องจากการสอนใช้เวลามากไป เนื้อหาไม่น่าสนใจ วิธีการสอนน่าเบื่อ เป็นต้น สาเหตุการออกระหว่างการทดลอง หรือเข้าร่วมบ้างไม่เข้าร่วมบ้างจะเกี่ยวข้องกับวิธีการสอนโดยตรง เป็นปัญหากับวิธีสอน การออกจากการทดลองจึงมีเฉพาะบางกลุ่ม ถ้านักวิจัยตัดอาสาสมัครที่มีปัญหากับวิธีสอนออกไปไม่นำมาวิเคราะห์ก็จะเกิดอคติในการสรุปผล เพราะอาสาสมัครที่การสอนวิธีนี้ไม่ได้ผลจะถูกตัดออกจากการสรุปผลทั้งหมด

วิธีการวิเคราะห์การทดลองที่มีอาสาสมัครออกระหว่างการทดลองโดยประยุกต์วิธีที่ใช้วิเคราะห์การทดลองทางคลินิกมี 2 วิธี คือ

1) **การวิเคราะห์โดยใช้หลัก intention to treat (ITT)** การวิเคราะห์แบบนี้จะนำอาสาสมัครทุกคน รวมทั้งที่ออกระหว่างการทดลอง และเข้าร่วมในการทดลองไม่สม่ำเสมอ เข้าร่วมในการวิเคราะห์ตามกลุ่มที่ถูกสุ่มเข้ารับการสอนตั้งแต่เริ่มแรก

อาสาสมัครที่มีการวัดผลหลังการทดลองทั้งผู้ที่เข้ารับการสอนสม่ำเสมอและไม่สม่ำเสมอ กลุ่มนี้มีข้อมูลที่จะใช้วิเคราะห์ แต่อาสาสมัครที่ออกระหว่างการทดลองทำให้ไม่สามารถวัดผลหลังการทดลอง จึงไม่มีข้อมูลไปคำนวณค่าเฉลี่ยหลังสอน อาสาสมัครกลุ่มนี้ถ้าได้รับการสอนบางส่วนควรมีความรู้เพิ่มขึ้นบ้าง หรืออย่างน้อยควรไม่ต่ำกว่าผลการวัดก่อนทดลอง ดังนั้นในการวิเคราะห์จึงใช้ผลการวัดความรู้ครั้งล่าสุด (last observation carried forward (LOCF)) แทนความรู้หลังทดลอง (ที่เก็บไม่ได้) ถ้าในการทดลองมีการวัดผล 3 ครั้ง คือ ก่อนทดลอง ช่วงกลาง และหลังทดลอง อาสาสมัครที่ออกระหว่างการทดลองที่มีผลการวัด 2 ครั้ง (ก่อนทดลองและช่วงกลาง) ในการวิเคราะห์จะใช้ผลการวัดครั้งที่ 2 ไปแทนผลการวัดหลังทดลอง ในกรณีที่มีผลการวัดก่อนทดลองเพียงครั้งเดียวก็ใช้ผลการวัดก่อนทดลองไปแทนผลการวัดหลังทดลองในการวิเคราะห์ ค่าเฉลี่ยหลังการทดลองที่มีหลักคิดแบบนี้จะเป็นผลลัพธ์ที่แย่ที่สุดของวิธีการนี้ (worst case scenario) ดังนั้นถ้าใช้วิธีวิเคราะห์แบบ ITT แล้วพบว่ากลุ่มทดลองมีผลการสอนดีกว่ากลุ่มควบคุม นักวิจัยจะสามารถเชื่อมั่นได้ว่าวิธีการสอนที่ใช้ในกลุ่มทดลองดีกว่าแน่ๆ

Using Statistics in Reseach 1_1 .indd 138 dummy 4 27/8/2557 11:54:14

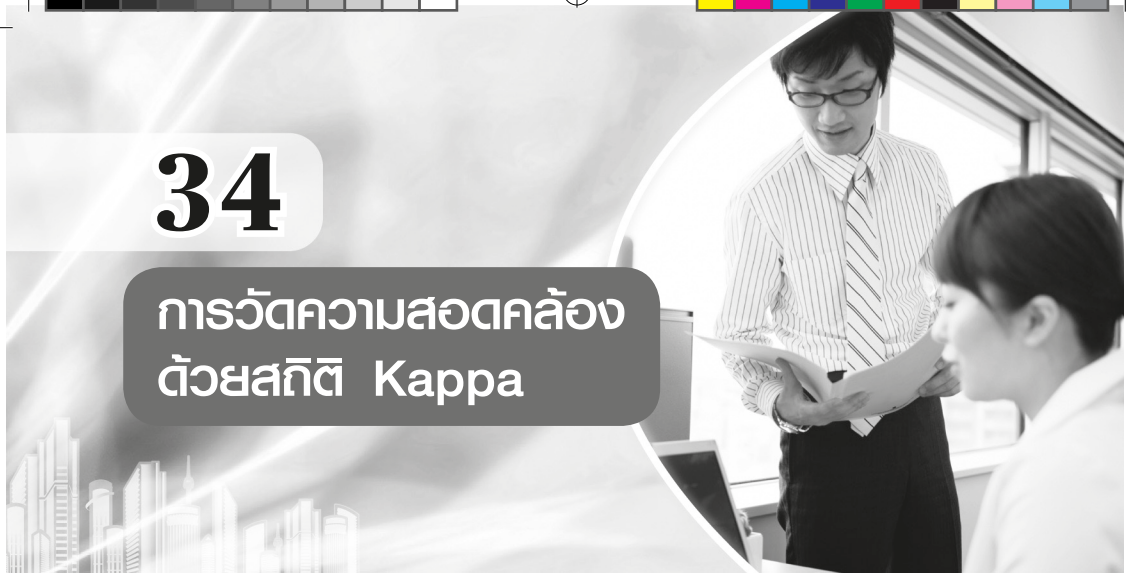
2) **การวิเคราะห์แบบตามแผนการทดลอง (per protocol analysis (PP))** เป็นการวิเคราะห์ที่นำเฉพาะอาสาสมัครที่ยินยอมรับการทดลองตามวิธีการที่กำหนดไว้ทั้งหมดของแต่ละกลุ่มมาวิเคราะห์ อาสาสมัครที่ออกระหว่างการทดลองหรือเข้าร่วมการทดลองไม่ครบทุกกิจกรรมจะถูกตัดออกจากการวิเคราะห์ หรืออาสาสมัครบางคนมาทุกครั้งแต่มาสายหรืออยู่ไม่ครบ ถ้านักวิจัยคิดว่าไม่สามารถรับประโยชน์การสอนได้ครบตามวิธีการที่กำหนดก็สามารถนำออกจากการวิเคราะห์ ผลการวิเคราะห์แบบนี้จะช่วยให้เข้าใจว่าถ้าอาสาสมัครเข้าร่วมการสอนตามวิธีการที่กำหนดไว้ทั้งหมดจะได้ผลอย่างไร

ในการนำเสนอผลการเปรียบเทียบที่ใช้สรุปผลการศึกษาจะใช้ผลการวิเคราะห์แบบ ITT เพราะจะสอดคล้องกับสถานการณ์จริงที่จะนำไปใช้มากที่สุด และค่าความต่างที่ได้เป็นความต่างที่ต่ำสุด ถ้านำไปใช้ควรจะได้ผลดีมากกว่านี้ ส่วนผลการวิเคราะห์แบบ PP จะนำเสนอประกอบเพื่อช่วยให้ผู้อ่านเข้าใจว่าถ้าสามารถหาวิธีสร้างความร่วมมือกับผู้เรียนได้ดีที่สุดแล้วกลุ่มทดลองจะได้ผลเพิ่มขึ้นเท่าใด

สรุป

การวิเคราะห์ผลการทดลองที่มีอาสาสมัครออกระหว่างการทดลอง จะให้ผลสรุปที่ถูกต้องไม่มีอคติ ต้องนำอาสาสมัครทุกคนเข้าร่วมในการวิเคราะห์ และใช้ผลการวิเคราะห์แบบ ITT สรุปผลการทดลอง โดยอาจนำเสนอผลการวิเคราะห์แบบ PP ร่วมในการสรุปและวิจารณ์ผล

Using Statistics in Reseach 1_1 .indd 139 dummy 4 30/6/2557 11:29:37



34

การวัดความสอดคล้องด้วยสถิติ Kappa

ในการศึกษาเรื่องความครอบคลุมของการค้นหาผู้ป่วยโรคเบาหวานโดยการตรวจระดับน้ำตาลในเลือด (FBS) นักวิจัยต้องการประเมินว่านักวิทยาศาสตร์การแพทย์ 2 คน ที่เป็นผู้ตรวจมีผลการตรวจเหมือนกันหรือไม่ การตรวจได้ผลเหมือนกัน หมายถึงตรวจวัดสิ่งเดียวกันให้ผลลัพธ์ที่เท่ากันหรือใกล้เคียงกัน

การประเมินว่าผู้วิเคราะห์ 2 คนตรวจวัดได้เหมือนกันเรียกว่าประเมินความสอดคล้อง (agreement) การประเมินความสอดคล้องของผู้วัดควรดำเนินการก่อนเริ่มเก็บข้อมูลวิจัย ตัวอย่างเช่น นำตัวอย่างเลือด 30 ตัวอย่าง โดยแต่ละตัวอย่างแบ่งให้นักวิทยาศาสตร์การแพทย์ 2 คน วิเคราะห์ ได้ผลการวิเคราะห์ดังนี้

ตัวอย่างที่	ผู้วิเคราะห์ 1	ผู้วิเคราะห์ 2	ตัวอย่างที่	ผู้วิเคราะห์ 1	ผู้วิเคราะห์ 2
1	132	92	16	131	128
2	99	106	17	89	89
3	102	104	18	128	126
4	80	80	19	94	94
5	125	125	20	140	142
6	94	94	21	93	95
7	72	78	22	86	88
8	107	97	23	126	124
9	121	115	24	116	120
10	84	86	25	139	136
11	132	128	26	79	82
12	115	120	27	101	104
13	87	84	28	86	88
14	93	90	29	131	130
15	79	79	30	92	91

ความไม่ถูกต้องที่พบในการใช้สถิติประเมินความสอดคล้องคือการที่นักวิจัยใช้ paired t-test เปรียบเทียบค่าเฉลี่ยผลการวัดของผู้วัด 2 คน ถ้าไม่ต่างกันจะสรุปว่าผู้วัดทั้งสองวัดได้เหมือนกัน การที่ไม่สามารถใช้ paired t-test (ทดสอบความแตกต่างของค่าเฉลี่ย) มาใช้ในการสรุปผลความสอดคล้องได้เนื่องจาก

1) ผลการตรวจไม่เหมือนกัน แต่ให้ค่าเฉลี่ยที่เท่ากันได้ เช่น จากข้อมูลตัวอย่างระดับน้ำตาลในเลือดของคนที่ 4, 5 และ 6 จากเดิมทั้ง 2 คนวัดได้ 80, 125 และ 94 ตรงกัน ถ้าเปลี่ยนให้คนที่ 2 วัดได้เป็น 84, 119 และ 96 ค่าเฉลี่ยที่ได้จะมีค่าเท่าเดิม

Using Statistics in Reseach 1_1 .indd 141 dummy 4 30/6/2557 11:29:37

(ตัวอย่างที่ 4 และ 6 มีค่าเพิ่มรวม 6 ส่วนตัวอย่างที่ 5 ได้ค่าลดลง 6) แต่ถ้าพิจารณาค่าของข้อมูลแต่ละตัวอย่างจะพบว่าความเหมือนกันเปลี่ยนไปแล้ว

2) ในการตรวจวินิจฉัยโรคเบาหวาน ความถูกต้องของการตัดสินว่าเป็นโรคหรือไม่เป็นโรคอยู่ที่ผลการตรวจช่วง 110 mg/dl ซึ่งเป็นเกณฑ์ที่ใช้ในการวินิจฉัยโรค ถึงแม้ผลการตรวจในช่วงอื่นๆ แตกต่างกันบ้าง แต่เมื่อนำมาสรุปผลการวินิจฉัยแล้วให้คำวินิจฉัยเหมือนกัน ก็ถือว่าการตรวจวิเคราะห์ของทั้ง 2 คนให้ผลสอดคล้องกัน ในกรณีที่มีเกณฑ์ควรแบ่งผลการตรวจให้เป็นกลุ่มตามเกณฑ์ก่อนนำมาหาค่าความสอดคล้อง

สถิติที่ใช้วัดความสอดคล้องในกรณีข้อมูลกลุ่มแบบมี 2 ค่า (dichotomous) คือ สถิติ Kappa คำนวณจากค่าสัดส่วนความสอดคล้องที่พบกับสัดส่วนความสอดคล้องจากความบังเอิญตามสูตรข้างล่างนี้

$$K = \frac{(\text{สัดส่วนความสอดคล้องที่พบจริง} - \text{สัดส่วนความสอดคล้องจากความบังเอิญ})}{1 - \text{สัดส่วนความสอดคล้องจากความบังเอิญ}}$$

ค่า K ที่คำนวณได้เป็นความสอดคล้องของผู้ตรวจทั้ง 2 คนที่หักความสอดคล้องที่เกิดจากความบังเอิญออกแล้ว ค่า K ที่คำนวณได้จะมีค่าไม่เกิน 1

วิธีการวิเคราะห์ต้องนำข้อมูลระดับน้ำตาลในเลือดมาแบ่งกลุ่มเป็น 2 กลุ่ม คือ กลุ่มไม่เป็นเบาหวานที่มีระดับน้ำตาลในเลือดน้อยกว่า 110 mg/dl และกลุ่มเบาหวานที่มีระดับน้ำตาลในเลือดเท่ากับหรือมากกว่า 110 mg/dl นำข้อมูลที่แบ่งกลุ่มแล้วมาทำเป็นตารางสำหรับคำนวณค่า Kappa ได้ดังนี้

Using Statistics in Reseach 1_1 .indd 142 dummy 4 27/8/2557 11:54:14

ตาราง 1 ผลการตรวจเลือดวิเคราะห์
เพื่อคัดกรองผู้ป่วยเบาหวาน จากผู้วิเคราะห์ 2 คน

		ผู้วิเคราะห์ 2	
		< 110 mg/dl	≥ 110 mg/dl
ผู้วิเคราะห์ 1	< 110 mg/dl	19 (a)	2 (b)
	≥ 110 mg/dl	1 (c)	8 (d)

สูตรคำนวณค่าสถิติ Kappa กรณีตาราง 2 × 2

$$K = \frac{2(ad - ac)}{n^2(p_1q_2 + p_2q_1)}$$

โดยที่ $p_1 = (a + b)/n$ และ $p_2 = (a + c)/n$

$q_1 = 1 - p_1$ และ $q_2 = 1 - p_2$

Altman* เสนอแนวทางการแปลผลค่า Kappa (โดยนำแนวทางที่เสนอโดย Landis and Koch** มาปรับเล็กน้อย) ไว้ดังนี้

---

* Douglas G. Altman, **Practical Statistics for Medical Research.** New York: Chapman & Hall/CRC, 1999.

** J. Richard Landis and Gary G. Koch, "The Measurement of Observer Agreement for Categorical Data." **Biometrics.** 1977; 33: 159–174.

Using Statistics in Reseach 1_1 .indd 143 dummy 4 27/8/2557 11:54:14

ค่า Kappa (K)	ระดับความสอดคล้อง
≤ 0.20	Poor
0.21–0.40	Fair
0.41–0.60	Moderate
0.61–0.80	Good
0.81–1.00	Very Good

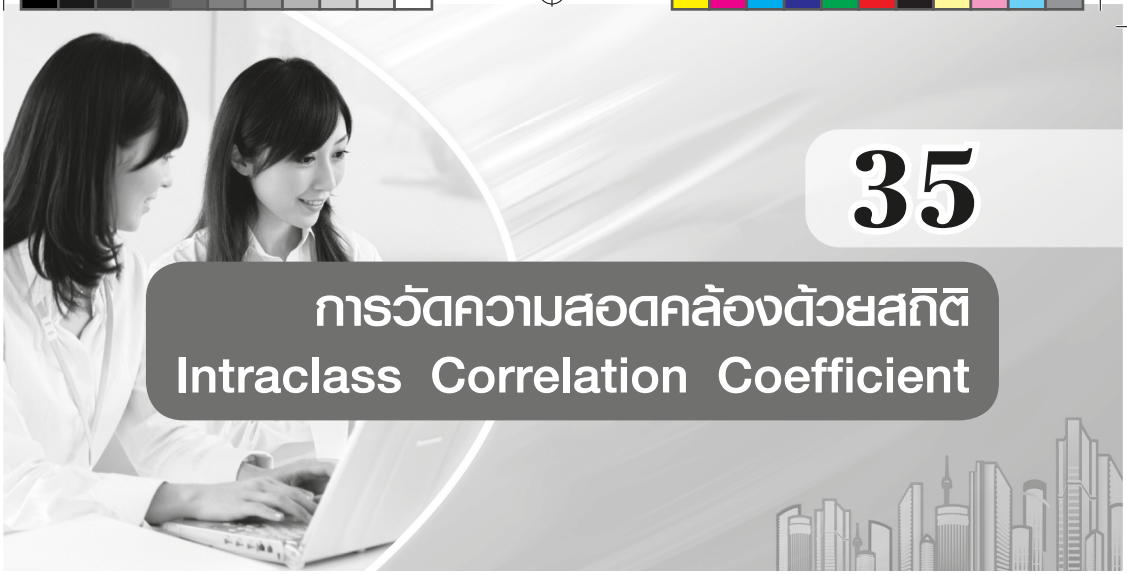
ข้อมูลจากตาราง 1 นำมาคำนวณค่า K ได้เท่ากับ 0.77 จะสรุปได้ว่าผลการวิเคราะห์ของผู้วิเคราะห์ทั้ง 2 คนมีความสอดคล้องดี

สถิติที่ใช้ในการวิเคราะห์ความสอดคล้องในกรณีมีผู้ประเมินมากกว่า 2 คน หรือมีผลการประเมินมากกว่า 2 กลุ่ม จะมีวิธีการคำนวณค่า Kappa ด้วยสูตรที่แตกต่างกัน ในกรณีที่ผลการวัดเป็นข้อมูลต่อเนื่องจะใช้สถิติ intraclass correlation coefficient ในการประเมินความสอดคล้อง ในกรณีที่ต้องการประเมินความสอดคล้องระหว่างวิธีการตรวจจะใช้วิธีการคำนวณค่า Kappa เหมือนกันโดยเปลี่ยนจากผู้ประเมินเป็นวิธีการตรวจ

สรุป

การประเมินความสอดคล้องระหว่างผู้ตรวจในกรณีที่มีเกณฑ์การวินิจฉัย ต้องนำผลที่วัดได้มาแบ่งเป็นกลุ่มตามเกณฑ์การวินิจฉัย แล้วจึงนำข้อมูลที่แบ่งกลุ่มไปคำนวณค่าสถิติ Kappa เพื่อสรุปความสอดคล้อง

Using Statistics in Reseach 1_1 .indd 144 dummy 4 27/8/2557 11:54:15



35

การวัดความสอดคล้องด้วยสถิติ Intraclass Correlation Coefficient

ในการประเมินความสอดคล้องระหว่างผู้ประเมิน 2 คนหรือมากกว่า ในกรณีที่ตัวแปรผลเป็นข้อมูลต่อเนื่อง หรือคะแนนที่วัดจากเครื่องมือวัดที่เป็น summated rating scale เช่น เครื่องมือวัดคุณภาพชีวิต เครื่องมือวัดความรู้ ทัศนคติ และการปฏิบัติตัว ฯลฯ จะมีการวิเคราะห์แตกต่างไปตามวิธีการประเมิน การสุ่มหรือไม่สุ่มผู้ประเมิน ในที่นี้จะขอยกกรณีของการประเมินตัวอย่างเดียวกันโดยผู้ประเมิน 2 คน ซึ่งใช้ในโครงการวิจัยที่จำเป็นต้องมีผู้ตรวจวิเคราะห์หรือผู้ประเมิน 2 คน และต้องการดูว่าผู้ตรวจวิเคราะห์/ประเมินทั้ง 2 คนให้ผลการตรวจที่สอดคล้องกันหรือไม่

สถิติที่ใช้ในกรณีนี้คือ intraclass correlation coefficient (ICC) โดยที่ ICC มีค่าน้อยสุดเท่ากับศูนย์ (0) มากที่สุดเท่ากับ 1 วิธีการอ่านแปลผลค่า ICC ให้ใช้เกณฑ์เดียวกับการอ่านผลค่า Kappa ที่ได้เสนออยู่ในเรื่องที่ 34 *การวัดความสอดคล้องด้วยสถิติ Kappa* การคำนวณค่า ICC ได้จากสูตรข้างล่างนี้

$$ICC = \frac{\text{variance due to subjects (patients)}}{(\text{variance due to subjects} + \text{variance due to raters} + \text{residual variance})}$$

แนวคิดของการคำนวณค่า ICC คือ เปรียบเทียบค่าความแปรปรวนที่เกิดจากความแตกต่างระหว่างคนที่มารับการตรวจ (variance due to subjects) กับความแปรปรวนรวมซึ่งมีความแปรปรวนที่เกิดจากความแตกต่างระหว่างผู้ตรวจ (variance due to raters) รวมอยู่ด้วย ดังนั้นค่า ICC จะมากหรือน้อยขึ้นอยู่กับความแปรปรวนที่เกิดจากความแตกต่างระหว่างผู้ประเมิน ถ้าความแปรปรวนระหว่างผู้ประเมินมีค่าน้อยๆ ค่า ICC จะมีค่าใกล้ 1 ถ้าความแปรปรวนระหว่างผู้ประเมินที่มีค่ามากๆ ค่า ICC จะมีค่าใกล้ศูนย์

ตัวอย่าง ในงานวิจัยเพื่อประเมินทักษะของแม่ในการดูแลลูกป่วย โดยมีพยาบาลเป็นผู้ประเมิน 2 คน ทีมวิจัยได้จัดอบรมผู้ประเมินและทดสอบผลการประเมินจากผู้ประเมินทั้ง 2 คน ได้ผลดังตาราง

ลำดับที่	ผลการประเมิน		ลำดับที่	ผลการประเมิน	
	ผู้ประเมิน 1	ผู้ประเมิน 2		ผู้ประเมิน 1	ผู้ประเมิน 2
1	9	8	16	7	7
2	6	5	17	5	6
3	8	8	18	9	8
4	7	6	19	9	8
5	5	6	20	6	6
6	10	9	21	5	6
7	6	7	22	5	5
8	5	5	23	7	6
9	4	4	24	6	5
10	5	7	25	7	6
11	5	6	26	6	5
12	8	8	27	7	7
13	9	7	28	5	6
14	8	7	29	9	8
15	6	5	30	9	8

Using Statistics in Reseach 1_1 .indd 146 dummy 4 8/9/2557 11:47:48

ในการคำนวณค่า ICC จะใช้ two-way ANOVA วิเคราะห์หาความแปรปรวน ระหว่างคนที่มารับการตรวจ ระหว่างผู้ประเมิน และจากการสุ่ม (residual) แล้วนำไปแทนค่าในสูตรคำนวณค่า ICC โปรแกรมสถิติบางโปรแกรมจะมีคำสั่งวิเคราะห์หาค่า ICC โดยตรง เช่น SPSS จะวิเคราะห์ด้วยคำสั่ง Reliability (ในเมนู เลือก Scale → Reliability analysis) นำข้อมูลจากตารางไปคำนวณค่า ICC ด้วยโปรแกรม SPSS ได้ค่า ICC เท่ากับ 0.796 โดยมีค่า 95% ช่วงเชื่อมั่นอยู่ระหว่าง 0.615 ถึง 0.897

จากค่า ICC ที่ได้แสดงว่าผลการประเมินความสอดคล้องของพยาบาลทั้ง 2 คนมีความสอดคล้องอยู่ในระดับดี แต่ค่า 95% ช่วงเชื่อมั่นกว้างเพราะขนาดตัวอย่างเล็ก แต่เนื่องจากค่าต่ำสุดของช่วงเชื่อมั่นก็อยู่ในระดับดีทำให้ไม่มีผลกระทบต่อการสรุปผล ในกรณีที่คำถามหลักของงานวิจัยเป็นการประเมินความสอดคล้อง นักวิจัยต้องทำการคำนวณขนาดตัวอย่าง* เพื่อให้ได้ขนาดตัวอย่างที่เหมาะสม และจะทำให้ได้ค่าช่วงเชื่อมั่นของ ICC ที่กระชับ

สิ่งที่นักวิจัยพึงตระหนักในการดำเนินการประเมินความสอดคล้องของผู้ประเมินหรือผู้ตรวจ คือ การฝึกอบรม และวิธีการตรวจวิเคราะห์หรือวิธีการประเมินควรเขียนเป็นวิธีมาตรฐาน (standard operating procedure) เพื่อให้ผู้ประเมินทุกคนทำตามวิธีมาตรฐานที่กำหนด ซึ่งมี

* S.D. Walter, M. Eliasziw, and A. Donner, "Sample Size and Optimal Designs for Reliability Studies." **Statistics in Medicine.** 1998; 17: 101–110.

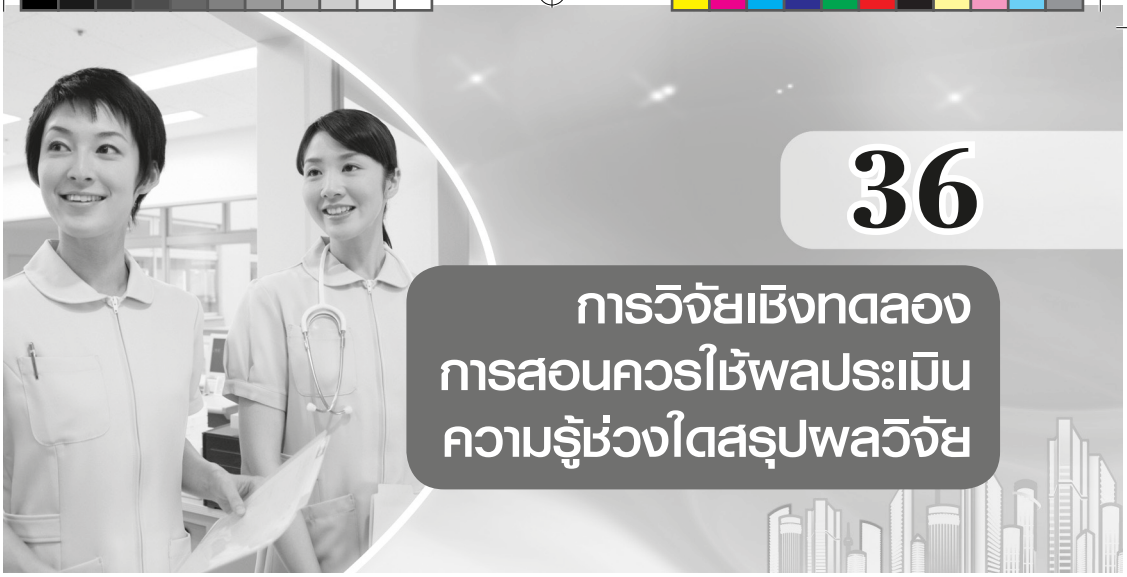
Using Statistics in Reseach 1_1 .indd 147 dummy 4 27/8/2557 11:54:15

ผลโดยตรงต่อคุณภาพของการตรวจวิเคราะห์หรือการประเมิน ในกรณีผลการประเมินมีความสอดคล้องกันน้อยควรพิจารณาว่าวิธีการตรวจวิเคราะห์หรือวิธีการประเมินมีข้อบกพร่องหรือไม่ ถ้าวิธีการชัดเจนดีแต่ผู้ประเมินมีทักษะต่างกันมาก หรือมีบางคนไม่ยินยอมทำตามวิธีมาตรฐานที่กำหนด ควรเปลี่ยนผู้ประเมิน

สรุป

กรณีข้อมูลตัวแปรผลวัดออกมาเป็นข้อมูลต่อเนื่อง หรือคะแนนจากเครื่องมือวัด (ชุดคำถาม) การประเมินความสอดคล้องของผู้ตรวจวัดทำโดยใช้สถิติ ICC ซึ่งมีค่าระหว่าง 0–1 ควรมีการคำนวณขนาดตัวอย่างเพื่อให้ได้จำนวนคนที่เหมาะสมสำหรับการประเมินความสอดคล้อง

Using Statistics in Reseach 1_1 .indd 148 dummy 4 30/6/2557 11:29:38



36

การวิจัยเชิงทดลอง การสอนควรใช้ผลประเมินความรู้ช่วงใดสรุปผลวิจัย

การทดลองเพื่อประเมินผลการสอน 2 วิธี โดยออกแบบงานวิจัยให้กลุ่มที่ 1 เป็นกลุ่มทดลอง (สอนแบบใหม่) และกลุ่มที่ 2 เป็นกลุ่มควบคุม (สอนแบบเดิม) ในการศึกษามีการประเมินความรู้ทั้งก่อนการทดลอง ระหว่างการทดลอง และหลังการทดลอง เนื่องจากมีการประเมินผลลัพธ์อยู่หลายช่วง จึงมีคำถามว่าจะใช้การวัดผลช่วงใดสรุปความสำเร็จของการทดลอง

แบบงานวิจัยที่ใช้สำหรับประเมินผลการสอน 2 กลุ่ม โดยปกติจะมีการวัดความรู้ก่อนการสอนและหลังการสอน บางกรณีอาจมีการวัดความรู้ระหว่างการสอน การที่มีการวัดความรู้หลายช่วงผลของแต่ละช่วงจะใช้ตอบคำถามที่แตกต่างกันดังนี้

1) **การวัดความรู้ก่อนการทดลอง** จะใช้ตอบคำถามว่า ตัวอย่าง 2 กลุ่มที่เลือกมามีความรู้แตกต่างกันหรือไม่ เพราะถ้าไม่เท่ากันตั้งแต่ต้นอาจจะมีผลต่อความสัมฤทธิ์ผลของการสอน เนื่องจากพื้นฐานไม่เท่ากันตั้งแต่แรก
2) **การวัดความรู้ระหว่างการทดลอง** ในกรณีที่มีการวัดความรู้เป็นช่วง ๆระหว่างการทดลอง โดยปกติจะมีจุดมุ่งหมายที่จะดูการเปลี่ยนแปลงความรู้ไปตามกิจกรรมที่ได้รับในแต่ละช่วงเวลา

Using Statistics in Reseach 1_1 .indd 149 dummy 4 30/6/2557 11:29:38

3) **การวัดความรู้โดยใช้ผลต่างระหว่างก่อนและหลังการสอนของแต่ละกลุ่ม** จะใช้ตอบคำถามว่า แต่ละกลุ่มเมื่อได้รับการสอนเสร็จสิ้นตามแผนแล้วมีความรู้เพิ่มขึ้นจากเดิมหรือไม่

4) **การวัดความรู้หลังการสอน** จะใช้ตอบคำถามว่า เมื่อสิ้นสุดการทดลอง กลุ่มทดลองมีความรู้มากกว่ากลุ่มควบคุมหรือไม่

5) **การวัดความรู้โดยใช้ผลต่างระหว่างก่อนและหลังการสอนของทั้ง 2 กลุ่มมาเปรียบเทียบกัน** จะใช้ตอบคำถามว่า กลุ่มทดลองและกลุ่มควบคุมมีขนาดของความรู้เพิ่มขึ้นต่างกันหรือไม่ ซึ่งความรู้ที่เพิ่มขึ้นมากกว่ามิได้หมายความว่ารู้มากกว่า เช่น

	ค่าเฉลี่ยก่อนสอน	ค่าเฉลี่ยหลังสอน	ค่าเฉลี่ยของผลต่างระหว่างก่อนและหลังสอน
กลุ่มทดลอง	25	40	15
กลุ่มควบคุม	15	35	20

ถ้าใช้ค่าเฉลี่ยของผลต่างระหว่างก่อนและหลังสอนเปรียบเทียบกัน จะพบว่ากลุ่มทดลองเพิ่มขึ้นน้อยกว่ากลุ่มควบคุม แต่ถ้าพิจารณาเฉพาะค่าเฉลี่ยหลังการสอน พบว่าค่าเฉลี่ยของกลุ่มทดลองสูงกว่ากลุ่มควบคุม ซึ่งสรุปได้ว่าเมื่อสิ้นสุดการทดลอง กลุ่มทดลองจะมีค่าเฉลี่ยความรู้มากกว่ากลุ่มควบคุม

โดยปกติในการทดลองเพื่อประเมินผลการสอนมีวัตถุประสงค์หลักเพื่อตอบคำถามว่า เมื่อสิ้นสุดการทดลองแล้วกลุ่มทดลองจะมีความรู้มากกว่ากลุ่มควบคุมหรือไม่ ดังนั้นตัววัดผลหลักจึงควรเป็นการเปรียบเทียบความรู้หลังการทดลอง โดยใช้การทดสอบด้วยสถิติ t กรณีที่ 2 กลุ่มเป็น

Using Statistics in Reseach 1_1 .indd 150 dummy 4 30/6/2557 11:29:38

อิสระต่อกัน หรือสถิติ paired t กรณีที่ 2 กลุ่มไม่เป็นอิสระต่อกัน เช่น การเลือกกลุ่มควบคุมโดยการจับคู่ให้มีเพศและกลุ่มอายุเหมือนกับกลุ่มทดลอง และมีคะแนนความรู้ก่อนทดลองต่างกันไม่เกิน 5 คะแนน

ในกรณีที่ความรู้ก่อนสอน 2 กลุ่มแตกต่างกัน ถ้าผู้วิจัยคาดว่าจะมีผลกระทบต่อความรู้หลังการทดลอง นักวิจัยควรใช้วิธีสุ่มแบบชั้นภูมิ (stratified random sampling) เพื่อให้ผู้เข้ารับการทดลองทั้ง 2 กลุ่ม มีระดับความรู้ก่อนทดลองใกล้เคียงกัน หรืออาจแก้ไขโดยปรับค่าผลกระทบด้วยการวิเคราะห์แบบ covariate adjustment

โดยปกติในการออกแบบหลักสูตรการสอนจะครอบคลุมเริ่มจากผู้ที่ไม่รู้เรื่องเลย เมื่อได้รับเนื้อหาตามหลักสูตรแล้วจะมีความรู้เพียงพอในเรื่องนั้นๆ ซึ่งถ้าประเมินผลตามเกณฑ์เพียงพอหรือไม่เพียงพอ ความรู้ก่อนการทดลองจะไม่มีผลกระทบต่อความรู้หลังการทดลอง

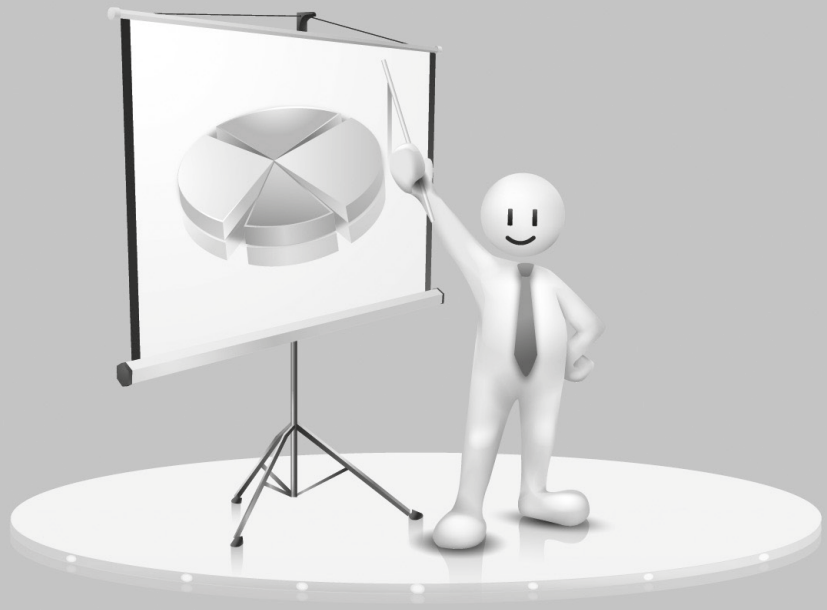
ความไม่เหมาะสมของการวิเคราะห์ที่พบคือนักวิจัยทำการทดสอบสมมติฐานเปรียบเทียบ ก่อนทดลองระหว่างกลุ่ม ก่อนทดลอง–หลังทดลองของแต่ละกลุ่ม และหลังการทดลองระหว่างกลุ่ม การที่มีการทดสอบหลายครั้ง (multiple testing) ทำให้การทดสอบสมมติฐานมีความผิดพลาด α เพิ่มขึ้น และคำตอบที่ได้ส่วนใหญ่ไม่เกี่ยวข้องกับการเปรียบเทียบระหว่างวิธีการสอน

สรุป

การวิเคราะห์และคำนวณขนาดตัวอย่างการทดลองเพื่อเปรียบเทียบการสอน 2 กลุ่ม จะต้องใช้ความรู้หลังการทดลองของทั้ง 2 กลุ่มเป็นตัวประเมินความสำเร็จหลักของการทดลอง ในกรณีที่นักวิจัยคาดว่าความรู้ก่อนสอนหรือปัจจัยอื่นจะมีผลกระทบต่อผลการสอน นักวิจัยสามารถปรับค่าผลกระทบด้วยการวิเคราะห์แบบ covariate adjustment

Using Statistics in Reseach 1_1 .indd 151 dummy 4 27/8/2557 11:54:15

สถิติในการสรุปแปลผล และนำเสนอผลวิจัย



37

ความแตกต่างอย่างมีนัยสำคัญ สำคัญจริงหรือ

การวิเคราะห์ข้อมูลงานวิจัยเชิงปริมาณมีการทดสอบสมมติฐานทางสถิติเพื่อเปรียบเทียบค่าเฉลี่ยหรือค่าสัดส่วนระหว่างกลุ่ม เมื่อผลการทดสอบปฏิเสธสมมติฐานจะสรุปว่า ค่าเฉลี่ยหรือค่าสัดส่วนที่ทดสอบมีความแตกต่างกันอย่างมีนัยสำคัญ (significant difference) และนำผลสรุปที่พบว่าแตกต่างนี้ไปวิจารณ์ผลพร้อมกับให้ข้อเสนอแนะโดยไม่นำขนาดความแตกต่างมาร่วมพิจารณา

ขนาดความแตกต่างที่พบในการทดสอบสมมติฐาน เช่น ผลต่างของค่าเฉลี่ยระหว่างกลุ่ม ($\overline{x}_t - \overline{x}_c$) หรือผลต่างของค่าสัดส่วนระหว่างกลุ่ม ($p_t - p_c$) เมื่อผลการทดสอบมีความแตกต่างกันอย่างมีนัยสำคัญ หมายความว่าผลสรุปที่ได้จากตัวอย่างแสดงว่ามีความแตกต่างกันจริงในประชากร เมื่อสรุปได้ว่าแตกต่างกันจริงแล้ว ขั้นต่อไปควรจะพิจารณาขนาดของความแตกต่าง เพราะขนาดความแตกต่างเป็นตัวสะท้อนว่า 2 วิธีให้ผลต่างกันมากน้อยเท่าไร เช่น ผลการรักษา 2 วิธีถ้าแตกต่างกันร้อยละ 5 อาจไม่มีผลต่างมากนักในการนำไปใช้รักษาผู้ป่วย แต่ถ้าแตกต่างกันร้อยละ 20 จะมีความสำคัญมากต่อผลการรักษา ความแตกต่างที่มีความสำคัญนี้ถ้าเป็นการรักษาโรคจะเรียกว่าความแตกต่างมีความสำคัญหรือมีนัยสำคัญทางคลินิก (clinical importance หรือ clinical significance)

การที่นักวิจัยนำเฉพาะข้อสรุปว่าแตกต่างอย่างมีนัยสำคัญทางสถิติ (statistical significance) ไปสรุปผลวิจัยอย่างเดียวโดยไม่นำขนาดความแตกต่างที่มีความสำคัญมาร่วมพิจารณาด้วย อาจทำให้ข้อเสนอแนะและการนำผลงานวิจัยไปใช้ไม่เหมาะสม เนื่องจากเหตุผลสำคัญ 2 ประการ คือ ประการแรกความแตกต่างอย่างมีนัยสำคัญเกิดขึ้นได้จากการมีจำนวนตัวอย่างขนาดใหญ่พอเพียง ดังนั้นการวิจัยทุกเรื่องถ้ามีจำนวนตัวอย่างขนาดใหญ่พอก็จะทำให้มีความแตกต่างทางสถิติได้ และประการที่ 2 ความแตกต่างอย่างมีนัยสำคัญทางสถิติบอกได้ว่ามีความแตกต่าง แต่ไม่สามารถบอกว่าแตกต่างมากน้อยเท่าใด การพิจารณาประโยชน์จะขึ้นอยู่กับขนาดความแตกต่าง ความแตกต่างขนาดเล็กไม่มีผลต่อการเปลี่ยนวิธีรักษาผู้ป่วยก็จะไม่มีความสำคัญทางคลินิก ในทำนองเดียวกันถ้าเป็นเรื่องการให้สุขศึกษาถ้าความรู้ที่เพิ่มขึ้นมีความแตกต่างอย่างมีนัยสำคัญ แต่ผู้ได้รับการสอนยังมีความรู้ไม่เพียงพอเหมือนเดิม ความแตกต่างที่พบจะไม่มีประโยชน์ต่อการใช้งานหรือไม่มีความสำคัญ

การทดสอบสมมติฐานทางสถิติเป็นการนำเอาความแตกต่างของค่าเฉลี่ย ($\overline{x}_t - \overline{x}_c$) หรือค่าสัดส่วน ($p_t - p_c$) ไปเทียบกับความคลาดเคลื่อนมาตรฐาน (standard error) ถ้าพบว่าความแตกต่างมีขนาดใหญ่กว่าที่ตั้งค่าระดับนัยสำคัญไว้ จะสรุปว่ามีความแตกต่างกันอย่างมีนัยสำคัญ ขนาดของความคลาดเคลื่อนมาตรฐานจะขึ้นอยู่กับขนาดของตัวอย่าง ดังนั้นในการศึกษาที่ใช้ตัวอย่างขนาดเล็กผลการทดสอบสมมติฐานพบความแตกต่างอย่างไม่มีนัยสำคัญ ถ้าเพิ่มขนาดตัวอย่างให้มากๆจะสามารถพบความแตกต่างอย่างมีนัยสำคัญได้

ตัวอย่างที่ 1 ในการวิจัยหารูปแบบการรณรงค์ให้สวมใส่หมวกนิรภัยในการขับขี่จักรยานยนต์ โดยเปรียบเทียบวิธีการรณรงค์รูปแบบใหม่กับวิธีการที่ใช้อยู่เดิม พบว่าวิธีการเดิมมีอัตรา

Using Statistics in Reseach 1_1 .indd 154 dummy 4 30/6/2557 11:29:39

การสวมใส่ฯร้อยละ 28.5 (p_1 = 0.285) วิธีการใหม่มีอัตราการสวมใส่ฯร้อยละ 30 (p_2 = 0.30) ในการทดสอบสมมติฐานเปรียบเทียบความแตกต่างของสัดส่วนการสวมใส่ฯ มีผลการคำนวณดังแสดงในตาราง

ขนาดตัวอย่าง	$(p_1 - p_2)$	$SE = \sqrt{\frac{p(1-p)}{n_1} + \frac{p(1-p)}{n_2}}$	$Z = \frac{p_1 - p_2}{SE}$	P value
1,000	0.015	0.0203	0.74	0.459
3,000	0.015	0.0117	1.28	0.201
5,000	0.015	0.0091	1.65	0.099
7,500	0.015	0.0074	2.02	0.043

จากตารางจะพบว่าเมื่อเพิ่มขนาดตัวอย่างขึ้นเรื่อยๆ โดยให้ค่า p_1 และ p_2 คงที่ ค่าความคลาดเคลื่อนมาตรฐานจะเล็กลง ทำให้ผลการทดสอบสมมติฐานอัตราการสวมใส่หมวกนิรภัยจะแตกต่างกันอย่างมีนัยสำคัญ ($P < 0.05$) เมื่อขนาดตัวอย่างเท่ากับ 7,500 คน/กลุ่ม

จากผลการศึกษาดังกล่าว ถ้าผลการทดสอบสมมติฐานทางสถิติพบว่ามีความแตกต่างอย่างมีนัยสำคัญ ผู้วิจัยเสนอแนะให้มีการเปลี่ยนวิธีการรณรงค์แบบใหม่ แต่ถ้าพิจารณาขนาดความแตกต่างของทั้ง 2 วิธี พบว่าแตกต่างกันเพียงร้อยละ 1.5 ซึ่งต่างกันค่อนข้างน้อย และอัตราการสวมใส่ฯ จากการรณรงค์วิธีใหม่มีเพียงร้อยละ 30 ข้อเสนอแนะคือ ไม่ควรจะแนะนำให้เปลี่ยนวิธีการรณรงค์รูปแบบใหม่ แต่ควรแนะนำให้หาวิธีใหม่ที่ได้ผลมากกว่านี้ การนำขนาดความแตกต่างที่มีความสำคัญต่อการใช้งานมาร่วมพิจารณาจะช่วยให้ข้อเสนอแนะถูกต้องสำหรับการแก้ปัญหา

Using Statistics in Reseach 1_1 .indd 155 dummy 4 8/9/2557 11:47:48

ตัวอย่างที่ 2 ในการวิจัยทดลองประสิทธิผลของหน้ากากป้องกันไอตะกั่ว พบว่ากลุ่มที่ไม่ได้ใช้หน้ากากป้องกันไอตะกั่วมีปริมาณตะกั่วในเลือด 63 mcg/dl ส่วนกลุ่มที่สวมหน้ากากมีปริมาณตะกั่วในเลือด 51 mcg/dl ผลการทดสอบสมมติฐานพบว่ามีความแตกต่างกันอย่างมีนัยสำคัญ ซึ่งแสดงว่าหน้ากากป้องกันไอตะกั่วสามารถลดปริมาณสารตะกั่วที่เข้าสู่ร่างกายได้ แต่ถ้าดูจากมาตรฐานความปลอดภัยในร่างกายมนุษย์กำหนดว่าจะมีตะกั่วได้ไม่เกิน 40 mcg/dl ดังนั้นเมื่อเอาข้อมูลนี้มาพิจารณาร่วมกับผลการทดลองจะเห็นได้ว่า หน้ากากดังกล่าวถึงแม้จะลดปริมาณตะกั่วลงได้อย่างมีนัยสำคัญทางสถิติ แต่ปริมาณที่ลดได้ยังไม่มากพอที่จะทำให้เกิดความปลอดภัยกับคนงาน เรียกว่าผลต่างที่ได้ไม่มีความสำคัญทางคลินิก จึงไม่ควรแนะนำให้ใช้หน้ากากดังกล่าวป้องกันไอตะกั่ว

สรุป

เมื่อผลการทดสอบสมมติฐานในการวิจัยพบว่ามีความแตกต่างอย่างมีนัยสำคัญ ผู้วิจัยควรพิจารณาขนาดความแตกต่างที่พบว่ามีความสำคัญทางคลินิกประกอบกันด้วย จึงจะช่วยให้ได้ข้อสรุปผลและข้อเสนอแนะที่ถูกต้อง

มีการนำหลักการพิจารณาความสำคัญทางคลินิกไปใช้ในงานวิจัยด้านต่างๆ ก็จะพิจารณาความสำคัญสำหรับศาสตร์นั้นๆ เช่น public health importance, nursing importance และ scientific importance เป็นต้น

Using Statistics in Reseach 1_1 .indd 156 dummy 4 27/8/2557 11:54:16

38

การแปลความหมาย P value ที่ไม่ถูกต้อง

ในการทดสอบสมมติฐานเพื่อเปรียบเทียบผลของสิ่งทดลอง (treatment) ในงานวิจัยทางด้านการแพทย์และสาธารณสุขส่วนใหญ่นิยมสรุปผลโดยใช้ P value เช่น ในการศึกษาว่าผู้ป่วยที่ได้รับการอธิบายขั้นตอนต่าง ๆของการผ่าตัดจะมีความวิตกกังวลน้อยกว่ากลุ่มควบคุมที่ไม่ได้รับการอธิบายหรือไม่ จากรายงานผลวิจัยที่ทำในต่างประเทศพบว่าการได้รับความรู้ช่วยให้ผู้ป่วยลดความวิตกกังวลลงได้อย่างมีนัยสำคัญ มีค่า P = 0.04 และรายงานผลการวิจัยที่ทำในประเทศไทยในเรื่องเดียวกันพบว่า การได้รับความรู้ช่วยให้ผู้ป่วยลดความวิตกกังวลลงได้อย่างมีนัยสำคัญ มีค่า P = 0.001 นักวิจัยสรุปวิจารณ์ผลว่าการศึกษาในคนไทยผู้ป่วยลดความวิตกกังวลมากกว่า (โดยดูจากค่า P value ของการศึกษาของไทยมีค่าน้อยกว่าที่พบในต่างประเทศ) ข้อสรุปดังกล่าวเป็นการแปลความหมายของ P value ที่ไม่ถูกต้อง การแปลความหมายของ P value ที่ไม่ถูกต้องที่พบบ่อยมีดังนี้

1) **มีการนำค่า P value ไปเป็นค่าความน่าจะเป็นของการเกิดความผิดพลาดชนิดที่ 1 (α error)** เช่น ถ้าพบว่า P = 0.03 จะสรุปว่าผลการทดสอบปฏิเสธสมมติฐาน (H_0) มีโอกาสผิดพลาดไม่เกิน 3% หรือมีโอกาสสรุปถูก 97% การแปลผล

ดังกล่าวไม่ถูกต้อง เพราะค่า P value เป็นความน่าจะเป็นที่จะสรุปผิด เมื่อสมมติฐานที่ตั้งไว้ (ว่าเท่ากัน) ถูกต้อง และค่า P value จะมีขนาดเล็กลงเมื่อขนาดตัวอย่างเพิ่มขึ้น ดังนั้นเมื่อพบว่า P value มีค่าน้อยๆ จะช่วยให้มั่นใจว่าในการตัดสินใจปฏิเสธสมมติฐานจะมีโอกาสเกิดความผิดพลาดน้อยเท่านั้น

2) **การที่ได้ค่า P value เล็กๆ เช่น 0.001 แล้วสรุปว่ามีโอกาสแตกต่างกันมากกว่า P value ขนาดใหญ่** เช่น ในตัวอย่างเรื่องการเปรียบเทียบความวิตกกังวลที่ได้กล่าวไว้ในตอนต้นเป็นข้อสรุปที่ไม่ถูกต้อง เพราะข้อสรุปที่ได้จากการตัดสินใจปฏิเสธสมมติฐานโดยพิจารณาจากค่า P value บอกได้เพียงแต่ว่า ถ้าค่า α ที่ตั้งไว้มีค่าแตกต่างจากค่า P value ยิ่งมากเท่าไรก็ยิ่งจะช่วยให้มั่นใจว่าการตัดสินมีโอกาสผิดพลาดยิ่งน้อยเท่านั้น

ถ้าจะดูว่าผู้ป่วยคนไทยที่ได้รับการอธิบายขั้นตอนต่างๆ ของการผ่าตัดสามารถลดความวิตกกังวลมากกว่าผลวิจัยที่ศึกษาในต่างประเทศหรือไม่ จะต้องดูที่ค่าช่วงเชื่อมั่น OR, RR หรือ absolute difference ของความวิตกกังวล

3) **เมื่อพบว่าค่า P value มีค่ามากกว่า 0.05 แล้วสรุปว่าค่าเฉลี่ยหรือค่าสัดส่วนที่เปรียบเทียบเท่ากัน** ตัวอย่างเช่น ในการเปรียบเทียบค่าเฉลี่ยความวิตกกังวลในขณะรอผ่าตัดระหว่างหญิง ($\overline{x}$ = 62.8) กับชาย ($\overline{x}$ = 45.1) พบว่าค่า P = 0.062 ผู้วิจัยสรุปว่าผลการทดสอบสมมติฐานแตกต่างกันอย่างไม่มีนัยสำคัญ จึงสรุปว่าค่าเฉลี่ยความวิตกกังวลในขณะรอผ่าตัดของชายหญิงเท่ากัน ซึ่งเป็นการสรุปที่ไม่ถูกต้อง เพราะค่า P value จะเปลี่ยนไปตามขนาดตัวอย่าง เมื่อขนาดตัวอย่างใหญ่ขึ้น P value จะมีขนาดเล็กลงจนพบความแตกต่างได้ จึงควรสรุปว่าผลการศึกษายังไม่สามารถระบุความแตกต่าง

Using Statistics in Reseach 1_1 .indd 158 dummy 4 30/6/2557 11:29:40

สรุป

ค่า P value มักจะมีการแปลความหมายผิดได้ง่าย ความเข้าใจวิธีการทางสถิติที่ถูกต้องจะช่วยให้การแปลความหมายค่า P value ได้อย่างถูกต้อง การสรุปผลการเปรียบเทียบโดยใช้ช่วงเชื่อมั่นเป็นอีกวิธีหนึ่งที่จะช่วยลดความผิดพลาดในการแปลความหมายค่า P value และยังช่วยให้ระบุขนาดของความแตกต่างได้ด้วย

Using Statistics in Reseach 1_1 .indd 159 dummy 4 30/6/2557 11:29:40

39

การนำเสนอผลต่างของการสอนด้วยขนาดอิทธิพล

ในการทดลองเพื่อเปรียบเทียบผลการสอน (ให้สุขศึกษา) 2 วิธี ว่าได้ผลต่างกันหรือไม่ งานวิจัยส่วนใหญ่นำเสนอแต่ค่า P value อย่างเดียวซึ่งสรุปได้ว่าความรู้ของ 2 กลุ่มต่างกันหรือไม่เท่านั้น การพิจารณานำผลงานวิจัยไปใช้งาน ขนาดความแตกต่างเป็นข้อมูลสำคัญที่ใช้พิจารณาตัดสินใจว่าถ้านำไปใช้แล้วจะช่วยทำให้ผู้รับการสอนมีความรู้เพิ่มขึ้นมากน้อยเพียงใด ขนาดอิทธิพล (effect size) เป็นสถิติใช้บอกขนาดความแตกต่างเมื่อผลการทดสอบสมมติฐานแตกต่างกันอย่างมีนัยสำคัญ

ขนาดอิทธิพลคืออะไร ขนาดอิทธิพลคือขนาดความแตกต่างของค่าเฉลี่ยระหว่างกลุ่มทดลองกับกลุ่มควบคุมหารด้วยส่วนเบี่ยงเบนมาตรฐาน $(\overline{x}_t - \overline{x}_c)/s$ ค่าขนาดอิทธิพลที่คำนวณได้จะไม่มีหน่วยเช่นเดียวกันกับคะแนนมาตรฐาน (standard score หรือ Z score) ทำให้สามารถใช้ในการระบุขนาดความแตกต่างได้ดีกว่าความแตกต่างระหว่างค่าเฉลี่ยอย่างเดียว เพราะขนาดความแตกต่างที่เท่ากัน ถ้ามาจากการศึกษาที่อาสาสมัคร (ผู้เข้ารับการทดลอง) เรียนรู้ได้ดีแตกต่างกันมาก

ค่าส่วนเบี่ยงเบนมาตรฐานจะใหญ่แต่ค่าขนาดอิทธิพลที่คำนวณได้จะมีขนาดเล็ก ส่วนการศึกษาที่อาสาสมัครเรียนรู้ได้ดีพอๆกันทุกคนค่าส่วนเบี่ยงเบนมาตรฐานจะเล็กทำให้ค่าขนาดอิทธิพลมีขนาดใหญ่ จึงสรุปได้ว่าวิธีการสอนที่มีขนาดความแตกต่างระหว่างค่าเฉลี่ยเท่ากัน แต่ถ้ากลุ่มคนที่เรียนมีความรู้พอๆกัน ย่อมจะดีกว่าวิธีสอนที่บางคนในกลุ่มเรียนรู้ได้ดีบางคนเรียนไม่รู้เรื่อง

ค่าส่วนเบี่ยงเบนมาตรฐานที่ใช้คำนวณค่าขนาดอิทธิพลโดยปกติจะใช้ค่าส่วนเบี่ยงเบนมาตรฐานของกลุ่มควบคุม เพราะกลุ่มควบคุมเป็นความรู้ในสภาวะปกติที่ไม่ได้ถูกแทรกแซง ค่าส่วนเบี่ยงเบนมาตรฐานของกลุ่มควบคุมจึงมีลักษณะเหมือนของประชากร ในการศึกษาถ้าใช้ขนาดตัวอย่างที่ใหญ่พอจะได้ค่าส่วนเบี่ยงเบนมาตรฐานที่เป็นตัวแทนที่ดีของประชากร ในกรณีที่มีการทำการวัดผลก่อนทดลองของทั้ง 2 กลุ่ม และนักวิจัยต้องการจะลดความคลาดเคลื่อนของการสุ่มตัวอย่าง ให้ใช้ค่าส่วนเบี่ยงเบนมาตรฐานจากความแปรปรวนร่วม (pool variance) ของคะแนนก่อนการทดลองในการคำนวณขนาดอิทธิพล

การแปลความหมายของขนาดอิทธิพล Jacob Cohen* ได้นำเสนอเป็นเกณฑ์กว้างๆในการแปลผลขนาดอิทธิพลดังนี้ 0.2 = น้อย, 0.5 = ปานกลาง และ 0.8 = มาก และบอกว่าผู้ใช้เกณฑ์นี้ต้องใช้ด้วยความระมัดระวังเพราะเกณฑ์กว้างๆนี้อาจแตกต่างกันไปตามสาขาวิชา นอกจากนี้ยังมีวิธีการแปลความหมายอีกหลายวิธีที่ใช้กันในงานวิจัยด้านการศึกษา

ขนาดอิทธิพลนอกจากจะบอกขนาดของความแตกต่างแล้ว ยังสามารถใช้ในการเปรียบเทียบผลการศึกษาเรื่องเดียวกันแต่วัดด้วยเครื่องมือ

---

* Jacob Cohen, **Statistical Power Analysis for the Behavioral Sciences.** 2nd ed. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1988.

Using Statistics in Reseach 1_1 .indd 161 dummy 4 30/6/2557 11:29:40

ต่างกัน เช่น ในการประเมินความรู้เรื่องการป้องกันโรคแทรกซ้อนของผู้ป่วยเบาหวาน งานวิจัยแรกวัดด้วยแบบวัดความรู้ 20 ข้อ ส่วนอีกหนึ่งงานวิจัยวัดด้วยแบบวัดความรู้ 30 ข้อ ผลต่างของค่าเฉลี่ยของ 2 งานวิจัยไม่สามารถเปรียบเทียบกันได้โดยตรงเพราะคะแนนเต็มไม่เท่ากัน แต่ขนาดอิทธิพลที่คำนวณจากงานวิจัยทั้งสองจะเป็นหน่วยเดียวกันจึงนำมาเปรียบเทียบกันได้ว่าการศึกษาใดอาสาสมัครได้ความรู้มากกว่าเมื่อเทียบกับกลุ่มควบคุม ถึงแม้เครื่องมือวัดเป็นคนละวิธี เช่น งานวิจัยที่ 1 ใช้คำตอบ 4 ตัวเลือก งานวิจัยที่ 2 ใช้คำตอบ 5 ตัวเลือกก็สามารถนำขนาดอิทธิพลมาเปรียบเทียบกันได้เพราะผลต่างถูกเปลี่ยนให้เป็นความแตกต่างต่อค่าส่วนเบี่ยงเบนมาตรฐาน ทำให้มีหน่วยวัดเดียวกัน

ความถูกต้องของการเปรียบเทียบด้วยขนาดอิทธิพล ตัวอย่างต้องสุ่มมาจากประชากรที่มีการแจกแจงปกติ ขนาดตัวอย่างที่ใช้ต้องใหญ่พอ และแบบงานวิจัยต้องเหมือนกัน

สรุป

ในการนำเสนอความแตกต่างของผลการทดสอบสมมติฐาน ควรนำเสนอขนาดอิทธิพลเพื่อแสดงขนาดของความแตกต่าง ในการพิจารณาผลที่คาดว่าจะได้รับถ้ามีการนำผลวิจัยไปใช้ควรพิจารณาจาก 95% ช่วงเชื่อมั่นของความแตกต่างของค่าเฉลี่ยระหว่างกลุ่มซึ่งเคยอธิบายไว้ในเรื่องที่ผ่านมา

ขนาดอิทธิพลของความแตกต่างไม่มีหน่วยการวัดจึงเหมาะที่จะใช้เปรียบเทียบขนาดความแตกต่างระหว่างงานวิจัย

Using Statistics in Reseach 1_1 .indd 162 dummy 4 30/6/2557 11:29:40

40

ประโยชน์ของการนำเสนอช่วงเชื่อมั่นของผลการเปรียบเทียบ

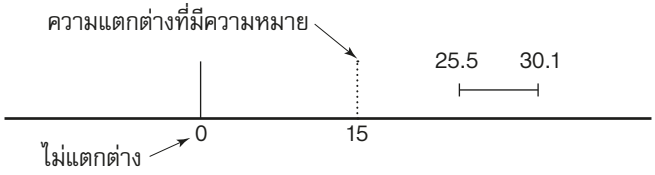
การเปรียบเทียบค่าเฉลี่ยหรือค่าสัดส่วนเพื่อสรุปความแตกต่างของประชากร 2 กลุ่ม นักวิจัยส่วนใหญ่คุ้นเคยกับการทดสอบสมมติฐานและนำเสนอผลการเปรียบเทียบด้วย P value การสรุปผลทำโดยนำค่า P value ไปเปรียบเทียบกับค่าระดับนัยสำคัญ (α) ถ้าค่า P value น้อยกว่าจะสรุปว่ามีความแตกต่างกันจริง ค่า P value ยิ่งเล็กลงเท่าไร การสรุปว่าแตกต่างจะมีโอกาสผิดพลาดยิ่งน้อยลงเท่านั้น แต่ P value ที่เล็กๆมิได้หมายความว่าผลการเปรียบเทียบมีความแตกต่างกันมาก

ในการเปรียบเทียบผลการทดลองระหว่าง 2 กลุ่มด้วย 95% ช่วงเชื่อมั่นของผลต่างของค่าเฉลี่ยหรือค่าสัดส่วนที่สร้างขึ้นจะให้ข้อมูลผลการเปรียบเทียบได้ 3 ส่วน ดังนี้

1) **ใช้พิจารณาเปรียบเทียบความแตกต่างของค่าเฉลี่ยหรือค่าสัดส่วนระหว่างกลุ่มได้** เช่นเดียวกันกับที่ได้จากค่า P value ของการทดสอบสมมติฐาน เช่น ในการเปรียบเทียบค่าเฉลี่ยคะแนนความรู้ระหว่างการสอนวิธีใหม่กับการสอนวิธีเดิมให้แก่ญาติผู้ป่วย มีค่า 95% ช่วงเชื่อมั่นของผลต่างของค่าเฉลี่ยเท่ากับ 25.5 ถึง 30.1 คะแนน ในการเปรียบเทียบถ้าค่าเฉลี่ยของ 2 กลุ่มเท่ากันผลต่างจะเท่ากับศูนย์ ดังนั้นการพิจารณาว่าการสอน

2 วิธีนี้ให้ผลการสอนแตกต่างกันอย่างมีนัยสำคัญหรือไม่ จะพิจารณาว่าช่วงเชื่อมั่นของผลต่างคร่อมศูนย์ (มีค่า 0 อยู่ในช่วงเชื่อมั่น) หรือไม่ ถ้าช่วงเชื่อมั่นไม่คร่อมศูนย์จะสรุปว่าคะแนนความรู้ทั้ง 2 กลุ่มแตกต่างกันอย่างมีนัยสำคัญ ดังนั้น 95% ช่วงเชื่อมั่นของผลต่างของค่าเฉลี่ยเท่ากับ 25.5 ถึง 30.1 ไม่คร่อมศูนย์ จึงสรุปว่าการสอนวิธีใหม่มีความรู้แตกต่างจากการสอนวิธีเดิมอย่างมีนัยสำคัญ

2) **บอกทิศทางและขนาดความแตกต่าง** จากภาพ 1 ค่า 95% ช่วงเชื่อมั่นของผลต่างการสอน 2 วิธี (วิธีใหม่ – วิธีเดิม) อยู่ระหว่าง 25.5 ถึง 30.1 แสดงว่าการสอนวิธีใหม่ผู้เรียนได้คะแนนสูงกว่าเพราะแตกต่างไปด้านบวก จึงสรุปว่าการสอนวิธีใหม่ช่วยให้ญาติผู้ป่วยมีความรู้มากขึ้นกว่าการสอนด้วยวิธีเดิมอย่างน้อย 25.5 คะแนน หรืออาจรู้มากกว่ากันถึง 30.1 คะแนน



ภาพ 1 ช่วงเชื่อมั่นของผลต่างของค่าเฉลี่ยคะแนนความรู้

ในการตัดสินใจนำผลการศึกษาไปใช้งาน นักวิจัยควรกำหนดขนาดความแตกต่างที่มีความสำคัญหรือมีประโยชน์ในการนำไปใช้ เช่น ถ้าความรู้จากการสอน 2 วิธีแตกต่างกันเกิน 15 คะแนน จึงจะทำให้เกิดการเปลี่ยนแปลงในการปฏิบัติ จากภาพ 1 ค่าต่ำสุดของช่วงเชื่อมั่นมีค่ามากกว่า 15 คะแนน แสดงว่าความแตกต่างที่พบจากการศึกษามีขนาดมากพอที่จะตัดสินใจนำวิธีการสอนแบบใหม่ไปใช้แทนการสอนแบบเดิม

Using Statistics in Reseach 1_1 .indd 164 dummy 4 30/6/2557 11:29:40

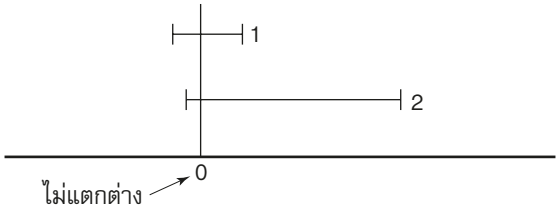
3) **ความกระชับในการสรุปผล** โดยพิจารณาจากความกว้างของช่วงเชื่อมั่นซึ่งขึ้นอยู่กับขนาดตัวอย่าง เช่น ช่วงเชื่อมั่นของผลต่างของสัดส่วนผู้สวมหมวกนิรภัยก่อนและหลังการอบรมมีค่าอยู่ระหว่าง 0.04 ถึง 0.45 แสดงว่าหลังการอบรมมีผู้สวมหมวกนิรภัยเพิ่มขึ้นร้อยละ 4 ถึงร้อยละ 45 ทำให้การให้ข้อเสนอแนะทำได้ยาก เพราะถ้าเพิ่มร้อยละ 4 ควรแนะนำให้หาวิธีสอนใหม่ แต่ถ้ามีผู้ใส่เพิ่มขึ้นถึงร้อยละ 45 จะแนะนำให้นำวิธีการสอนดังกล่าวไปใช้ให้ครอบคลุมทั้งพื้นที่ ดังนั้นความกระชับของช่วงเชื่อมั่นจึงเป็นข้อมูลอีกส่วนหนึ่งที่ช่วยในการตัดสินใจนำผลวิจัยไปใช้งาน

นอกจากนี้ ความกระชับยังช่วยในการพิจารณาความเพียงพอของขนาดตัวอย่างเมื่อพบว่าผลการทดสอบแตกต่างอย่างไม่มีนัยสำคัญ ในกรณีที่ผลการทดสอบสมมติฐานพบว่ามีความแตกต่างกันอย่างไม่มีนัยสำคัญ ข้อสรุปเป็นไปได้ 2 ทาง คือ ไม่แตกต่างกันจริง หรือขนาดตัวอย่างไม่พอที่จะระบุความแตกต่าง โดยพิจารณาจากอำนาจการทดสอบ ถ้าได้ค่าอำนาจการทดสอบสูงเกิน 65% จะสรุปว่าค่าเฉลี่ยหรือค่าสัดส่วนของทั้ง 2 กลุ่มไม่แตกต่างกันจริง ถ้าได้ค่าอำนาจการทดสอบต่ำจะสรุปว่าการศึกษามีขนาดตัวอย่างไม่พอที่จะระบุความต่าง

ในการพิจารณาจากค่าช่วงเชื่อมั่นของผลต่าง ถ้าพบว่าช่วงเชื่อมั่นคร่อมศูนย์จะสรุปว่าแตกต่างกันอย่างไม่มีนัยสำคัญ และจะพิจารณาความกระชับ (ความกว้าง) ของช่วงเชื่อมั่นเพื่อสรุปว่าไม่แตกต่างกันจริง หรือขนาดตัวอย่างไม่พอที่จะระบุความแตกต่าง จากภาพ 2 ช่วงเชื่อมั่นของทั้ง 2 การศึกษาคร่อมศูนย์ แสดงว่าแตกต่างกันอย่างไม่มีนัยสำคัญ ถ้าพิจารณาความกระชับของช่วงเชื่อมั่นของการศึกษาที่ 1 พบว่าช่วงเชื่อมั่น

Using Statistics in Reseach 1_1 .indd 165 27/8/2557 11:54:16
dummy 4

แคบ จะสรุปว่าทั้ง 2 กลุ่มไม่แตกต่างกันจริง สำหรับการศึกษาที่ 2 ค่าช่วงเชื่อมั่นกว้างและค่าศูนย์อยู่ที่ปลายด้านค่าน้อย ถ้าทำการศึกษาซ้ำโดยเพิ่มขนาดตัวอย่างอาจมีโอกาสพบความแตกต่าง



ภาพ 2 ช่วงเชื่อมั่นของผลต่างของสัดส่วน

สรุป

ปัจจุบันการรายงานผลงานวิจัยที่ตีพิมพ์ในวารสารในต่างประเทศมีการกำหนดแนวปฏิบัติให้นักวิจัยนำเสนอผลการเปรียบเทียบระหว่าง 2 กลุ่มด้วยช่วงเชื่อมั่น เพราะนอกจากจะสามารถสรุปผลได้เช่นเดียวกันกับการทดสอบสมมติฐานแล้ว ยังให้ข้อมูลที่เป็นประโยชน์ต่อการตัดสินใจของผู้อ่านที่จะนำผลงานวิจัยดังกล่าวไปใช้ประโยชน์

Using Statistics in Reseach 1_1 .indd 166 dummy 4 27/8/2557 11:54:16

41

ทำไมต้องรายงานช่วงเชื่อมั่นในรายงานวิจัย

ในการวิจัยเชิงบรรยายที่ต้องการประมาณค่าลักษณะประชากร เช่น สตรีวัยรุ่นในเขตเมืองภาคกลางมีอัตราการสูบบุหรี่ร้อยละเท่าใด ในการวิเคราะห์ข้อมูล เมื่อนักวิจัยคำนวณอัตราการสูบบุหรี่จากตัวอย่างที่ศึกษาได้แล้ว เช่น ร้อยละ 7 แล้วนำค่าดังกล่าวไปเป็นค่าประมาณอัตราการสูบบุหรี่ของสตรีวัยรุ่นในเขตเมืองภาคกลางทั้งหมดโดยตรง นักวิจัยจะไม่สามารถระบุได้ว่าคำตอบมีโอกาสถูกต้องเท่าไร เพราะถ้ามีการทำซ้ำด้วยวิธีเหมือนกันอาจจะได้คำตอบที่แตกต่างกัน เช่น ร้อยละ 8, 7.2, 6, 7.8 หรือ 9 เป็นต้น

วิธีการทางสถิติที่เหมาะสมกับการวิเคราะห์ดังกล่าวคือ การประมาณค่าด้วยช่วงเชื่อมั่น โดยนิยมกำหนดความเชื่อมั่นไว้ที่ร้อยละ 95 จึงเรียกว่า 95% ช่วงเชื่อมั่น วิธีคำนวณจะใช้ความน่าจะเป็นของการแจกแจงของตัวแปรผล (outcome variable) มาใช้ประมาณค่าช่วงเชื่อมั่น ตัวอย่างในการศึกษาอัตราการสูบบุหรี่ของสตรีวัยรุ่นภาคกลาง ถ้ามีจำนวนตัวอย่างที่ใช้ศึกษาทั้งหมด 1,500 คน พบว่ามีอัตราการสูบบุหรี่ร้อยละ 7 ผลการคำนวณ 95% ช่วงเชื่อมั่นของการสูบบุหรี่จะมีค่าอยู่ระหว่างร้อยละ 5.79–8.44 ซึ่งมีความมั่นใจร้อยละ 95 ว่าค่าช่วงที่

คำนวณได้มีความถูกต้อง (มีค่าจริงของอัตราการสูบบุหรี่ของสตรีวัยรุ่นในเขตเมืองอยู่ในช่วง) ถ้ามีผู้ทำการศึกษาซ้ำด้วยวิธีการเหมือนกัน 100 ครั้ง ช่วงเชื่อมั่นที่คำนวณได้เป็นช่วงที่ถูกต้องไม่ต่ำกว่า 95 ครั้ง

ดังนั้นในงานวิจัยเชิงพรรณนาหรือการสำรวจที่ต้องการนำเสนอผลการศึกษาว่ามีอัตราหรือค่าเฉลี่ยของตัวแปรที่ศึกษาเป็นเท่าไร การประมาณค่าพารามิเตอร์ของประชากรต้องนำเสนอในรูปของช่วงเชื่อมั่นจึงจะเป็นการวิเคราะห์ที่เหมาะสม

สำหรับงานวิจัยที่ต้องการเปรียบเทียบค่าสัดส่วนหรือค่าเฉลี่ยระหว่าง 2 ประชากร เช่น สตรีวัยรุ่นในเขตเมืองภาคเหนือกับภาคกลางมีอัตราการสูบบุหรี่ต่างกันหรือไม่ ผลการศึกษาวัยรุ่นในเขตเมืองภาคกลางจำนวน 1,500 คน พบว่ามีอัตราการสูบบุหรี่ร้อยละ 7 วัยรุ่นในเขตเมืองภาคเหนือจำนวน 1,500 คน มีอัตราการสูบบุหรี่ร้อยละ 12.3 ผลการทดสอบสมมติฐานเปรียบเทียบค่าสัดส่วนได้ค่า $P < 0.001$ ซึ่งสรุปได้ว่ามีความแตกต่างกันอย่างมีนัยสำคัญเท่านั้น ไม่สามารถบอกได้ว่าแตกต่างกันมากน้อยเท่าไร

ถ้าวิเคราะห์คำตอบดังกล่าวด้วยวิธีการประมาณค่าผลต่างของค่าสัดส่วนการสูบบุหรี่ของสตรีวัยรุ่นในเขตเมืองภาคเหนือกับภาคกลาง 95% ช่วงเชื่อมั่นของผลต่างของอัตราการสูบบุหรี่จะอยู่ระหว่าง 3.13–7.47 จากค่าช่วงเชื่อมั่นที่ได้ไม่คร่อมศูนย์ (0) แสดงว่ามีความแตกต่างกันอย่างมีนัยสำคัญ ข้อสรุปที่ได้จะเหมือนกับที่ได้จากการทดสอบสมมติฐาน และยังสามารถบอกขนาดของความแตกต่างได้ว่าอย่างน้อยที่สุดจะแตกต่างกันร้อยละ 3.13 หรือแตกต่างกันมากที่สุดไม่เกินร้อยละ 7.47

จากการที่ช่วงเชื่อมั่นสามารถบอกได้ว่าแตกต่างกันหรือไม่ และยังสามารถบอกขนาดความแตกต่างว่าแตกต่างกันมากน้อยเท่าไร ทำให้ข้อสรุปที่ได้จากการวิเคราะห์ด้วยช่วงเชื่อมั่นให้ข้อมูลที่ครบถ้วนสำหรับการพิจารณานำผลการวิจัยไปใช้งาน ดังนั้นรายงานผลการวิจัยในปัจจุบัน

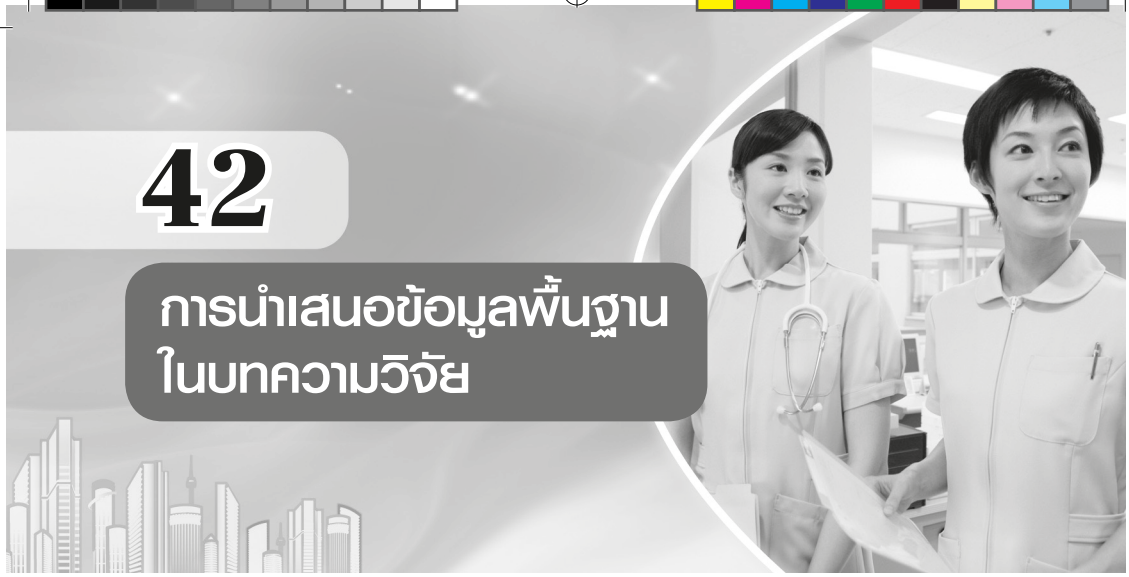
Using Statistics in Reseach 1_1 .indd 168 dummy 4 27/8/2557 11:54:17

จึงนิยมรายงานผลด้วยช่วงเชื่อมั่น โดยเฉพาะรายงานวิจัยที่จะตีพิมพ์ลงในวารสารต่างประเทศจะถูกกำหนดให้ต้องรายงานผลด้วยช่วงเชื่อมั่น ในทางปฏิบัติถึงแม้ข้อสรุปที่ได้จากช่วงเชื่อมั่นเพียงอย่างเดียวก็เพียงพอในการสรุปผลการเปรียบเทียบ แต่เนื่องจากยังมีผู้อ่านผลการวิจัยจำนวนหนึ่งคุ้นเคยกับค่า P value ดังนั้นการเพิ่มค่า P value ลงในรายงานผลพร้อมค่าช่วงเชื่อมั่นก็ช่วยให้ผู้อ่านทุกกลุ่มสามารถอ่านและเข้าใจผลสรุปได้

สรุป

ในรายงานผลการวิจัยไม่ว่าจะเป็นการวิจัยเชิงพรรณนา การวิจัยที่ต้องการเปรียบเทียบ หรือการหาปัจจัยเสี่ยง ที่ต้องการให้ผู้อ่านมีข้อมูลครบถ้วนในการพิจารณานำผลการศึกษาไปใช้ นักวิจัยต้องรายงานผลการศึกษาด้วยช่วงเชื่อมั่น

Using Statistics in Reseach 1_1 .indd 169 dummy 4 30/6/2557 11:29:41



42

การนำเสนอข้อมูลพื้นฐานในบทความวิจัย

ข้อมูลพื้นฐานถูกนำเสนอเป็นส่วนแรกของผลการวิจัย การนำเสนอข้อมูลพื้นฐานดูเหมือนจะง่ายและไม่ค่อยมีคำถามที่เกี่ยวข้องกับปัญหาในส่วนนี้ แต่ยังพบว่ามีการนำเสนอข้อมูลพื้นฐานที่ไม่เหมาะสมและไม่เพียงพอ เช่น รายงานเป็นข้อความสรุปโดยไม่มีตารางข้อมูลประกอบ

ข้อมูลพื้นฐานมีความสำคัญมากสำหรับผู้อ่านบทความวิจัยที่ต้องการนำผลวิจัยไปใช้งาน เพราะจะทำให้ทราบว่าลักษณะของตัวอย่างที่ใช้ศึกษามีลักษณะคล้ายกับประชากรหรือผู้ป่วยที่ต้องการนำผลวิจัยไปใช้หรือไม่ การมีข้อมูลพื้นฐานที่ครบถ้วนจะช่วยในการตัดสินใจใช้ผลงานวิจัยได้อย่างเหมาะสม นอกจากนี้ข้อมูลพื้นฐานยังสามารถใช้พิจารณาคุณภาพของวิธีการศึกษาได้อีกด้วย

ข้อมูลพื้นฐานในบทความวิจัยทางด้านการแพทย์และสาธารณสุขจะประกอบด้วยตัวแปรที่ใช้พรรณนาลักษณะตัวอย่าง 3 ส่วน คือ

1) **ลักษณะทางประชากรและเศรษฐกิจสังคม** เช่น เพศ อายุ อาชีพ ฯลฯ
2) **ลักษณะทางคลินิก** (ในกรณีการศึกษาเกี่ยวกับผู้ป่วย) ระดับสารชีวเคมีในเลือด เช่น คอเลสเตอรอล เฮโมโกลบิน ข้อมูล

ผลตรวจทางร่างกาย เช่น ความดันโลหิต ชีพจร เป็นต้น

3) **ปัจจัยที่มีอิทธิพลต่อตัวแปรผล** ได้แก่ ตัวแปรกวน ปัจจัยเสี่ยง หรือตัวแปรที่ใช้ในการแบ่งกลุ่ม

จะเห็นได้ว่าในแต่ละส่วนมีข้อมูลอยู่หลายตัวแปร การพิจารณาเลือกตัวแปรใดในการนำเสนอควรพิจารณาเลือกตัวแปรที่เกี่ยวข้องกับผลการวิจัยเฉพาะส่วนที่นำเสนอในบทความเท่านั้น

ข้อมูลพื้นฐานสามารถใช้พิจารณาลักษณะตัวอย่างและคุณภาพงานวิจัยได้ดังนี้

1) **ใช้พิจารณาลักษณะของตัวอย่าง**

1.1) **ข้อมูลประชากรและข้อมูลเศรษฐกิจสังคม** จะใช้พิจารณาว่าลักษณะตัวอย่างที่ศึกษาคล้ายกับประชากรที่จะนำผลวิจัยไปประยุกต์ใช้หรือไม่ เช่น ผลการสำรวจเรื่องสุขภาพช่องปาก ถ้าพบว่าตัวอย่างที่ศึกษามีสัดส่วนของคนในกลุ่มอายุต่าง ๆไม่เหมือนกับประชากรที่จะนำผลไปใช้ ก็จะไม่สามารถนำผลไปใช้ได้โดยตรง

1.2) **ข้อมูลทางคลินิก** จะใช้พิจารณาตัดสินว่าลักษณะการเจ็บป่วยของตัวอย่างเหมือนผู้ป่วยที่จะทำการรักษาหรือไม่ เช่น ความรุนแรงของอาการป่วย ระยะเวลาที่ป่วย

1.3) **ข้อมูลปัจจัยเสี่ยง** เช่น วิธีการดูแลหญิงตั้งครรภ์ไม่ให้คลอดทารกน้ำหนักน้อย ถ้าตัวอย่างในการศึกษามีแม่วัยรุ่นมาก แต่ในโรงพยาบาลที่จะนำไปใช้มีแม่วัยรุ่นน้อย ซึ่งมีความเสี่ยงต่างกัน

2) **ใช้พิจารณาคุณภาพของการสุ่มตัวอย่าง**

2.1) **งานวิจัยเชิงสำรวจ** หรือการศึกษาระยะยาวที่กำหนดประชากรศึกษาไว้ตามลักษณะภูมิศาสตร์ เช่น กำหนด

Using Statistics in Reseach 1_1 .indd 171 dummy 4 30/6/2557 11:29:41

ให้ประชากรในเขตเทศบาลของจังหวัดในภาคตะวันออกเฉียงเหนือเป็นประชากรศึกษา ผู้อ่านอาจต้องการตรวจสอบว่าตัวอย่างที่สุ่มมาศึกษามีโครงสร้างด้านประชากร อายุ เพศ และอาชีพ เหมือนข้อมูลจากสำมะโนประชากรหรือไม่ ถ้าตัวอย่างที่ศึกษามีลักษณะคล้ายกัน แสดงว่าผลการสุ่มดี ทำให้ผู้อ่านมั่นใจว่าผลการศึกษาจากตัวอย่างที่ได้น่าจะสามารถอ้างอิงไปสู่ประชากรได้อย่างถูกต้อง

2.2) **การวิจัยเชิงทดลอง** ต้องการดูว่าข้อมูลพื้นฐาน ตัวแปรกวน และปัจจัยเสี่ยง มีความเท่าเทียมกันระหว่างกลุ่มหรือไม่ ถ้าเท่าเทียมกันจะทำให้เชื่อมั่นได้ว่าข้อสรุปที่ได้เกิดจากอิทธิพลของสิ่งทดลองเท่านั้น ไม่ได้เกิดจากอิทธิพลของตัวแปรอื่น

3) **ใช้พิจารณาข้อมูลที่เกี่ยวข้องและปัจจัยที่มีอิทธิพลต่อผลของการศึกษา** เช่น นำเสนอรายงานวิจัย เรื่อง ความชุกของโรคหัวใจและหลอดเลือด สิ่งที่ควรนำเสนอคือกลุ่มอายุที่ถูกจัดตามความเสี่ยงของการเป็นโรคหัวใจและหลอดเลือด โดยต้องแยกกลุ่มอายุตามเพศ เช่น ผู้ชายแบ่งอายุตามความเสี่ยง คือ ต่ำกว่า 45 ปี และกลุ่มที่สูงกว่าหรือเท่ากับ 45 ปี ส่วนผู้หญิงแบ่งอายุตามความเสี่ยง คือ ต่ำกว่า 55 ปี และสูงกว่าหรือเท่ากับ 55 ปี การนำเสนอข้อมูลส่วนนี้จะช่วยให้ทราบลักษณะของปัจจัยเสี่ยงของตัวอย่างที่ศึกษา ทำให้การอนุมานผลไปใช้กับประชาชนที่อยู่นอกพื้นที่การศึกษาได้เหมาะสมยิ่งขึ้น

การนำเสนอกลุ่มอายุในบทความวิจัยหนึ่งๆอาจแบ่งมากกว่า 1 แบบ ขึ้นอยู่กับวัตถุประสงค์การนำเสนอ เช่น แบ่งกลุ่มอายุช่วงละ 5 หรือ 10 ปี เพื่อใช้บรรยายลักษณะตาม

Using Statistics in Reseach 1_1 .indd 172 dummy 4 30/6/2557 11:29:41

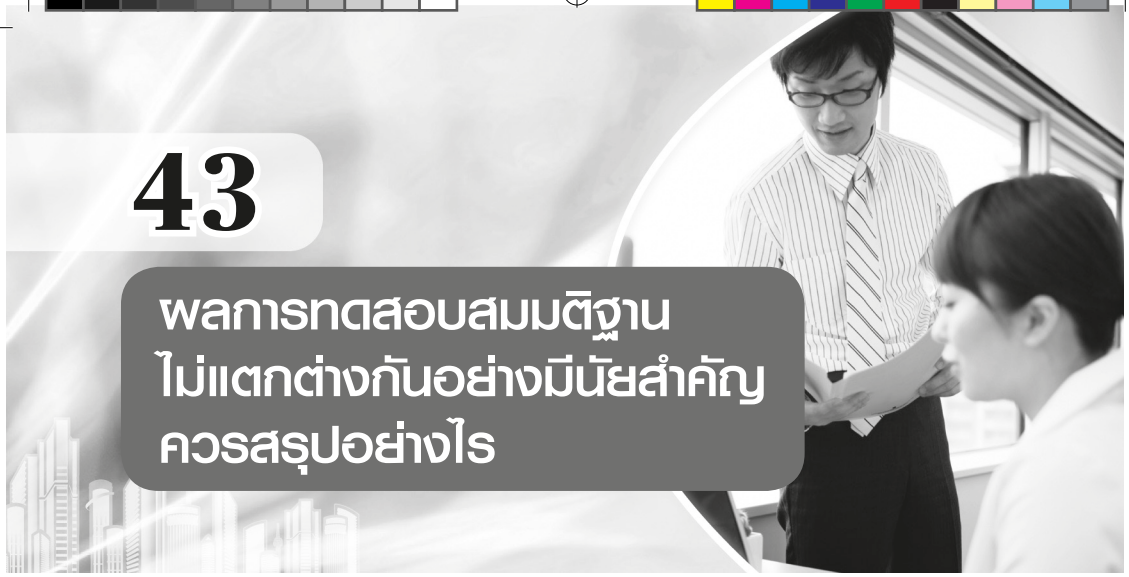
โครงสร้างอายุในประชากร และอาจแบ่งกลุ่มอายุตามความเสี่ยงเพื่อใช้แสดงจำนวนประชากรกลุ่มเสี่ยง

การนำเสนอข้อมูลพื้นฐานไม่ควรนำเสนอโดยการเขียนสรุปอย่างเดียว ควรนำเสนอเป็นตารางหรือแผนภูมิประกอบด้วย เพราะผู้เขียนบทความวิจัยสรุปเฉพาะส่วนที่ต้องการสรุปแต่ผู้อ่านอาจต้องการทราบในส่วนย่อยอื่น เช่น ในการศึกษาความชุกโรคซึมเศร้า ผู้เขียนบรรยายสรุปข้อมูลพื้นฐานว่า *"ตัวอย่างส่วนใหญ่ (23%) อายุ 35-50 ปี ประกอบอาชีพเกษตรกรรม (35%) ..."* โดยไม่มีตารางประกอบ ถ้าผู้อ่านสนใจการเกิดโรคซึมเศร้าในคนงานโรงงานว่ามีเท่าไรจะไม่สามารถทราบได้จากข้อความบรรยายสรุป แต่สามารถหาได้จากตารางแสดงจำนวนโรคซึมเศร้าจำแนกตามการประกอบอาชีพ

สรุป

การกำหนดข้อมูลพื้นฐานที่นำเสนอในบทความวิจัยมีความสำคัญต่อผู้อ่านในการพิจารณานำผลงานวิจัยไปประยุกต์ใช้งาน ผู้เขียนบทความต้องนำเสนอข้อมูลพื้นฐานด้านต่าง ๆให้เพียงพอสำหรับการพิจารณาลักษณะตัวอย่าง การเลือกนำเสนอตัวแปรใดบ้างจะต่างกันไปตามเป้าหมายของบทความและแบบงานวิจัย โครงการวิจัยใหญ่ที่แบ่งการนำเสนอเป็นหลายบทความควรเลือกนำเสนอตัวแปรที่เกี่ยวข้องกับผลการวิจัยเฉพาะส่วนที่นำเสนอในบทความเท่านั้น และไม่ควรนำเสนอข้อมูลพื้นฐานโดยการเขียนสรุปอย่างเดียวควรนำเสนอเป็นตารางหรือแผนภูมิประกอบด้วย

Using Statistics in Reseach 1_1 .indd 173 dummy 4 27/8/2557 11:55:23



43

ผลการทดสอบสมมติฐานไม่แตกต่างกันอย่างมีนัยสำคัญ ควรสรุปอย่างไร

มีคำถามจากนักวิจัยปรึกษาเรื่องการสรุปผลการทดลองใช้ยาสมุนไพรเปรียบเทียบกับยามาตรฐานว่าสามารถลดความดันโลหิตแตกต่างกันหรือไม่ ถ้าผลการทดสอบสมมติฐานค่าเฉลี่ยความดันโลหิตระหว่างกลุ่มที่ใช้ยาสมุนไพรเปรียบเทียบกับกลุ่มใช้ยามาตรฐาน พบว่าผลการทดสอบไม่แตกต่างกันอย่างมีนัยสำคัญ จะสามารถสรุปได้หรือไม่ว่ายาสมุนไพรลดความดันโลหิตได้เท่ากับยามาตรฐาน

เมื่อพบว่าการทดสอบแตกต่างกันอย่างไม่มีนัยสำคัญ ไม่ได้หมายความว่ายาทั้งสองให้ผลเท่ากันหรือไม่แตกต่างกัน แต่หมายความว่าจำนวนตัวอย่างที่ใช้ศึกษาไม่สามารถระบุความแตกต่างของยาทั้งสองได้ ในการสรุปผลการศึกษาควรสรุปว่า ผลการศึกษายังไม่สามารถระบุว่ายาสมุนไพรให้ผลแตกต่างจากยามาตรฐาน

ทำไมเมื่อสรุปผลการทดสอบสมมติฐานที่ไม่สามารถระบุความแตกต่าง จึงสรุปว่าค่าเฉลี่ยที่เปรียบเทียบเท่ากันไม่ได้ ในที่นี้จะอธิบายหลักทางสถิติเพื่อแสดงให้เห็นว่าการที่ไม่สามารถระบุความแตกต่าง มีความหมายแตกต่างกันกับการสรุปว่าเท่ากันอย่างไร

ในการตอบคำถามนี้จะใช้ตัวอย่างการทดสอบสมมติฐานค่าเฉลี่ยระหว่าง 2 ประชากรที่เป็นอิสระต่อกัน โดยมีสมมติฐาน $[H_0: \mu_1 = \mu_2, H_A: \mu_1 \neq \mu_2]$ เป็นแบบ 2 ทาง ทดสอบโดยใช้สถิติ t โดยมีสูตรคำนวณค่าสถิติดังนี้ $t = \bar{x}_1 - \bar{x}_2 / \sqrt{s_p^2/n_1 + s_p^2/n_2}$

ตัวอย่างข้อมูลและผลการคำนวณแสดงดังตาราง

	งานวิจัย 1 n/กลุ่ม = 30	งานวิจัย 2 n/กลุ่ม = 50
$\bar{x}_1 (s_1)$	19.2 (5.8)	19.2 (5.8)
$\bar{x}_2 (s_2)$	22.4 (7.3)	22.4 (7.3)
$\bar{x}_1 - \bar{x}_2$	3.2	3.2
$\sqrt{s_p^2/n_1 + s_p^2/n_2}$	1.7	1.3
t (P value)	1.88 (0.065)	2.43 (0.017)

การสรุปผลการทดสอบสมมติฐาน ถ้ากำหนดระดับนัยสำคัญเท่ากับ 0.05 การทดสอบแบบ 2 ทางจึงนำค่า $\alpha/2$ (0.025) ไปเทียบกับค่า P value ที่คำนวณได้จากค่าสถิติ จากตารางในช่องงานวิจัย 1 เมื่อ n ต่อกลุ่มเท่ากับ 30 ค่า P value (0.065) ซึ่งมีค่ามากกว่าค่า $\alpha/2$ (0.025) ผลการทดสอบไม่สามารถปฏิเสธสมมติฐาน จึงสรุปว่าจากผลการศึกษายังไม่สามารถระบุว่าค่าเฉลี่ยของทั้ง 2 กลุ่มแตกต่างกัน

ถ้ากำหนดให้ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานเท่าเดิม แต่เพิ่มขนาดตัวอย่างให้ใหญ่ขึ้นเป็น 50 รายต่อกลุ่ม ผลการคำนวณอยู่ในตารางช่องงานวิจัย 2 พบว่า ค่า P value (0.017) น้อยกว่าค่า $\alpha/2$ (0.025) ผลการทดสอบปฏิเสธสมมติฐานแสดงว่าค่าเฉลี่ยของทั้ง 2 กลุ่ม

Using Statistics in Reseach 1_1 .indd 175 dummy 4 8/9/2557 11:47:49

แตกต่างกัน

จากตัวอย่างจะเห็นว่าขนาดตัวอย่างมีผลต่อการทดสอบความแตกต่าง ในกรณีตัวอย่างเล็กผลการทดสอบยังไม่สามารถระบุความแตกต่าง ถ้ามีการเพิ่มขนาดตัวอย่างให้มากขึ้น ผลการทดสอบอาจพบความแตกต่างได้ ทั้งนี้ เพราะค่า t คำนวณได้จากความแตกต่างของค่าเฉลี่ยหารด้วยความคลาดเคลื่อนมาตรฐาน ($\sqrt{s_p^2/n_1 + s_p^2/n_2}$) เมื่อขนาดตัวอย่างเพิ่มขึ้น ค่าของความคลาดเคลื่อนมาตรฐานจะเล็กลง ทำให้ t มีค่าเพิ่มขึ้น ถ้าเพิ่มขนาดตัวอย่างมากพอจะทำให้ได้ค่า t ใหญ่พอที่จะระบุความแตกต่างระหว่างกลุ่มได้

สรุป

การที่ไม่ปฏิเสธสมมติฐานจึงไม่ได้หมายความว่าเท่ากัน เพราะถ้าเพิ่มขนาดตัวอย่างที่เพียงพอก็จะสามารถพบความแตกต่างได้

Using Statistics in Reseach 1_1 .indd 176 dummy 4 30/6/2557 11:29:42

44

การแปลผลการทดสอบสมมติฐานของค่าเฉลี่ย

ในรายงานวิจัยที่มีการทดสอบสมมติฐานเพื่อเปรียบเทียบค่าเฉลี่ยระหว่างกลุ่มในงานวิจัยเชิงทดลอง ถ้าพบว่าค่า P value มีค่าต่ำกว่า 0.05 นักวิจัยสรุปว่ามีความแตกต่างกันอย่างมีนัยสำคัญ แสดงว่ากลุ่มทดลองได้ผลลัพธ์ดีกว่ากลุ่มควบคุม และมักมีข้อเสนอแนะให้นำวิธีการที่ศึกษาไปใช้ การที่นักวิจัยพิจารณาจากความแตกต่างกันอย่างมีนัยสำคัญอย่างเดียวแล้วสรุปว่าวิธีการใหม่ได้ผลลัพธ์ดีกว่า อาจได้ข้อสรุปที่ไม่ถูกต้อง

การสรุปผลการทดสอบสมมติฐานที่ถูกต้อง ควรนำช่วงเชื่อมั่นของค่าเฉลี่ยของตัวอย่าง ขนาดอิทธิพล (effect size) และค่า P value มาพิจารณาร่วมกัน

ตัวอย่าง การสรุปผลการทดสอบสมมติฐานในการเปรียบเทียบค่าเฉลี่ยความรู้ของการสอนทันตสุขศึกษาของกลุ่มทดลอง (สอนด้วยวิธีใหม่) และกลุ่มควบคุม (สอนด้วยวิธีเดิม) โดยใช้แบบทดสอบมีคะแนนเต็ม 50 คะแนน มีผลการวิเคราะห์ดังในตาราง

	n	$\bar{x}$	SD	95% CI	P value (t-test)
กลุ่มทดลอง	40	22.5	2.4	21.7, 23.3	
กลุ่มควบคุม	40	20.1	3.1	19.1, 21.1	
ผลต่าง		2.4	SE = 0.626	1.2, 3.6	0.0002

โดยผู้วิจัยสรุปผลว่า *"จากผลการทดสอบสมมติฐาน พบว่า ความรู้ในกลุ่มทดลองมากกว่ากลุ่มควบคุมแตกต่างกันอย่างมีนัยสำคัญ (P = 0.0002) ที่ระดับความเชื่อมั่น 95% ดังนั้นการสอนด้วยวิธีใหม่จะทำให้ผู้เรียนมีความรู้มากกว่าการสอนด้วยวิธีเดิม จึงควรมีการนำการสอนทันตสุขศึกษาไปใช้ในการให้บริการของ รพ.สต."*

การสรุปผลดังกล่าวนำค่า P value มาพิจารณาเพียงค่าเดียว ซึ่งข้อสรุปดังกล่าวไม่ถูกต้อง ทั้งนี้ เพราะเมื่อพิจารณาค่าเฉลี่ยคะแนนทันตสุขศึกษาของตัวอย่างกลุ่มทดลองได้ 22.5 จากคะแนนเต็ม 50 คิดเป็นร้อยละ 45 ของคะแนนเต็ม ถ้าพิจารณาค่า 95% ช่วงเชื่อมั่นของกลุ่มทดลอง พบว่าค่าสูงสุดของช่วงเชื่อมั่นเท่ากับ 23.3 คิดเป็นร้อยละ 47 ของคะแนนเต็ม หมายความว่า ถ้ามีการใช้จริงทั้งประเทศ ค่าเฉลี่ยความรู้สูงสุดของประชาชนจะได้เท่ากับร้อยละ 47 ซึ่งยังไม่ถึงครึ่งความรู้ที่สอน

ถ้าพิจารณาจากค่าผลต่างของค่าเฉลี่ยความรู้ระหว่างกลุ่มควบคุมและกลุ่มทดลอง ค่าสูงสุดของ 95% ช่วงเชื่อมั่นของผลต่างของค่าเฉลี่ยที่ได้ คือ 3.6 เท่ากับร้อยละ 7.2 ของคะแนนเต็ม แสดงว่าการสอนทั้ง 2 วิธีมีความรู้แตกต่างกันเพียงร้อยละ 7.2 ของคะแนนเต็ม

Using Statistics in Reseach 1_1 .indd 178 dummy 4 27/8/2557 11:55:24

ดังนั้นการสรุปผลเปรียบเทียบค่าเฉลี่ยระหว่างกลุ่มที่ดีควรพิจารณาจากข้อมูลต่าง ๆดังนี้

1. ผลการทดสอบสมมติฐานว่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติหรือไม่
2. เมื่อพบว่ามีความแตกต่างกันอย่างมีนัยสำคัญแล้ว ควรพิจารณาดูว่าความแตกต่างที่พบมีความสำคัญพอ (ใหญ่พอ) หรือไม่ เช่น จากค่า 95% ช่วงเชื่อมั่น การสอนวิธีใหม่ให้รู้เพิ่มจากแบบเดิมร้อยละ 1.2–7.2 มีความเป็นไปได้ที่รู้เพิ่มมากที่สุดเพียงร้อยละ 7.2 ถ้าคิดว่าความรู้ที่เพิ่มขึ้นเท่านี้ไม่มีผลต่อการเปลี่ยนแปลงพฤติกรรมมากนัก ก็จะสรุปว่าวิธีการสอนวิธีใหม่มีความรู้เพิ่มจากเดิมไม่มากพอที่จะนำไปใช้ ควรไปพัฒนาหาวิธีการสอนใหม่ที่ดีกว่านี้
3. ต้องพิจารณาช่วงเชื่อมั่นของค่าเฉลี่ยของตัวอย่างแต่ละกลุ่มประกอบด้วย จากค่า 95% ช่วงเชื่อมั่น พบว่าค่าสูงสุดของช่วงเชื่อมั่นของค่าเฉลี่ยความรู้ของกลุ่มทดลองได้ร้อยละ 47.0 ส่วนกลุ่มควบคุมได้ร้อยละ 42.2 ซึ่งทั้ง 2 กลุ่มมีความรู้ไม่ถึงครึ่งหนึ่งของที่สอน แสดงว่าวิธีการสอนที่ใช้ยังดีไม่พอทั้ง 2 วิธี

ถ้ามีการนำข้อมูลจากผลการวิเคราะห์ในตารางมาพิจารณาทั้ง 3 ขั้นตอนแล้ว จะสรุปใหม่ได้ว่า *“ถึงแม้วิธีการสอนวิธีใหม่ได้คะแนนความรู้สูงกว่าการสอนวิธีเดิมอย่างมีนัยสำคัญ แต่คะแนนความแตกต่างที่พบไม่แตกต่างกันมากนัก ประกอบกับค่าเฉลี่ยความรู้ของกลุ่มทดลองและกลุ่มควบคุมมีค่าน้อยกว่าร้อยละ 50 ของคะแนนเต็ม แสดงว่าวิธี*

Using Statistics in Reseach 1_1 .indd 179 dummy 4 30/6/2557 11:29:42

การสอนที่ใช้อยู่และการสอนวิธีใหม่ยังไม่ดีพอ จึงควรพัฒนาวิธีการสอนใหม่ให้ได้ผลมากกว่านี้”

สรุป

ในการสรุปผลการทดสอบสมมติฐานทางสถิติไม่ควรนำข้อสรุปจากค่า P value เพียงอย่างเดียว ควรนำค่า 95% ช่วงเชื่อมั่นของทั้งผลต่าง และค่า 95% ช่วงเชื่อมั่นของค่าเฉลี่ยแต่ละกลุ่มมาร่วมพิจารณาสรุปผลด้วย จึงจะทำให้ได้ข้อสรุปที่ถูกต้อง

Using Statistics in Reseach 1_1 .indd 180 dummy 4 27/8/2557 11:55:24

45

การนำเสนอข้อมูลพื้นฐานในงานวิจัยเชิงทดลอง

ในการนำเสนอผลการวิจัยเชิงทดลอง มักมีคำถามและข้อถกเถียงกันในเรื่องความจำเป็นที่จะต้องมีการทดสอบสมมติฐานว่าข้อมูลพื้นฐานแต่ละตัวมีความแตกต่างกันระหว่างกลุ่มควบคุมและกลุ่มทดลองหรือไม่ ผู้เขียนจึงสรุปข้อคิดเห็นและวิธีการนำเสนอข้อมูลพื้นฐานในงานวิจัยเชิงทดลองดังนี้

ข้อมูลพื้นฐานของงานวิจัยเชิงทดลองมักประกอบไปด้วยข้อมูล 3 ส่วน คือ ส่วนแรกเป็นข้อมูลลักษณะประชากร ส่วนที่สองเป็นลักษณะทางคลินิก และส่วนที่สามเป็นข้อมูลปัจจัยกวนหรือปัจจัยที่มีอิทธิพลต่อผลการศึกษา นักวิจัยต้องนำเสนอข้อมูลพื้นฐานทั้ง 3 ส่วนของแต่ละกลุ่มการทดลองให้อยู่ในตารางหรือแผนภูมิเดียวกัน เพื่อช่วยให้ผู้อ่านสามารถเปรียบเทียบความเหมือนกันของข้อมูลพื้นฐานระหว่างกลุ่มได้ง่าย เช่น การนำเสนอข้อมูลพื้นฐานด้านประชากรในงานวิจัยเปรียบเทียบผลของโปรแกรมการทดลองเลิกสูบบุหรี่ ดังตัวอย่างในตาราง 1 และ 2

ตาราง 1 เพศและกลุ่มอายุ

ลักษณะทางประชากร		กลุ่มควบคุม		กลุ่มทดลอง	
		n	%	n	%
เพศ	ชาย	50	60.0	52	63.4
	หญิง	32	40.0	30	36.6
อายุ	15–20	15	18.2	12	14.6
	20–30	24	29.3	16	19.5
	30–40	19	23.2	25	30.5
	40+	24	29.3	29	35.4

ตาราง 2 ความมั่นใจในตนเองและความสมัครใจเข้าร่วมโครงการ

ปัจจัย		กลุ่มควบคุม		กลุ่มทดลอง	
		n	%	n	%
ความมั่นใจในตนเอง	ต่ำ	14	17.1	8	9.8
	ปานกลาง	40	48.8	32	39.0
	สูง	28	34.1	42	51.2
ความสมัครใจ	สมัครใจ	57	69.5	55	67.1
	ไม่สมัครใจ	25	30.5	27	32.9

Using Statistics in Reseach 1_1 .indd 182 dummy 4 30/6/2557 11:29:43

จากตัวอย่างเป็นการนำเสนอในรูปตาราง แต่อาจนำเสนอในรูปของแผนภูมิก็ได้ เพื่อช่วยให้เปรียบเทียบลักษณะความแตกต่างของข้อมูลพื้นฐานได้ง่าย สำหรับการนำเสนอเป็นตารางควรใส่ทั้งค่าจำนวน (n) และร้อยละ (%) ลงในตารางเพื่อความสะดวกในการเปรียบเทียบด้วย

เมื่อพิจารณาค่าร้อยละของอายุในตาราง 1 และความมั่นใจในตนเองในตาราง 2 พบว่ากลุ่มควบคุมและกลุ่มทดลองมีความแตกต่างกัน นักวิจัยต้องไปพิจารณาต่อว่าสองปัจจัยที่แตกต่างกันนี้จะมีผลต่อความสำเร็จของการเลิกสูบบุหรี่หรือไม่ ถ้าปัจจัยดังกล่าวเป็นตัวกวนอาจทำให้ข้อสรุปการทดลองมีความผิดพลาดได้ ก็ควรใช้วิเคราะห์เพื่อปรับอิทธิพลของตัวแปรกวนดังกล่าว

จึงมีคำถามว่า จำเป็นหรือไม่ที่จะต้องทดสอบสมมติฐานตัวแปรข้อมูลพื้นฐานระหว่างกลุ่มทดลองและกลุ่มควบคุมว่าแตกต่างกัน คำตอบคือไม่จำเป็นต้องทดสอบ ด้วยเหตุผล 2 ประการ คือ

ประการแรก ถ้าดำเนินการทดสอบสมมติฐานความแตกต่างของข้อมูลพื้นฐานแล้วพบว่ามีความแตกต่างอย่างมีนัยสำคัญ จะสรุปได้เพียงว่าวิธีการสุ่มตัวอย่างที่ใช้ไม่ดีพอ ทำให้ลักษณะพื้นฐานของทั้ง 2 กลุ่มแตกต่างกัน หรืออาจมีสาเหตุมาจากการศึกษาใช้ตัวอย่างขนาดเล็ก ทำให้ไม่ครอบคลุมความแตกต่างภายในตัวแปรพื้นฐานได้ทั้งหมด ดังนั้นเมื่อทดสอบสมมติฐานข้อมูลพื้นฐานแล้วพบว่า 2 กลุ่มแตกต่างกัน ผู้วิจัยทำได้แต่เพียงอธิบายว่าอาจจะเกิดจากการสุ่มไม่ดีพอ หรือเกิดจากตัวอย่างมีขนาดเล็ก ซึ่งไม่เชื่อมโยงถึงอิทธิพลของตัวแปรดังกล่าวกับการสรุปผลการวิจัย

ประการที่สอง การทดสอบในแต่ละครั้งจะมีโอกาสเกิดความผิดพลาดในการสรุปผล ยิ่งถ้าทำการทดสอบหลายตัวแปรยิ่งทำให้มีโอกาสสรุปผิดพลาดมากขึ้นเรื่อยๆ เช่น ในกรณีทดสอบข้อมูลพื้นฐาน 20

Using Statistics in Reseach 1_1 .indd 183 dummy 4 30/6/2557 11:29:43

ตัวแปร และกำหนดให้มีความผิดพลาดในการทดสอบแต่ละตัวแปรไว้ร้อยละ 5 จะทำให้การสรุปผลจากการทดสอบ 20 ตัวแปรมีโอกาสสรุปผิด (ไม่แตกต่างแต่สรุปว่าแตกต่าง) อย่างน้อย 1 ตัว

ในการศึกษาที่คาดว่าวิธีการสุ่มตัวอย่างแบบง่ายแล้วจะทำให้ตัวแปรกวนกระจายอยู่ใน 2 กลุ่มไม่เท่ากัน ควรเปลี่ยนวิธีการสุ่มเป็นการสุ่มแบบชั้นภูมิ (stratified random sampling) จะสามารถปรับอิทธิพลของตัวแปรกวนได้ดีที่สุด

ในกรณีที่ไม่ได้วางแผนไว้ก่อน หรือการใช้วิธีสุ่มแบบชั้นภูมิไม่สามารถทำได้ทุกตัวแปรกวน นักวิจัยควรวางแผนให้มีการวิเคราะห์ผลแบบมีการปรับปัจจัยที่มีอิทธิพล หรือเรียกว่าการทำ covariate adjustment เพื่อให้สามารถเปรียบเทียบผลการวิจัยโดยควบคุมอิทธิพลจากปัจจัยกวนที่สำคัญ และนักวิจัยควรระบุวิธีการวิเคราะห์ดังกล่าวไว้ในแผนการวิเคราะห์ข้อมูลในโครงร่างงานวิจัยด้วย ในการวิเคราะห์ให้ดำเนินการวิเคราะห์แบบ covariate adjustment ไปเลย โดยไม่จำเป็นจะต้องมีการทดสอบความแตกต่างของข้อมูลพื้นฐานก่อน

สรุป

ในการเสนอข้อมูลพื้นฐานของการวิจัยเชิงทดลอง ควรนำเสนอข้อมูลพื้นฐานทางด้านประชากรและปัจจัยกวนที่มีอิทธิพลต่อผลการศึกษาในลักษณะเปรียบเทียบในตารางหรือแผนภูมิเดียวกัน โดยไม่จำเป็นจะต้องทดสอบสมมติฐานว่าข้อมูลพื้นฐานตัวใดแตกต่างกันอย่างมีนัยสำคัญบ้าง และควรแก้ไขปัญหาอิทธิพลของตัวแปรกวนด้วยการเปลี่ยนวิธีการสุ่ม ในกรณีที่ไม่สามารถแก้ไขโดยเปลี่ยนวิธีการสุ่ม ให้ใช้วิธีการวิเคราะห์ผลแบบมีการปรับปัจจัยที่มีอิทธิพล

Using Statistics in Reseach 1_1 .indd 184 dummy 4 30/6/2557 11:29:43

46

การแปลผลการทดสอบสมมติฐานของค่าสัดส่วน

ในการทดสอบสมมติฐานเพื่อเปรียบเทียบผลการทดลองระหว่างประชากร 2 กลุ่มที่ผลการทดลองเป็นตัวแปรกลุ่มแบบมี 2 ค่า (dichotomous) เช่น หาย/ไม่หาย รู้/ไม่รู้ วิธีการทดสอบสมมติฐานที่นักวิจัยส่วนใหญ่ใช้ คือ การทดสอบด้วยสถิติ chi-square

ตัวอย่าง ผลการสอนสุขศึกษาเรื่องการป้องกันโรคไข้หวัดนกด้วยวิธีการสอนแบบใหม่และวิธีการสอนแบบเดิม กลุ่มละ 350 คน พบว่าการสอนแต่ละวิธีมีผู้นำไปปฏิบัติได้ถูกต้องดังตารางข้างล่างนี้

	ปฏิบัติถูกต้อง จำนวน (%)	ปฏิบัติไม่ถูกต้อง จำนวน (%)	P value (chi-square)
กลุ่มทดลอง	205 (58.57)	145 (41.43)	0.048
กลุ่มควบคุม	178 (50.86)	172 (49.14)	

"จากผลการทดสอบสมมติฐานพบว่าวิธีการสอนมีความสัมพันธ์กับการปฏิบัติถูกต้องอย่างมีนัยสำคัญ (P = 0.048) ที่ระดับความเชื่อมั่น 95% ดังนั้นการสอนวิธีใหม่มีผู้ปฏิบัติถูกต้องแตกต่างจากวิธีเดิม"

การสรุปผลด้วย P value ของสถิติ chi-square เพียงอย่างเดียวอาจให้ผลสรุปที่ไม่เพียงพอต่อการพิจารณาความสำคัญของการนำผลงานวิจัยไปใช้งาน เพราะจะบอกได้เพียงแต่ว่าแตกต่างหรือไม่แตกต่าง ไม่สามารถสรุปขนาดและทิศทางของความแตกต่าง ในการสรุปควรนำค่าความแตกต่างของค่าสัดส่วนระหว่างกลุ่มจากช่วงเชื่อมั่นมาร่วมพิจารณาด้วย

วิธีการทางสถิติในการเปรียบเทียบความแตกต่างของค่าสัดส่วนของประชากรอีกวิธีหนึ่ง คือ การทดสอบด้วยสถิติ Z ซึ่งการใช้สถิติ Z จะสามารถคำนวณ 95% ช่วงเชื่อมั่นของค่าสัดส่วนและผลต่างของค่าสัดส่วน ซึ่งจะช่วยให้ได้ข้อสรุปถูกต้อง จากตัวอย่างนำมาคำนวณค่าสถิติต่างๆได้ดังนี้

	n	สัดส่วนผู้ปฏิบัติถูกต้อง (p)	95% CI	P value
กลุ่มทดลอง	350	0.586	0.532, 0.637	
กลุ่มควบคุม	350	0.509	0.455, 0.562	
ผลต่าง		0.077	0.08, 0.154	0.048

จากผลการวิเคราะห์พบว่าสัดส่วนการปฏิบัติถูกต้องของทั้ง 2 กลุ่มแตกต่างกันจริง และจากค่า 95% ช่วงเชื่อมั่นของผลต่างมีค่าอยู่

Using Statistics in Reseach 1_1 .indd 186 dummy 4 30/6/2557 11:29:43

ระหว่าง 0.08–0.154 หมายความว่าจากการสอนวิธีใหม่จะทำให้มีผู้ปฏิบัติได้ถูกต้องกว่าการสอนแบบเดิมอย่างน้อยร้อยละ 8 มากที่สุดไม่เกินร้อยละ 15 จากขนาดความแตกต่างที่คำนวณได้จะช่วยพิจารณาว่าควรจะเสนอให้เปลี่ยนวิธีการสอนหรือไม่ กรณีนี้เป็นการวัดการเปลี่ยนพฤติกรรมของการปฏิบัติตัวซึ่งโดยปกติจะเปลี่ยนได้ยาก (มีความรู้เพิ่มแต่ไม่ทำ) ถ้านักวิจัยหารือผู้เชี่ยวชาญในเรื่องนี้แล้วตั้งเกณฑ์ว่าถ้าวิธีการใหม่มีผู้ปฏิบัติได้ถูกต้องเพิ่มขึ้นมากกว่าร้อยละ 5 ก็ควรนำมาใช้แทน การพิจารณาว่าได้ผลต่างเท่าไรจึงจะถือว่ามีความสำคัญ มีประโยชน์ต่อการใช้งาน ควรพิจารณาความยากง่ายของการเปลี่ยนแปลงเรื่องนั้นด้วย ซึ่งต้องพิจารณาโดยผู้ที่ปฏิบัติงานประจำหรือผู้เชี่ยวชาญในเรื่องนั้นๆ จะพิจารณาจากแค่ตัวเลขค่าร้อยละไม่ได้

เมื่อพิจารณาค่า 95% ช่วงเชื่อมั่นของกลุ่มทดลองพบว่ามีค่าอยู่ระหว่าง 0.53–0.64 ซึ่งหมายความว่าการสอนวิธีใหม่จะทำให้มีผู้ปฏิบัติได้ถูกต้องสูงสุดไม่เกิน 64%

ดังนั้นเมื่อนำผลการวิเคราะห์ทั้งหมดมาพิจารณาร่วมกัน จะสรุปได้ว่า *“การสอนวิธีใหม่ทำให้มีผู้ปฏิบัติได้ถูกต้องเพิ่มขึ้นอย่างน้อย 8% ซึ่งควรจะมีการนำการสอนวิธีใหม่ไปใช้สอนแทนวิธีเดิม และจาก 95% ช่วงเชื่อมั่นของกลุ่มทดลองที่พบว่ามีผู้ปฏิบัติได้ถูกต้องสูงสุดไม่เกิน 64% ยังเป็นประเด็นที่ควรมีการพัฒนาการสอนแบบใหม่ที่จะช่วยทำให้มีผู้ปฏิบัติได้ถูกต้องมากกว่านี้ เพราะโรคนี้มีอัตราการตายสูงและการระบาดที่รุนแรง”*

Using Statistics in Reseach 1_1 .indd 187 dummy 4 27/8/2557 11:55:24

สรุป

ในการสรุปผลการทดสอบสมมติฐานค่าสัดส่วน ไม่ควรนำค่า P value จากการทดสอบด้วยสถิติ chi-square มาใช้สรุปผลเพียงอย่างเดียว การที่จะช่วยให้ได้ข้อสรุปที่ถูกต้องควรนำค่า 95% ช่วงเชื่อมั่นของค่าความแตกต่างของค่าสัดส่วนระหว่างกลุ่มและของแต่ละกลุ่มมาพิจารณาร่วมด้วย การพิจารณาขนาดความแตกต่างว่ามีประโยชน์ต่อการใช้งานหรือไม่ต้องพิจารณาความยากง่ายของการเปลี่ยนแปลงเรื่องนั้นด้วย ไม่ควรพิจารณาจากแค่ตัวเลขร้อยละที่เปลี่ยนเพียงอย่างเดียว

Using Statistics in Reseach 1_1 .indd 188 dummy 4 30/6/2557 11:29:43

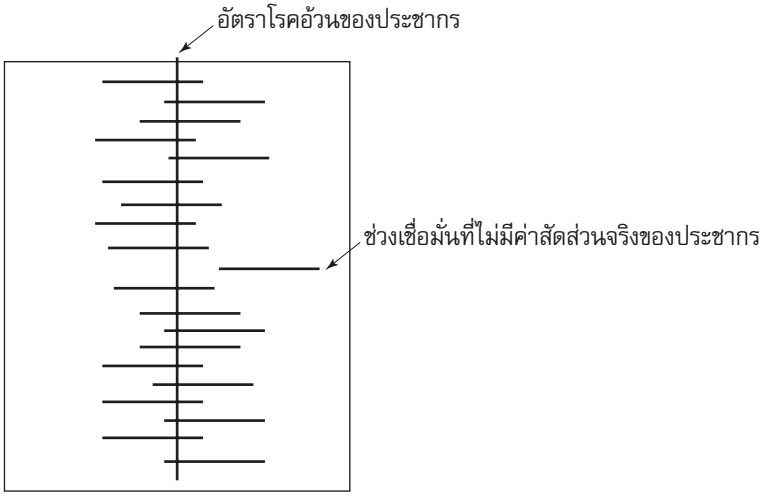
47

ทำไมจึงควรนำเสนอช่วงเชื่อมั่นร่วมกับอัตราหรือค่าเฉลี่ยในการนำเสนอผลการสำรวจด้วยตัวอย่าง

การประเมินสภาวะสุขภาพของประชาชนในพื้นที่โดยใช้วิธีการสำรวจด้วยตัวอย่าง ส่วนใหญ่จะรายงานผลการสำรวจ/วิจัยในรูปของอัตรา เช่น อัตราความชุก อุบัติการณ์ ความครอบคลุม ที่คำนวณได้จากข้อมูลตัวอย่างเพียงค่าเดียว ทำให้ไม่สามารถระบุความน่าเชื่อถือและความถูกต้องของการแปลผลอัตราต่างๆที่ได้จากการสำรวจ

ช่วงเชื่อมั่นเป็นวิธีการทางสถิติที่นำค่าสัดส่วน (อัตราความชุก/ความครอบคลุม) หรือค่าเฉลี่ยของตัวอย่างมาประมาณไปเป็นค่าสัดส่วน/ค่าเฉลี่ยของประชากรด้วยระดับความเชื่อมั่นที่กำหนด เช่น ในการสำรวจอุบัติการณ์ของโรคอ้วนในเด็กวัยเรียนจำนวน 500 คน พบว่าเป็นโรคอ้วน 80 คน คิดเป็นร้อยละ 16 เมื่อนำมาคำนวณ 95% ช่วงเชื่อมั่นอยู่ระหว่างร้อยละ 13.0 ถึง 15.6

การแปลความหมายของช่วงเชื่อมั่นให้พิจารณาจากภาพ 1 เป็นผลสำรวจอัตราโรคอ้วนจากทีมวิจัย 20 ทีมในพื้นที่เดียวกัน ใช้จำนวนตัวอย่างและวิธีสุ่มตัวอย่างเหมือนกัน นำผลการศึกษาของทั้ง 20 ทีมมาคำนวณค่า 95% ช่วงเชื่อมั่นของอัตราโรคอ้วนแล้วนำมาสร้างแผนภูมิโดยเส้นแนวนอนเป็นช่วงเชื่อมั่น เส้นแนวตั้งเป็นค่าอัตราโรคอ้วนที่แท้จริงของประชากรในพื้นที่นั้น



ภาพ 1 แสดง 95% ช่วงเชื่อมั่นของสัดส่วนโรคอ้วน

ภาพ 1 เส้นแนวนอนจำนวน 20 เส้น ใช้แสดง 95% ช่วงเชื่อมั่นของผลการสำรวจเรื่องโรคอ้วนจำนวน 20 ครั้ง โดยกึ่งกลางของแต่ละเส้นจะเป็นค่าสัดส่วนของตัวอย่างแต่ละตัวอย่าง ส่วนเส้นแนวตั้งกลางภาพเป็นเส้นบอกตำแหน่งสัดส่วนโรคอ้วนของประชากร จากภาพจะเห็นได้ว่ามี 19 ครั้งที่ช่วงเชื่อมั่นมีค่าสัดส่วนจริงของประชากรอยู่ในช่วง มีเพียงครั้งเดียวที่ 95% ช่วงเชื่อมั่นไม่มีค่าจริงของประชากร (ครั้งที่ไม่มีเส้นบอกตำแหน่งสัดส่วนโรคอ้วนของประชากรพาดผ่าน) ผลศึกษาที่ทำแบบเดียวกัน 20 ครั้งมีโอกาสที่ได้ช่วงที่ผิด (ไม่มีค่าจริงของประชากรอยู่) 1 ครั้ง คิดเป็นร้อยละ 5 นักวิจัยมีความมั่นใจร้อยละ 95 ว่าช่วงเชื่อมั่นที่คำนวณได้มีโอกาสเป็นช่วงเชื่อมั่นที่มีค่าจริงของประชากรอยู่

การอ่านผลค่าช่วงเชื่อมั่นจากภาพ 1 ช่วงเชื่อมั่นแต่ละช่วงมีค่าจริงของประชากรอยู่ส่วนใดของช่วงก็ได้ จากผลการสำรวจได้ค่า 95% ช่วงเชื่อมั่นของอัตราโรคอ้วน (13.0, 19.6) จึงสรุปได้ว่าอัตราโรคอ้วนของ

Using Statistics in Reseach 1_1 .indd 190 dummy 4 30/6/2557 11:29:43

เด็กนักเรียนในพื้นที่ศึกษาพบอัตราโรคอ้วนสูงสุดไม่เกินร้อยละ 19.6 หรืออาจพบอัตราโรคอ้วนไม่ต่ำกว่าร้อยละ 13

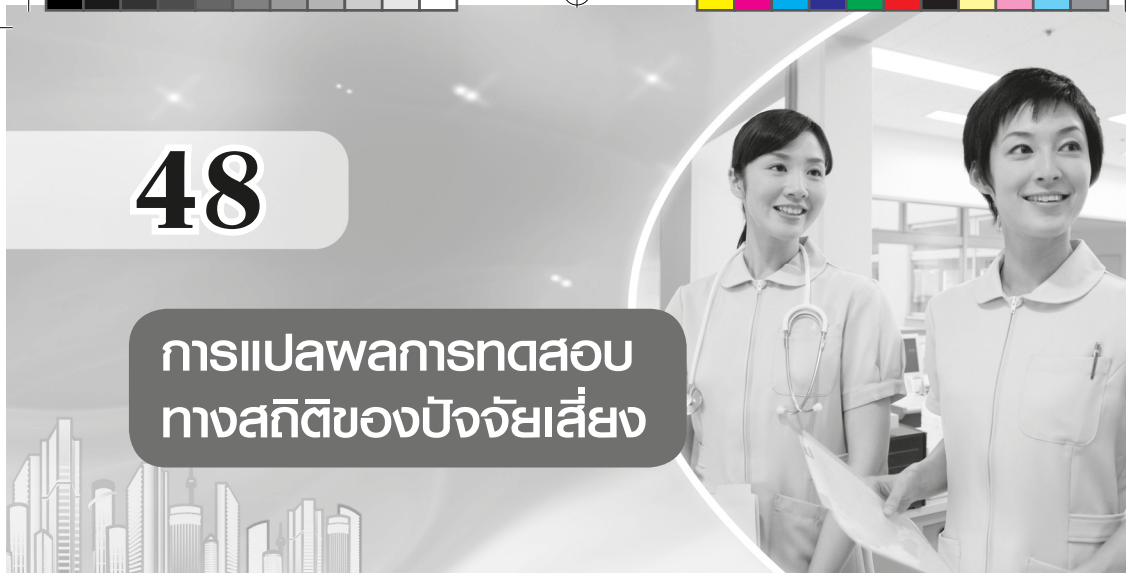
การนำค่าด้านน้อยหรือด้านมากไปใช้ขึ้นอยู่กับวัตถุประสงค์ของการใช้ผลสำรวจ เช่น กรณีการสำรวจเพื่อประเมินผลการดำเนินงานว่าได้ตามเป้าหมายที่ตั้งไว้ว่าจะลดโรคอ้วนในระดับนักเรียนให้เหลือร้อยละ 10 ในกรณีนี้ควรนำค่าต่ำของช่วงเชื่อมั่นไปใช้อภิปรายผล จากการศึกษาพบว่าค่าต่ำสุดของ 95% ช่วงเชื่อมั่นเท่ากับ 13 แสดงว่าอัตราโรคอ้วนของประชากรมีมากกว่าร้อยละ 13 ซึ่งเกินร้อยละ 10 ของเป้าหมาย จึงสรุปว่าผลการดำเนินงานไม่สามารถลดอัตราโรคอ้วนได้ตามเป้าที่กำหนดไว้

ถ้าใช้การประมาณค่าแบบจุดประมาณค่าสัดส่วนของประชากร จากภาพ 1 จะพบว่าช่วงเชื่อมั่นทั้ง 20 ค่าไม่มีตัวอย่างใดมีอัตราโรคอ้วนของตัวอย่างตรงกับของประชากร (เส้นแนวตั้งพาดผ่านตรงกลาง) หมายถึงค่าประมาณแบบจุดประมาณค่าสัดส่วนของประชากรไม่ถูกต้องทั้งหมด ในทางทฤษฎีจะไม่สามารถบอกได้ว่าการประมาณค่าแบบจุดมีโอกาสถูกต้องมากน้อยเท่าไร ดังนั้นจึงไม่ควรใช้การประมาณค่าแบบจุดที่ใช้สัดส่วนจากตัวอย่างเพียงค่าเดียวประมาณค่าสัดส่วนของประชากร

สรุป

ในการรายงานผลการสำรวจ/วิจัย นักวิจัยจะต้องนำเสนอค่าอัตรา (สัดส่วน) พร้อมค่า 95% ช่วงเชื่อมั่น ไม่ควรใช้การประมาณค่าแบบจุด การอภิปรายผลค่า 95% ช่วงเชื่อมั่นจะใช้ทั้งค่าต่ำและค่าสูง หรือจะใช้ค่าใดค่าหนึ่งเพียงค่าเดียวก็ได้ ขึ้นอยู่กับวัตถุประสงค์การใช้ผลสำรวจ

Using Statistics in Reseach 1_1 .indd 191 dummy 4 30/6/2557 11:29:43



48

การแปลผลการทดสอบทางสถิติของปัจจัยเสี่ยง

ความไม่เหมาะสมของการหาปัจจัยที่มีความสัมพันธ์หรือการหาปัจจัยเสี่ยงของการเกิดโรคที่พบในบทความวิจัยมีอยู่ด้วยกัน 2 ลักษณะ คือ ตัวแปรปัจจัยที่มีอยู่ในงานวิจัยไม่ครอบคลุมปัจจัยที่สำคัญ และในการแปลผลการทดสอบไม่ใช้องค์ความรู้ทางเนื้อหาช่วยสรุปลักษณะความสัมพันธ์

การหาปัจจัยเสี่ยงที่ผู้วิจัยนำปัจจัยบางส่วนที่ไม่ครอบคลุมปัจจัยที่สำคัญมาทำการวิเคราะห์สรุปผล โปรแกรมวิเคราะห์จะคำนวณค่าสถิติที่ใช้หาความสัมพันธ์จากข้อมูลของตัวแปรที่ใส่เข้าไปเท่านั้น การแปลผลจะสรุปได้แต่เพียงว่าในบรรดาตัวแปรทั้งหมดที่ใช้ในการศึกษาพบว่ามีปัจจัยใดมีความสัมพันธ์กับการเกิดโรค เช่น ในการศึกษาปัจจัยที่มีผลต่อการตัดสินใจเลือกซื้ออาหารที่มีประโยชน์ต่อสุขภาพ นักวิจัยไม่ได้กำหนดให้มีตัวแปรรสชาติของอาหารอยู่ในข้อมูลที่ทำการศึกษา จึงมีข้อสรุปจากการวิเคราะห์ว่าเพศและฉลากแสดงปริมาณสารอาหารมีผลต่อการตัดสินใจซื้ออาหาร

ถ้านักวิจัยมีตัวแปรรสชาติของอาหารอยู่ในการศึกษาด้วย ผลการศึกษาจะเปลี่ยนไป โดยพบว่ารสชาติและฉลากแสดงปริมาณสารอาหารมีผลต่อการตัดสินใจซื้ออาหาร และยังพบอีกว่าตัวแปรรสชาติเป็นปัจจัยหลักที่ผู้ซื้อให้น้ำหนักร้อยละ 80 ในการตัดสินใจซื้อ

ดังนั้นในการศึกษาปัจจัยเสี่ยงหรือปัจจัยที่มีอิทธิพล การใช้สถิติที่ถูกต้องในการวิเคราะห์อย่างเดียวอาจให้คำตอบที่ผิดพลาดได้ถ้าตัวแปรปัจจัยสำคัญไม่ถูกนำมาศึกษาให้ครบถ้วน นักวิจัยควรทบทวนวรรณกรรมให้ครอบคลุมเพื่อจะได้กำหนดตัวแปรปัจจัยสำคัญได้ครบถ้วน

การแปลผลความสัมพันธ์ในกรณีที่ผลการทดสอบทางสถิติระบุว่ามีความสัมพันธ์ ตัวอย่างเช่น การศึกษาความสัมพันธ์ของการดื่มกาแฟกับการเป็นโรคหัวใจ มีความเป็นไปได้ 5 กรณี คือ

กรณีที่ 1 ความสัมพันธ์ที่พบเป็นความผิดพลาดของการทดสอบสมมติฐาน เพราะในการทดสอบสมมติฐานนักวิจัยต้องกำหนดโอกาสในการสรุปผิดในลักษณะนี้ไว้เท่ากับระดับนัยสำคัญ ความผิดพลาดในกรณีนี้แก้ไขได้ด้วยการเพิ่มขนาดตัวอย่าง และแปลผลด้วยความเข้าใจวิธีการสรุปผลด้วยสถิติ

กรณีที่ 2 ความสัมพันธ์ที่พบเป็นความผิดพลาดจากอคติที่เกิดขึ้นในขั้นตอนต่าง ๆของการทำวิจัย เช่น การกำหนดประชากรศึกษา การสุ่มตัวอย่าง การวัดผล และการวัดปัจจัย ความผิดพลาดที่เกิดจากอคตินี้ไม่สามารถใช้วิธีการทางสถิติในการแก้ไข การแก้ไขต้องทำด้วยการออกแบบงานวิจัยที่ดี

Using Statistics in Reseach 1_1 .indd 193 dummy 4 30/6/2557 11:29:44

กรณีที่ 3 ความสัมพันธ์ที่พบอาจเป็นเหตุเกิดจากผล คนที่เป็นโรคหัวใจรู้สึกเหนื่อยพอดื่มกาแฟทำให้สดชื่น ในการศึกษาจึงพบว่ามีความสัมพันธ์กัน ดังนั้นการดื่มกาแฟอาจมีผลมาจากการเป็นโรคหัวใจ

กรณีที่ 4 ความสัมพันธ์ที่พบเกิดจากทั้งเหตุและผลไปมีความสัมพันธ์กับปัจจัยที่ 3 (ปัจจัยกวนหรือปัจจัยปรับแก้ผล) เช่น การศึกษาความสัมพันธ์ของการดื่มกาแฟกับการเป็นโรคหัวใจ ความสัมพันธ์ที่พบอาจเกิดจากการดื่มกาแฟมีความสัมพันธ์กับการทำงานในห้องปรับอากาศ และผู้ที่ทำงานในห้องปรับอากาศจะไม่ค่อยออกกำลังกาย และบริโภคอาหารไขมันสูง โปรตีนสูง ซึ่งมีความสัมพันธ์กับการเป็นโรคหัวใจ ในการวิเคราะห์จึงพบว่าการดื่มกาแฟกับการเป็นโรคหัวใจมีความสัมพันธ์กัน แต่ในความเป็นจริงการดื่มกาแฟและการเป็นโรคหัวใจไม่ได้มีความสัมพันธ์กันโดยตรง ถ้าผู้แปลผลไม่เข้าใจลักษณะความสัมพันธ์และไม่ได้ทำการวิเคราะห์เพื่อแก้ปัญหาปัจจัยกวนหรือปัจจัยปรับแก้ผล จะทำให้สรุปลักษณะความสัมพันธ์ไม่ถูกต้อง

กรณีที่ 5 ความสัมพันธ์ที่พบเป็นผลเกิดจากเหตุ การดื่มกาแฟเป็นสาเหตุของการเกิดโรคหัวใจ

ดังนั้นการแปลผลความสัมพันธ์ในกรณีที่ผลการทดสอบทางสถิติระบุว่ามีความสัมพันธ์ นักวิจัยต้องมีเหตุผลในการอธิบายว่าความสัมพันธ์ที่พบเป็นผลเกิดจากเหตุจริง

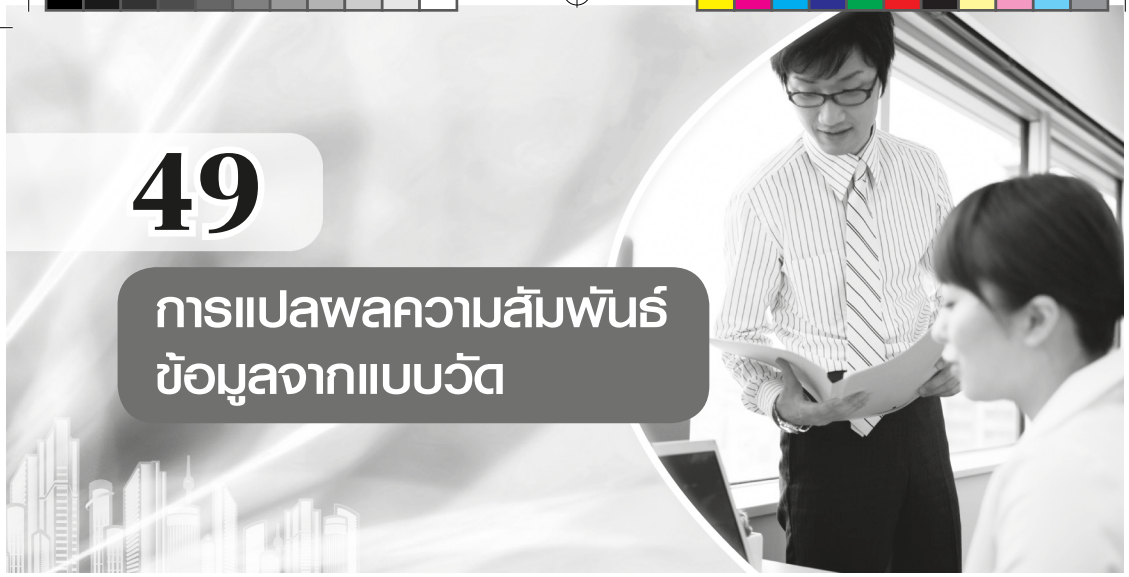
Using Statistics in Reseach 1_1 .indd 194 dummy 4 30/6/2557 11:29:44

สรุป

ในการศึกษาเรื่องความสัมพันธ์ ความครอบคลุมของปัจจัยที่สำคัญช่วยให้การระบุปัจจัยได้ถูกต้อง การออกแบบงานวิจัยและวิธีการวิเคราะห์จะช่วยให้สรุปลักษณะความสัมพันธ์และการแปลผล

ในบทความวิจัยควรอธิบายว่าตัวแปรปัจจัยกำหนดอย่างไร เพื่อแสดงความครอบคลุมของตัวแปรปัจจัยที่สำคัญทั้งหมด และให้รายละเอียดในการออกแบบ รวมถึงวิธีการวิเคราะห์ที่ช่วยควบคุมอิทธิพลปัจจัยกวนและปัจจัยปรับแก้ผล

Using Statistics in Reseach 1_1 .indd 195 dummy 4 30/6/2557 11:29:44



49

การแปลผลความสัมพันธ์ข้อมูลจากแบบวัด

การวิเคราะห์ความสัมพันธ์ข้อมูลที่ได้จากแบบวัด เช่น แบบวัดคุณภาพชีวิตของผู้ป่วย สามารถทำได้ 2 ลักษณะ คือ คำนวณค่าสถิติแสดงความสัมพันธ์จากคะแนนที่วัดได้โดยตรง หรือนำข้อมูลที่วัดได้มาจัดเป็นกลุ่มก่อนนำไปคำนวณค่าสถิติ ในที่นี้จะกล่าวถึงกรณีตัวแปรต้นและตัวแปรตามเป็นข้อมูลจากแบบวัด และคำนวณค่าสถิติจากคะแนนที่วัดได้โดยตรง

ตัวอย่าง ในการศึกษาความสัมพันธ์ระหว่างตัวแปรอิสระ การรับรู้ความรุนแรงของโรคหัวใจและหลอดเลือด กับการควบคุมการบริโภคอาหาร ทั้ง 2 ตัวแปรวัดด้วยแบบวัด ใช้ค่าสถิติสหสัมพันธ์ (r) แสดงระดับความสัมพันธ์

ในการวิเคราะห์มีสมมติฐานว่าถ้าผู้ป่วยรับรู้ความรุนแรงมากจะมีการควบคุมอาหารได้ดี ถ้ารับรู้ความรุนแรงน้อยก็จะควบคุมอาหารได้ไม่ดี เมื่อนำข้อมูลผลการศึกษามาวิเคราะห์พบว่ามีความสัมพันธ์ ($P < 0.05$) แต่ค่าความสัมพันธ์ที่ได้ค่อนข้างต่ำ ($r = 0.12$) จึงอยากทราบว่าควรสรุปผลอย่างไร

การแปลผลนักวิจัยต้องทราบว่าค่าสถิติที่คำนวณได้จะขึ้นอยู่กับคุณภาพเครื่องมือที่ใช้วัด ขนาดตัวอย่าง และการควบคุมปัจจัยกวน ดังนั้นในการสรุปผลนักวิจัยต้องพิจารณาผลกระทบที่อาจเกิดจากปัจจัยดังกล่าวด้วย

จากตัวอย่าง เมื่อพิจารณาเครื่องมือพบว่า นักวิจัยนำเครื่องมือที่เคยใช้วัดการรับรู้ความรุนแรงของโรคหัวใจและหลอดเลือดในผู้ป่วยกลุ่มที่มีความรุนแรงมาใช้ในงานวิจัยที่เป็นการศึกษาในกลุ่มผู้ป่วยใหม่ที่มีความรุนแรงน้อย ผู้ป่วยใหม่กลุ่มนี้หลังจากรักษาแล้วสามารถมีชีวิตได้เทียบเท่าคนปกติ อาการต่าง ๆที่เกิดขึ้นในกลุ่มที่มีความรุนแรงมากแตกต่างจากกลุ่มผู้ป่วยใหม่ที่มีความรุนแรงน้อย ทำให้ผลการรับรู้ความรุนแรงที่ได้อาจมีค่าคะแนนต่ำ การกระจายของคะแนนการรับรู้เรื่องความรุนแรงพบว่า เกือบทั้งหมดมีค่าคะแนนต่ำใกล้เคียงกันทั้งกลุ่ม แต่คะแนนพฤติกรรมการบริโภคอาหารมีการกระจายตั้งแต่ควบคุมได้ดีจนถึงควบคุมได้ไม่ดี ลักษณะข้อมูลแบบนี้เมื่อนำมาคำนวณค่า r จะได้ค่า r ที่ค่อนข้างต่ำเพราะตัวแปรการรับรู้เรื่องความรุนแรงมีการเปลี่ยนแปลงค่อนข้างน้อย แต่พฤติกรรมการบริโภคอาหารมีการเปลี่ยนแปลงมาก

ดังนั้นในการอภิปรายผลควรวิเคราะห์เรื่องคุณสมบัติของเครื่องมือด้วย ในกรณีนี้เครื่องมืออาจไม่ไวพอ (sensitive to change) ไม่สามารถแยกแยะผู้ป่วยกลุ่มที่มีอาการไม่รุนแรงที่มีระดับความรับรู้ต่างกันให้มีคะแนนที่ต่างกันได้ ผลการวิเคราะห์จึงได้ค่าความสัมพันธ์ที่ต่ำ

ขนาดตัวอย่างเป็นอีกปัจจัยหนึ่งที่มีผลต่อความน่าเชื่อถือของผลการศึกษา ในกรณีที่นักวิจัยใช้ตัวอย่าง

Using Statistics in Reseach 1_1 .indd 197 dummy 4 30/6/2557 11:29:44

ขนาดเล็ก (ไม่ได้มีการคำนวณขนาดตัวอย่าง หรือมีการคำนวณขนาดตัวอย่างที่ไม่เหมาะสม) อาจมีผลทำให้ไม่พบความสัมพันธ์ หรืออาจพบความสัมพันธ์ที่มีค่าช่วงเชื่อมั่นกว้าง แปลผลยาก เช่น r = 0.41 95% ช่วงเชื่อมั่น (0.14–0.62) ค่าช่วงเชื่อมั่นที่ได้แสดงว่ามีความเป็นไปได้ที่จะมีความสัมพันธ์ต่ำ (r = 0.14) หรือสูง (r = 0.62) ทำให้สรุปวิจารณ์ผลไม่ได้

ส่วนปัจจัยที่มีอิทธิพลหรือปัจจัยกวนต่างๆ นักวิจัยจะต้องพิจารณาจาก conceptual framework หรือ theoretical framework ว่ามีปัจจัยใดบ้าง ถ้ามีในการวิเคราะห์ต้องใช้วิธีการวิเคราะห์ที่ปรับอิทธิพลของปัจจัยกวน

นอกจากที่ได้กล่าวมาแล้ว ในการแปลผลนักวิจัยจะต้องทราบลักษณะระดับความสัมพันธ์ที่วัดได้จากเครื่องมือแต่ละประเภท เช่น ถ้าเป็นข้อมูลที่วัดจากเครื่องมือทางวิทยาศาสตร์ ความสัมพันธ์สูงจะต้องมีค่า r ตั้งแต่ 0.7 ขึ้นไป อาจมีค่าใกล้ 1 แต่ถ้าเครื่องมือเป็นแบบวัด (สร้างจากชุดคำถาม) ค่า r 0.4 ขึ้นไปก็นับว่ามีความสัมพันธ์สูงแล้ว ดังนั้นถ้านักวิจัยทราบว่างานวิจัยที่ผ่านมาค่า r ที่พบมีค่าเท่าใดบ้าง ค่าสูงสุดและต่ำสุดเป็นเท่าใด จะช่วยในการอภิปรายผลได้มาก

ข้อควรระวังในการแปลผลอีกเรื่องหนึ่งคือวิธีการให้ค่าคะแนนกับคำตอบ จากตัวอย่างถ้าให้ค่าการรับรู้ความรุนแรงของโรคหัวใจและหลอดเลือด คะแนนต่ำหมายถึงการรับรู้น้อย คะแนนสูงหมายถึงการรับรู้มาก ส่วนการควบคุมการบริโภคอาหาร ถ้าคะแนนต่ำหมายถึงควบคุมได้ไม่ดี คะแนนสูงหมายถึงควบคุมได้ดี ค่า r ที่ได้จะมีค่าเป็นบวก (positive correlation) แต่ถ้าการให้ค่าคะแนนของการควบคุมการบริโภคอาหาร

Using Statistics in Reseach 1_1 .indd 198 dummy 4 30/6/2557 11:29:44

เปลี่ยนเป็นคะแนนต่ำหมายถึงควบคุมได้ดี คะแนนสูงหมายถึงควบคุมได้ไม่ดี ค่า r ที่ได้จะมีค่าเป็นลบ (negative correlation) ดังนั้นการแปลผลต้องทราบวิธีการกำหนดรหัสหรือวิธีการให้คะแนนคำตอบด้วย

สรุป

ในการแปลผลความสัมพันธ์ของตัวแปรที่ข้อมูลได้จากแบบวัด นักวิจัยจะพิจารณาเฉพาะค่าสถิติและผลการทดสอบสมมติฐานเพียง 2 อย่างไม่ได้ การแปลผลที่เหมาะสมจะต้องพิจารณาการกระจายของข้อมูลที่วัดได้ (เกี่ยวกับคุณภาพของเครื่องมือ) ขนาดตัวอย่าง อิทธิพลของตัวแปรกวน วิธีการให้คะแนน และลักษณะระดับความสัมพันธ์ที่วัดได้จากเครื่องมือแต่ละประเภทด้วย จึงจะสามารถแปลผลค่าสถิติที่คำนวณได้อย่างถูกต้องและเหมาะสม

Using Statistics in Reseach 1_1 .indd 199 dummy 4 30/6/2557 11:29:44

50

การพิจารณานำงานวิจัยไปใช้

เมื่อผู้ปฏิบัติงานอ่านบทความวิจัยแล้วพบว่าตรงกับปัญหาที่ประสบอยู่ จึงนำวิธีที่ใช้ในรายงานวิจัยไปใช้แก้ปัญหา เมื่อนำไปใช้แล้วพบว่ามีทั้งที่ได้ผลเช่นเดียวกันกับที่พบในงานวิจัยและไม่ได้ผล จึงเกิดคำถามว่าจะมีวิธีการเลือกงานวิจัยอย่างไรดี ต่อไปนี้เป็นบทสนทนาเกี่ยวกับเรื่องดังกล่าว

คำถาม: อยากทราบว่าผลการวิจัยที่สรุปว่าวิธีการดูแลผู้ป่วยเบาหวานรูปแบบใหม่สามารถช่วยป้องกันการเกิดแผลที่เท้าแตกต่างจากวิธีเดิมอย่างมีนัยสำคัญ ที่ระดับความเชื่อมั่น 95% ($P < 0.05$) ถ้านำวิธีการดูแลผู้ป่วยเบาหวานรูปแบบดังกล่าวไปใช้แล้วจะได้ผลหรือไม่

คำตอบ: ก่อนจะตอบปัญหานี้อยากให้เข้าใจก่อนว่างานวิจัยที่มีคำตอบที่เชื่อถือได้ต้องนำมาจากงานวิจัยที่มีวิธีการวิจัยที่ดี ซึ่งผู้อ่านจะต้องมีความรู้วิธีวิภาคย์ความน่าเชื่อถือของงานวิจัย เพื่อเลือกเฉพาะงานวิจัยที่มีวิธีการวิจัยที่ดีมาพิจารณานำผลไปใช้งาน ถ้าพิจารณาตามหลักการ

วิภาคย์ความน่าเชื่อถือของรายงานวิจัยแล้วพบว่างานวิจัยเรื่องการดูแลผู้ป่วยเบาหวานรูปแบบใหม่เป็นงานวิจัยที่ดี ผลสรุปว่าวิธีการดูแลผู้ป่วยเบาหวานป้องกันการเกิดแผลที่เท้าดีกว่าวิธีเดิมอย่างมีนัยสำคัญที่ระดับความเชื่อมั่น 95% หมายถึงข้อมูลจากงานวิจัยให้การยืนยันว่าวิธีการดูแลผู้ป่วยด้วยวิธีใหม่ให้ผลดีกว่าวิธีเดิม แต่บอกไม่ได้ว่าดีขึ้นกว่าเดิมมากน้อยเท่าไร

คำถาม: ถ้าผลต่างที่พบในงานวิจัยไม่มาก เวลาไปใช้ในงานประจำอาจได้ผลไม่ต่างจากเดิมใช่หรือไม่

คำตอบ: ใช่ และนักวิจัยควรพิจารณาขนาดของผลต่างจาก 95% ช่วงเชื่อมั่น ในกรณีที่ผู้นิพนธ์ไม่ได้นำเสนอช่วงเชื่อมั่นในรายงานวิจัย ผู้อ่านที่จะนำผลไปใช้ต้องคำนวณเองจากข้อมูลที่มีอยู่ในรายงานวิจัย

คำถาม: จะรู้ได้อย่างไรว่าการดูแลระบบใหม่จะมีการเกิดแผลลดลงจากเดิมมากน้อยเท่าไร

คำตอบ: ให้พิจารณาจาก 95% ช่วงเชื่อมั่นของผลต่างของอัตราการเกิดแผลที่เท้า เช่น ถ้า 95% ช่วงเชื่อมั่นเท่ากับร้อยละ 12–18 หมายถึงถ้านำเอาผลงานวิจัยไปใช้อาจจะพบผู้เป็นแผลที่เท้าลดลงอย่างน้อยร้อยละ 12 หรืออาจสูงถึงร้อยละ 18 อัตราการลดลงที่พบนี้เป็นการดูแลในบริบทของการทำวิจัย เมื่อไปใช้ภายใต้บริบทของงานประจำจะไม่ได้ผลดีเท่ากับที่พบในงานวิจัย ถ้าคาดว่าเมื่อนำเอาไปใช้ในบริบทงานประจำได้ผลลดลงอีก

Using Statistics in Reseach 1_1 .indd 201 dummy 4 27/8/2557 11:55:25

ร้อยละ 20 ผลที่ได้จะลดลงเหลือร้อยละ 9.6–14.4 ให้พิจารณาการเกิดแผลลดลงจากเดิมจากค่าต่ำของช่วงเชื่อมั่นที่ปรับแล้ว (9.6)

คำถาม: จะรู้ได้อย่างไรว่าวิธีการดูแลผู้ป่วยจากโครงการวิจัยที่นำไปใช้จะทำได้เช่นเดียวกันกับที่นักวิจัยใช้

คำตอบ: ในบทความวิจัยผู้นิพนธ์ต้องเขียนวิธีการดูแลผู้ป่วย (intervention) ที่มีรายละเอียดวิธีดำเนินงานที่ชัดเจนพอที่ผู้อ่านสามารถนำไปใช้ได้

คำถาม: ถ้าในบทความวิจัยมีรายละเอียดไม่พอจะทำอย่างไร

คำตอบ: ในกรณีที่มีรายละเอียดของวิธีการดูแลผู้ป่วยไม่เพียงพอ ผู้ที่จะนำผลงานไปใช้สามารถติดต่อขอรายละเอียดต่างๆเพิ่มเติมจากนักวิจัย โดยปกติในบทความจะมีชื่อที่อยู่ของนักวิจัยที่จะเป็นผู้ตอบคำถาม ถ้าไม่มีให้ติดต่อขอจากผู้ที่มีชื่อแรกในบทความ

การนำวิธีการดูแลผู้ป่วยไปใช้จริงอาจได้ผลไม่เท่ากับที่พบในงานวิจัยเพราะความใส่ใจของผู้ให้บริการในบริบทของการทำงานประจำ ปัจจัยที่มีอิทธิพล และลักษณะผู้ป่วยแตกต่างกัน ดังนั้นในการพิจารณานำไปใช้จะต้องพิจารณาปัจจัยเหล่านี้ประกอบด้วย

คำถาม: จะรู้ได้อย่างไรว่าวิธีการดูแลนี้เหมาะกับผู้ป่วยที่จะนำรูปแบบนี้ไปใช้

Using Statistics in Reseach 1_1 .indd 202 dummy 4 27/8/2557 11:55:25

คำตอบ: ในบทความวิจัยจะนำเสนอข้อมูลพื้นฐานผู้ป่วย ซึ่งเป็นส่วนที่บอกลักษณะของตัวอย่างที่ใช้ในการศึกษา นอกจากลักษณะทางด้านประชากรแล้วยังมีลักษณะทางคลินิกและปัจจัยเสี่ยง ข้อมูลเหล่านี้จะใช้ในการพิจารณาว่าผู้ป่วยกลุ่มที่ใช้ศึกษามีลักษณะเหมือนกับผู้ป่วยที่ต้องการจะนำรูปแบบนี้ไปใช้หรือไม่ ถ้าคล้ายกันก็มีโอกาสได้ผลเช่นเดียวกันกับที่พบในบทความวิจัย

คำถาม: การพิจารณาลักษณะของผู้ป่วยจากข้อมูลพื้นฐานด้านต่างๆที่นำเสนอในบทความวิจัยเพียงพอสำหรับการพิจารณาหรือไม่

คำตอบ: โดยปกติผู้นิพนธ์จะนำเสนอรายละเอียดเป็นตารางและเขียนบรรยายสรุป อาจมีภาพประกอบบรรยายสรุปลักษณะของตัวอย่างเพื่อให้ผู้อ่านเข้าใจโดยไม่ต้องอ่านจากตาราง แต่ข้อสรุปที่นักวิจัยเขียนอาจไม่ครอบคลุมรายละเอียดทั้งหมด หรืออาจไม่ตรงประเด็นสำหรับใช้พิจารณา ดังนั้นผู้จะนำผลงานวิจัยไปใช้จึงควรอ่านรายละเอียดโดยตรงจากตารางข้อมูลพื้นฐาน ในกรณีต้องการรู้ข้อมูลเพิ่มเติมนอกเหนือที่มีอยู่ในบทความวิจัย ผู้อ่านสามารถติดต่อขอข้อมูลเพิ่มเติมได้จากผู้นิพนธ์

คำถาม: การพิจารณาว่าลักษณะกลุ่มผู้ป่วยในกลุ่มทดลองเหมือนกับกลุ่มผู้ป่วยในกลุ่มควบคุม จะต้องดูว่าคล้ายกันทุกตัวแปรหรือไม่

Using Statistics in Reseach 1_1 .indd 203 dummy 4 30/6/2557 11:29:45

คำตอบ: ในการพิจารณาความคล้ายกันของผู้ป่วยจะพิจารณาจากการกระจายของตัวแปรกวนที่สำคัญว่าเท่าเทียมกันในแต่ละกลุ่มการทดลองหรือไม่ เช่น สัดส่วนความหนักเบาของโรค การควบคุมระดับน้ำตาลในเลือด ฯลฯ

สรุป

จากบทสนทนาข้างต้นจะเป็นแนวทางการพิจารณานำผลงานวิจัยไปใช้กับการแก้ปัญหาต่างๆ และถ้าผู้อ่านเป็นผู้เขียนบทความวิจัยจะได้เข้าใจว่าข้อมูลแต่ละเรื่องที่มีอยู่ในบทความวิจัยใช้พิจารณาอะไรบ้าง จะได้เขียนบทความให้มีข้อมูลครบถ้วนในการพิจารณานำผลวิจัยไปใช้งาน

Using Statistics in Reseach 1_1 .indd 204 dummy 4 30/6/2557 11:29:45

การออกแบบแบบสอบถามสำหรับงานวิจัย

ผู้เขียน รองศาสตราจารย์อรุณ จิรวัฒน์กุล
เหมาะสำหรับ นักศึกษา ครู อาจารย์ นักวิจัย นักบริหาร นักการตลาด และผู้ที่ต้องทำงานเกี่ยวข้องกับการสร้างแบบสอบถามเพื่อการวิจัย
ขนาด 6" x 8" กระดาษปรู๊ฟ
ราคา 100 บาท จำนวนหน้า 144 หน้า

หนังสือ **การออกแบบแบบสอบถามสำหรับงานวิจัย** เล่มนี้เขียนขึ้นจากประสบการณ์ของผู้เขียนทั้งในฐานะนักวิจัยและที่ปรึกษาโครงการวิจัยจำนวนมากมากว่า 30 ปี เนื้อหาภายในเล่มให้ความรู้เกี่ยวกับการออกแบบแบบสอบถามที่ถูกต้องและเหมาะสมเพื่อให้ได้ผลการวิจัยที่มีคุณภาพที่สุด เนื้อหาครอบคลุมเรื่อง *ลักษณะของแบบสอบถาม ประเภทของข้อมูล รูปแบบและวิธีเขียนคำถาม วิธีกำหนดการตอบคำถาม การเขียนข้อคำถาม การออกแบบหน้ากระดาษ การออกแบบรหัสสำหรับประมวลผล การทดสอบแบบสอบถาม การเขียนคำถามที่ช่วยกระตุ้นความจำในการถามเหตุการณ์ย้อนหลัง หลักจริยธรรมการวิจัยในมนุษย์ที่เกี่ยวข้องกับการใช้แบบสอบถามในงานวิจัย*

การออกแบบแบบสอบถามสำหรับงานวิจัย เล่มนี้ใช้ภาษาที่เข้าใจง่ายและอธิบายอย่างเป็นลำดับขั้นตอน สามารถศึกษาได้เองและนำไปใช้ได้จริง จึงเหมาะสำหรับผู้ที่ต้องการให้ผลการวิจัยถูกต้อง เชื่อถือได้ และตรงตามความเป็นจริงมากที่สุด

สถิติทางวิทยาศาสตร์สุขภาพเพื่อการวิจัย

ผู้เขียน รองศาสตราจารย์อรุณ จิรวัฒน์กุล
เหมาะสำหรับ นิสิต นักศึกษาทั้งระดับปริญญาตรีและปริญญาโท และนักวิจัยด้านวิทยาศาสตร์สุขภาพ
ขนาด 7" x 10" กระดาษปอนด์
ราคา 230 บาท จำนวนหน้า 312 หน้า

ตำรา **สถิติทางวิทยาศาสตร์สุขภาพเพื่อการวิจัย** เล่มนี้เขียนขึ้นจากประสบการณ์ในการเป็นผู้สอน นักวิจัย นักสถิติ และผู้อ่านพิจารณาการใช้สถิติในบทความวิจัยของวารสารวิชาการ เนื้อหาภายในเล่มให้ความรู้เกี่ยวกับการใช้สถิติต่าง ๆสำหรับงานวิจัยทางวิทยาศาสตร์สุขภาพ อาทิ ความรู้พื้นฐานเกี่ยวกับสถิติสำหรับงานวิจัยทางวิทยาศาสตร์สุขภาพ การพรรณนาลักษณะตัวอย่าง ความน่าจะเป็นและช่วงความเชื่อมั่น การทดสอบสมมติฐาน การเปรียบเทียบผลลัพธ์ การหาความสัมพันธ์ของตัวแปร การคำนวณขนาดตัวอย่าง วิธีเลือกใช้สถิติสำหรับแบบงานวิจัย สถิติที่ควรมีในโครงร่างงานวิจัย การวิเคราะห์ความแปรปรวน การวิเคราะห์การถดถอยพหุคูณ การวิเคราะห์ระยะเวลาการรอดชีพ สถิติแบบไม่ใช้พารามิเตอร์ การวางแผนประมวลผล การวิเคราะห์ข้อมูล และการแปลงค่าข้อมูล เนื้อหาเขียนอย่างเป็นระบบ ใช้ภาษาที่เข้าใจง่าย มีภาพประกอบเป็นจำนวนมากเพื่อเสริมความเข้าใจ อีกทั้งเนื้อหาเชื่อมโยงเข้ากับแบบงานวิจัย รวมทั้งมีการแปลผลที่สอดคล้องกับการนำไปใช้ในชีวิตจริง ตำราเล่มนี้จึงเหมาะสำหรับผู้ที่เริ่มต้นศึกษาวิชาสถิติด้านวิทยาศาสตร์สุขภาพ ตลอดจนนักศึกษาระดับบัณฑิตศึกษาและนักวิจัยที่ต้องทำงานวิจัยด้านวิทยาศาสตร์สุขภาพ

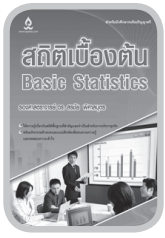
วิจัยจากงานประจำ เรียนแบบทีละขั้น

ผู้เขียน ศาสตราจารย์ ดร. ศิริพร จิรวัฒน์กุล
รองศาสตราจารย์อรุณ จิรวัฒน์กุล
เหมาะสำหรับ บุคลากรในสาขาวิทยาศาสตร์สุขภาพ สาขาการศึกษา และผู้ที่ต้องการพัฒนางานประจำที่จำเจ
ขนาด 6" x 8" กระดาษปอนด์
ราคา 110 บาท จำนวนหน้า 160 หน้า

อาร์ทูอาร์ (R to R หรือ R 2 R) คือการวิจัยจากงานประจำ (Routine to Research) คำว่างานประจำมักทำให้เกิดความรู้สึกว่าจำเจ รวมทั้งคำว่างานวิจัยมักทำให้เกิดความรู้สึกว่ายาก น่ากลัว หรือน่าเบื่อหน่าย แต่การวิจัยจากงานประจำไม่ใช่เรื่องที่ยาก จำเจ หรือว่าน่าเบื่อหน่าย ตรงกันข้าม กลับทำให้งานประจำที่หลายคนอาจรู้สึกว่าจำเจกลายเป็นงานที่ท้าทายขึ้นมา

อาร์ทูอาร์เป็นเครื่องมืออย่างหนึ่งในการสร้างความรู้เพื่อนำมาใช้ในการปรับปรุง แก้ไข หรือพัฒนางานและตัวบุคคล การทำอาร์ทูอาร์ไม่ใช่เรื่องที่ยากเกินความสามารถหากแต่เป็นเทคนิคที่สามารถเรียนรู้ได้ **วิจัยจากงานประจำ เรียนแบบทีละขั้น** เล่มนี้เป็นหนังสือที่เขียนขึ้นจากประสบการณ์จริงในการทำอาร์ทูอาร์และนำมาถ่ายทอดแบบทีละขั้นเพื่อให้ผู้สนใจสามารถติดตามและเรียนรู้ไปตั้งแต่ขั้นเริ่มต้นไปจนถึงขั้นการนำผลไปใช้เพื่อพัฒนาคนและพัฒนางาน

วิจัยจากงานประจำ เรียนแบบทีละขั้น เล่มนี้เขียนด้วยภาษาที่อ่านเข้าใจง่าย พร้อมทั้งมีตัวอย่างโครงร่างการวิจัยที่จะช่วยให้เข้าใจการนำอาร์ทูอาร์ไปใช้กับปัญหาลักษณะต่าง ๆในงานประจำ เหมาะสำหรับเป็นคู่มือสำหรับผู้ที่ต้องการพัฒนาผลงานจากงานประจำด้วยการทำวิจัย ไม่ว่าจะเป็นบุคลากรในสาขาวิทยาศาสตร์สุขภาพ สาขาการศึกษา รวมทั้งผู้ปฏิบัติงานในสาขาอื่น ๆที่ต้องการนำอาร์ทูอาร์ไปใช้ให้ประสบผลสำเร็จ



สถิติเบื้องต้น

ผู้เขียน รองศาสตราจารย์ ดร. สรชัย พิศาลบุตร
เหมาะสำหรับ นักศึกษาระดับปริญญาตรี
ขนาด 7" x 10" กระดาษปอนด์
ราคา 150 บาท จำนวนหน้า 256 หน้า

สถิติมีความสำคัญและจำเป็นอย่างมากต่อการดำเนินงานทางธุรกิจ เพราะต้องตัดสินใจ วางแผน และแก้ปัญหาต่าง ๆอยู่ตลอดเวลา ซึ่งต้องอาศัยข้อมูลที่ผ่านการเก็บรวบรวมและวิเคราะห์อย่างมีระบบ และถูกต้องตามหลักวิชา

ตำรา **สถิติเบื้องต้น** เล่มนี้มีเนื้อหาเกี่ยวกับ *ความสำคัญของสถิติที่มีต่อการบริหารธุรกิจ ข้อมูลและการวิเคราะห์ข้อมูลธุรกิจเบื้องต้น ความน่าจะเป็นและการนำไปใช้เพื่อการตัดสินใจทางธุรกิจ การเก็บรวบรวมข้อมูลทางธุรกิจ การประมาณค่าและการทดสอบสมมติฐานทางธุรกิจ* รวมทั้ง *การหาความสัมพันธ์และการพยากรณ์ทางธุรกิจ* นอกจากนี้แต่ละบทยังมีกิจกรรมท้ายบทและแบบฝึกหัดให้ทบทวนและฝึกฝนความเข้าใจด้วย ตำราเล่มนี้จึงเหมาะสำหรับนักศึกษาระดับปริญญาตรี โดยเฉพาะสาขาวิชาบริหารธุรกิจและเศรษฐศาสตร์ ตลอดจนผู้ที่ต้องใช้ข้อมูลและสารสนเทศทางสถิติในการบริหารและประกอบธุรกิจ

สถิติธุรกิจ

ผู้เขียน **รองศาสตราจารย์ ดร. สรชัย พิศาลบุตร**
เหมาะสำหรับ **นักศึกษาระดับปริญญาตรี**
ขนาด 7" x 10" **กระดาษปรู๊ฟ**
ราคา 250 บาท **จำนวนหน้า 412 หน้า**

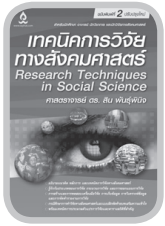
ตำรา **สถิติธุรกิจ** เล่มนี้มีเนื้อหาครอบคลุมหลักสูตรวิชาสถิติธุรกิจระดับปริญญาตรีที่ใช้ในมหาวิทยาลัยทั้งของรัฐและเอกชนเกือบทุกแห่ง เนื้อหาภายในเล่มให้ความรู้เกี่ยวกับทฤษฎีและวิธีการเชิงสถิติต่าง ๆที่จำเป็นสำหรับการเก็บรวบรวมและการวิเคราะห์ข้อมูลเพื่อการตัดสินใจและวางแผนทางธุรกิจ ได้แก่ สถิติกับการตัดสินใจทางธุรกิจ ความน่าจะเป็นและการนำไปใช้ประโยชน์ทางธุรกิจ การสำรวจตัวอย่าง การประมาณค่า การทดสอบสมมติฐานทางสถิติ การวิเคราะห์ความแปรปรวน การวิเคราะห์การถดถอยและสหสัมพันธ์เบื้องต้น การวิเคราะห์การถดถอยพหุคูณ การวิเคราะห์อนุกรมเวลา การวิเคราะห์ข้อมูลที่อยู่ในรูปความถี่ การทดสอบสมมติฐานที่ไม่ใช้พารามิเตอร์ เลขดัชนี และการตัดสินใจเชิงสถิติ โดยเนื้อหาในแต่ละบทอธิบายด้วยภาษาที่เข้าใจง่าย พร้อมตัวอย่างการประยุกต์ทางธุรกิจมากมาย ทั้งยังมีแบบฝึกหัดท้ายบทเพื่อเสริมความเข้าใจ และกิจกรรมในชั้นเรียนของทุกบทเพื่อการสร้างเสริมนักศึกษาในการเป็นศูนย์กลางของการเรียนรู้อีกด้วย

ตำรา **สถิติธุรกิจ** เล่มนี้จึงเป็นตำราที่ครบครันทั้งด้านเนื้อหาและตัวอย่างการประยุกต์ เหมาะสำหรับผู้ที่ต้องการปรับพื้นฐานวิชาสถิติธุรกิจและสถิติวิเคราะห์ทั้งในระดับอาชีวศึกษา ปริญญาตรี และปริญญาโท เพื่อให้เกิดความเข้าใจจนถึงขั้นนำไปใช้ประโยชน์ได้

การวิจัยทางธุรกิจ

ผู้เขียน **รองศาสตราจารย์ ดร. สรชัย พิศาลบุตร**
เหมาะสำหรับ **นักศึกษาระดับปริญญาตรี-โท เนื้อหาครอบคลุมหลักสูตรของมหาวิทยาลัยราชภัฏและมหาวิทยาลัยเทคโนโลยีราชมงคล**
ขนาด 7" x 10" **กระดาษปอนด์**
ราคา 180 บาท **จำนวนหน้า 228 หน้า**

ตำรา **การวิจัยทางธุรกิจ** เล่มนี้ให้ความรู้เกี่ยวกับวิธีการทำวิจัยทางธุรกิจที่ทันสมัยและครบถ้วน เนื้อหาภายในเล่มประกอบด้วยเทคนิคและขั้นตอนการทำวิจัยธุรกิจอย่างละเอียด เริ่มตั้งแต่ความรู้พื้นฐานไปจนถึงหลักการและวิธีปฏิบัติจนได้งานวิจัยที่มีความเชื่อถือได้ สามารถนำไปแก้ปัญหาต่าง ๆทางธุรกิจ ได้แก่ ความสำคัญและประโยชน์ของการวิจัยทางธุรกิจ วิธีการทำวิจัยทางธุรกิจ เครื่องมือและวิธีเก็บรวบรวมข้อมูลทางธุรกิจ การวิเคราะห์ข้อมูลเชิงสถิติกับการวิจัยทางธุรกิจ คอมพิวเตอร์กับการวิจัยทางธุรกิจ และการเขียนรายงานการวิจัยและการประเมินผลงานวิจัย นอกจากนี้ ยังมีกิจกรรมท้ายบทให้ฝึกฝนเพื่อเสริมความเข้าใจในการทำวิจัยแต่ละขั้นตอนอีกด้วย ตำราเล่มนี้เรียบเรียงขึ้นด้วยภาษาที่กระชับ เข้าใจง่าย โดยผู้เขียนที่มีประสบการณ์ในการทำงานวิจัยและการสอนวิชาวิธีวิจัยทางธุรกิจในมหาวิทยาลัยมานานปี เหมาะสำหรับนักศึกษาระดับปริญญาตรี ปริญญาโท ตลอดจนผู้ที่ทำงานเกี่ยวข้องกับการทำวิจัยทางธุรกิจทั่วไปที่ต้องการผลงานวิจัยที่มีคุณภาพมาใช้ในการดำเนินธุรกิจขององค์กรให้มีประสิทธิภาพมากยิ่งขึ้น



เทคนิคการวิจัยทางสังคมศาสตร์

ผู้เขียน ศาสตราจารย์ ดร. สิน พันธุ์พินิจ
เหมาะสำหรับ นักศึกษา อาจารย์ นักวิชาการ และนักวิจัยทางสังคมศาสตร์
ขนาด 7" x 10" กระดาษปอนด์
ราคา 340 บาท จำนวนหน้า 504 หน้า

ในปัจจุบันสังคมไทยมีความซับซ้อนมากขึ้น ปัญหาสังคมเพิ่มทวีอันเนื่องมาจากการเปลี่ยนแปลงทางสังคม อาทิ ปัญหาสถาบันครอบครัวล่มสลาย ปัญหาโสเภณีเด็ก ปัญหาประชาชนติดยาเสพติด ปัญหาการฉ้อราษฎร์บังหลวง ฯลฯ ภารกิจที่สำคัญยิ่งประการหนึ่งของครู อาจารย์ และนักวิชาการก็คือ การทำวิจัยเพื่อค้นหาคำตอบของปัญหาต่าง ๆ เพื่อนำไปสู่การพัฒนาและการแก้ไขปัญหาอย่างยั่งยืน

ตำรา **เทคนิคการวิจัยทางสังคมศาสตร์** เล่มนี้ เป็นตำราที่ให้ความรู้ทั้งด้านทฤษฎีและด้านการปฏิบัติในการวิจัยทางสังคมศาสตร์อย่างครบถ้วน เนื้อหาภายในเล่มประกอบด้วย สารัตถะทางการวิจัย ประเภทของการวิจัยทางสังคมศาสตร์ กระบวนการวิจัย การออกแบบการวิจัย ประชากรและกลุ่มตัวอย่าง เครื่องมือที่ใช้ในการวิจัย การทดสอบคุณภาพของเครื่องมือวิจัย การเก็บข้อมูล การวิเคราะห์ข้อมูลเชิงปริมาณและเชิงคุณภาพ การวิจัยเชิงประเมินผล การเขียนโครงการวิจัย การเขียนรายงานและการเผยแพร่ผลการวิจัย รวมถึงกรณีศึกษาการทำวิจัยทางสังคมศาสตร์ นอกจากนี้ ในภาคผนวกท้ายเล่มยังได้รวบรวมเทคนิคการประมวลตัวแปรการวิจัยและตารางสถิติต่าง ๆที่จำเป็นในการทำวิจัยอีกด้วย ดังนั้นตำราเล่มนี้จึงเป็นทั้งตำราและคู่มือการทำงานวิจัยทางสังคมศาสตร์อย่างสมบูรณ์

เทคนิคการวิจัยทางวิทยาศาสตร์

ผู้เขียน ศาสตราจารย์ ดร. สิน พันธุ์พินิจ
เหมาะสำหรับ นักศึกษาระดับปริญญาตรี-โท อาจารย์ นักวิชาการ และนักวิจัยในสายวิทยาศาสตร์
ขนาด 7" x 10" กระดาษปอนด์
ราคา 370 บาท จำนวนหน้า 572 หน้า

ตำรา **เทคนิคการวิจัยทางวิทยาศาสตร์** เล่มนี้ให้ความรู้เกี่ยวกับการวิจัยทางวิทยาศาสตร์อย่างครบถ้วนทั้งด้านแนวคิด ทฤษฎี และการปฏิบัติ เนื้อหาภายในเล่มประกอบด้วย สารัตถะทางการวิจัย สาขาของการวิจัยทางวิทยาศาสตร์ กระบวนการวิจัย การออกแบบการวิจัย ประชากรและกลุ่มตัวอย่าง การวิเคราะห์ข้อมูลด้วยสถิติพรรณนาและสถิติอนุมาน การทดสอบสมมติฐาน หลักการออกแบบการทดลอง แบบแผนการทดลองทางวิทยาศาสตร์ ได้แก่ แบบแผนการทดลองสุ่มสมบูรณ์ แบบแผนการทดลองสุ่มสมบูรณ์ในบล็อก แบบแผนการทดลองลาตินสแควร์ แบบแผนการทดลองแฟกทอเรียล แบบแผนการทดลองสปลิตพล็อต แบบแผนการทดลองประเภทบล็อกไม่สมบูรณ์ และการประยุกต์เทคโนโลยีสารสนเทศในการวิจัย เนื้อหาอธิบายด้วยภาษาที่เข้าใจง่าย สามารถศึกษาได้ด้วยตนเอง นอกจากนี้ ในภาคผนวกยังให้ความรู้เรื่องการเขียนข้อเสนอโครงการเพื่อขอทุนวิจัยและการเขียนบรรณานุกรมในรายงานการวิจัยทางวิทยาศาสตร์อีกด้วย นับว่า **เทคนิคการวิจัยทางวิทยาศาสตร์** เป็นตำราและคู่มือการทำงานวิจัยสายวิทยาศาสตร์ที่สมบูรณ์ ให้ทั้งพื้นฐานไปจนถึงเทคนิคต่าง ๆ จึงเหมาะอย่างยิ่งสำหรับผู้ที่ต้องทำการวิจัยทางวิทยาศาสตร์และเทคโนโลยี นักศึกษาในทุกระดับทั้งปริญญาตรีและปริญญาโท อาจารย์ นักวิชาการ ตลอดจนนักวิจัย

หนังสือ **สถิติในงานวิจัย เลือกใช้อย่างไรให้เหมาะสม** เล่มนี้ เขียนจากประสบการณ์ยาวนานกว่า 20 ปีในด้านการทำวิจัยของผู้เขียน ทั้งในฐานะผู้ทำวิจัย ผู้สอนวิชาสถิติ ผู้ให้คำปรึกษาด้านการใช้สถิติ และผู้ประเมินโครงการและบทความวิจัย รวบรวมปัญหาที่มักพบในงานวิจัย เช่น ข้อผิดพลาด ความไม่เหมาะสม ความเข้าใจผิดต่างๆ พร้อมอธิบายเหตุผลและแนะนำวิธีการแก้ไขให้ถูกต้อง เรียบเรียงเนื้อหาตามลำดับขั้นตอนในการทำวิจัย ตั้งแต่การออกแบบงานวิจัย การวิเคราะห์ข้อมูล ไปจนถึงการสรุปแปลผลและนำเสนอผลวิจัย อ่านเข้าใจง่าย สามารถศึกษาได้ด้วยตนเอง เหมาะสำหรับเป็นคู่มือให้ผู้ที่ต้องทำวิจัย ซึ่งไม่เฉพาะแต่ในสาขาวิทยาศาสตร์สุขภาพเท่านั้น แต่ยังมีประโยชน์อย่างยิ่งกับการทำวิจัยสาขาอื่นๆที่เป็นงานวิจัยเชิงปริมาณอีกด้วย

เกี่ยวกับผู้เขียน

รองศาสตราจารย์อธุณ จิรวัฒน์กุล

จบการศึกษาปริญญาตรีสาขาอาชีวอนามัย และปริญญาโทสาขาวิชาชีวสถิติ จากคณะสาธารณสุขศาสตร์ มหาวิทยาลัยมหิดล และ Master of Science (Clinical Epidemiology) จาก McMaster University ประเทศแคนาดา ดำรงตำแหน่งรองศาสตราจารย์ อดีตอาจารย์ประจำภาควิชาสถิติและประชากรศาสตร์ คณะสาธารณสุขศาสตร์ มหาวิทยาลัยขอนแก่น ปัจจุบันเป็นอาจารย์พิเศษในสถาบันการศึกษาหลายแห่ง และวิทยากรด้านการวิจัยและสถิติของหน่วยงานในกระทรวงสาธารณสุข

สถิติในงานวิจัย
เลือกใช้อย่างไรให้เหมาะสม

หมวด
การวิจัย/สถิติศาสตร์

ISBN 978-616-7136-71-4

9 786167 136714

ราคา 130 บาท

www.wphat.com