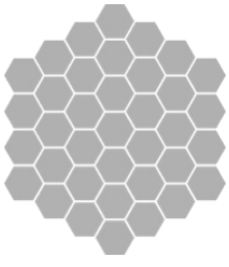
ISBN : 978-616-565-574-3

ตำรา

วิทยาการระบาดและชีวสถิติ ในงานวิจัยวิทยาศาสตร์สุขภาพ

Epidemiology and Biostatistics in Health Science Research



คณะแพทยศาสตร์ มหาวิทยาลัยมหาสารคาม
นิรันดร์ อินทรัตน์

วิทยาการระบาดและชีวสถิติ ในงานวิจัยวิทยาศาสตร์สุขภาพ

Epidemiology and Biostatistics in Health Science Research

ISBN	978-616-565-574-3
ผู้แต่ง	นิรันดร์ อินทรัตน์
พิมพ์ครั้งที่	1
จำนวนพิมพ์	400 เล่ม
พิมพ์ที่	หจก.โรงพิมพ์คลังนานาวิทยา 232/199 หมู่ 6 ถ.ศรีจันทร์ ต.ในเมือง อ.เมือง จ.ขอนแก่น 4000 โทร. 0-4346-6444, 0-4346-6860-61 โทรสาร 0-4346-6863 klungpress@hotmail.com www.klungnana.com

คำนำ

ตำราเล่มนี้ผู้เขียนได้ตั้งใจรวบรวมข้อมูลเพื่อนำมาใช้เป็นส่วนหนึ่งของวิชาชีวสถิติและระบาดวิทยาประยุกต์ของนิสิตระดับบัณฑิตศึกษา สาขาวิชาวิทยาศาสตร์สุขภาพและสาขาวิชาระบาดวิทยาและการวิจัยของนิสิตสาขาแพทยศาสตรบัณฑิต ซึ่งจะมีเนื้อหาเกี่ยวกับวิทยาการทางระบาดและชีวสถิติเบื้องต้น ทั้งนี้เนื้อหาในตำราเล่มนี้ได้เน้นเกี่ยวกับรูปแบบการศึกษาและอคติที่เกิดขึ้นในขั้นตอนการทำวิจัย ส่วนชีวสถิติเป็นการเลือกใช้สถิติและการวิเคราะห์ข้อมูล ทั้งนี้การฝึกปฏิบัติวิเคราะห์ข้อมูล สามารถดาวน์โหลดชุดข้อมูลและชุดคำสั่งจากลิ้งนี้ https://doi.org/10.7910/DVN/UK0PZU หรือติดต่อ nirun.i@msu.ac.th

ผู้เขียนหวังว่าตำราเล่มนี้จะมีประโยชน์ต่อตัวผู้ศึกษาและผู้สนใจเกี่ยวกับวิทยาการทางระบาดและชีวสถิติเบื้องต้น หากมีข้อเสนอแนะประการใด ผู้เขียนขอน้อมรับไว้ด้วยความขอบพระคุณยิ่ง

นิรันดร์ อินทรัตน์

กิตติกรรมประกาศ

ผู้เขียนขอขอบพระคุณผู้ประพันธ์ทั้งหนังสือและงานวิจัยที่ได้นำมาอ้างอิงในตำราเล่มนี้ ขอขอบพระคุณ คณะแพทยศาสตร์ มหาวิทยาลัยมหาสารคาม ที่ได้สนับสนุนในการจัดทำตำรา

ผู้เขียนขอสำนึกในพระคุณของครูบาอาจารย์ ซึ่งเป็นผู้ให้โดยไม่หวังสิ่งตอบแทน ขอขอบพระคุณ ศาสตราจารย์นายแพทย์สมรัตน์ จารุลักษณานันท์ เป็นผู้ให้โอกาส จุดประกายประสบการณ์ทางด้านงานวิจัยและให้โอกาสผู้เขียนได้เข้ามาทำงานด้านนี้ ขอขอบพระคุณท่านอาจารย์ ศาสตราจารย์ ดร.มาลินี เหล่าไพบูลย์ และ ศาสตราจารย์นายแพทย์ภิเศก ลุมพิกานนท์ ได้ให้โอกาสเข้ามาทำงานและเรียนรู้งานด้านวิจัย ให้คำแนะนำหลายๆ ด้านทั้งการเรียน การทำงาน และขอขอบพระคุณ ศาสตราจารย์นายแพทย์วีระศักดิ์ จงสู่วิวัฒวงศ์ ที่ได้ให้โอกาสเข้ามาศึกษาในระดับปริญญาเอก สาขาวิชาระบาดวิทยา ท่านอาจารย์เป็นแบบอย่างที่ดีแก่ลูกศิษย์ทั้งด้านการงาน การดำเนินชีวิตและครอบครัว

สุดท้ายขอขอบพระคุณบิดา มารดา ซึ่งคอยเป็นกำลังใจในการดำเนินชีวิต ขอบพระคุณพี่สาว ที่คอยดูแลสนับสนุนน้องชายมาจนถึงปัจจุบันและขอบคุณภรรยาและลูกสาว ที่คอยเป็นกำลังใจและสนับสนุนในทุกๆ เรื่อง

สารบัญ

หน้า

คำนำ .....ก
กิตติกรรมประกาศ .....ข
สารบัญ .....ค
บทที่ 1 วิทยาการทางระบาดและชีวสถิติ 1
บทที่ 2 รูปแบบการศึกษาเชิงสังเกตการณ์ 5
บทที่ 3 รูปแบบการศึกษาเชิงทดลอง 15
บทที่ 4 อคติในงานวิจัย 27
บทที่ 5 ชีวสถิติเบื้องต้น 39
บทที่ 6 การใช้โปรแกรม R และชุดคำสั่งเบื้องต้น 47
บทที่ 7 การวิเคราะห์ข้อมูลต่อเนื่อง 59
บทที่ 8 การวิเคราะห์ข้อมูลจำแนกประเภท 81
บทที่ 9 การวิเคราะห์สหสัมพันธ์ของเพียร์สันและการถดถอยเชิงเส้น 99
บทที่ 10 การวิเคราะห์การถดถอยลอจิสติก 127
บทที่ 11 การวิเคราะห์อัตราการรอดชีพ..... 145
ดัชนี 161
เฉลย 167

บทที่ 1
วิทยาการทางระบาดและชีวสถิติ

วิทยาการทางระบาดเป็นศาสตร์ที่ศึกษาเกี่ยวกับลักษณะการเกิดโรค ความรุนแรงของการเกิดโรค ส่วนชีวสถิติเป็นศาสตร์ที่ศึกษาเกี่ยวกับกระบวนการทางสถิติทางด้านชีวภาพ เช่น ด้านการแพทย์ ด้านสาธารณสุข เป็นต้น วัตถุประสงค์ของบทนี้เพื่อให้ผู้ศึกษาได้เรียนรู้ภาพรวมของวิทยาการทางระบาดและชีวสถิติ

1.1 วิทยาการทางระบาด

วิทยาการระบาด หรือ ระบาดวิทยา ศัพท์ทางภาษาอังกฤษเรียกว่า “epidemiology”[1-3] มีรากศัพท์มาจากภาษากรีกคือ “epi” หมายถึง “ขึ้นอยู่กับ” “demos” หมายถึง “คน” และ “logos” หมายถึง “เป็นการศึกษาของ” เมื่อรวมรากศัพท์เข้าด้วยกันแล้วมีความหมายว่า การศึกษาที่เกี่ยวข้องกับคน เพื่ออธิบายการกระจายของโรค (Distribution) และปัจจัยที่เกี่ยวข้องกับการเกิดโรคหรือปัจจัยที่เกี่ยวข้องกับการกระจายของการเกิดโรค (Determinants)[4-6]

การกระจายของโรคในแง่มุมของวิทยาการทางระบาดจะมุ่งเน้นอธิบายถึง ความถี่ของการเกิดโรค (Frequency) เช่น เกิดขึ้นบ่อยไหม เกิดขึ้นเวลาไหน (Time) สถานที่ใด (Place) และเกิดขึ้นกับคนกลุ่มใด (Person) ลักษะณะดังกล่าวเรียกว่าเป็น การศึกษาทางระบาดวิทยาเชิงพรรณนา (Descriptive epidemiology) [7, 8]

นอกจากนั้นแล้วการศึกษาปัจจัยที่เกี่ยวข้องกับการเกิดโรคหรือสาเหตุที่ทำให้เกิดโรคต่างๆ นั้นจะเรียกว่าเป็นการศึกษาทางด้านระบาดวิทยาเชิงวิเคราะห์ (Analytical epidemiology) ซึ่งรูปแบบการศึกษาหลักๆ จะแบ่งเป็นการศึกษาเชิงสังเกต (Observational study) และการศึกษาเชิงทดลอง (Experimental study) [9-11]

การศึกษาเชิงสังเกตประกอบไปด้วย การศึกษาภาคตัดขวาง (Cross-sectional study) การศึกษาไปข้างหน้า (Cohort study) และการศึกษาย้อนหลัง (Case-control study) ส่วนการศึกษาเชิงทดลองประกอบไปด้วย การศึกษาเชิงสุ่มแบบมีกลุ่มควบคุม (Randomized controlled trial) เป็นต้น [12, 13]

จากที่ได้กล่าวมาเบื้องต้นวิทยาการทางระบาดเป็นวิชาที่มุ่งเน้นอธิบายถึงการกระจายของโรคประกอบไปด้วย ความถี่และรูปแบบของการกระจาย (เวลา สถานที่ และคน) นอกจากนั้นแล้ว

ยังพยายามอธิบายถึงปัจจัยที่มีความสัมพันธ์ต่อการเกิดโรคนั้นด้วย ทั้งนี้จะใช้ศาสตร์ทางด้านชีวสถิติมาช่วยอธิบายเหตุการณ์การเกิดโรคหรือทดสอบความสัมพันธ์ของปัจจัยต่างๆ ให้เข้าใจมากยิ่งขึ้น ทั้งนี้ ผู้ศึกษาจะได้ศึกษาวิทยาการทางระบาดเบื้องต้นในบทที่ 2 ถึง บทที่ 4

1.2 ชีวสถิติ

ชีวสถิติเป็นการประยุกต์ใช้ทฤษฎีหรือแนวทางของวิชาสถิติ ในงานวิจัยทางด้านสุขภาพ เช่น ด้านการแพทย์ ด้านสาธารณสุข ด้านเภสัชศาสตร์ หรือด้านระบบสุขภาพ เป็นต้น เมื่อเรากำหนดหัวข้องานวิจัย ทำวิจัย เก็บข้อมูล เราจะใช้ความรู้ทางด้านชีวสถิติเป็นเครื่องมือในการวิเคราะห์ผลการศึกษาและสรุปผลการทดลองหรือสมมุติฐานงานวิจัยของเราว่าเป็นอย่างไร โดยกระบวนการทางชีวสถิติจะถูกนำมาใช้ตั้งแต่เริ่มต้น ออกแบบงานวิจัย การคำนวณขนาดตัวอย่าง การสุ่มตัวอย่าง การเก็บข้อมูล การวิเคราะห์ข้อมูล การรายงานผลการวิเคราะห์ข้อมูลและการสรุปผล[11,12]

ในการวิเคราะห์ข้อมูล จะทำการพรรณนาข้อมูลโดยใช้สถิติพรรณนา (Descriptive statistics)[13-15] เช่น ความถี่ ร้อยละ หรือ ค่าเฉลี่ย เป็นต้น จากนั้นจะทำการอนุมานข้อมูลที่ได้ (ตัวอย่าง) ไปสู่ข้อมูลประชากรที่ศึกษาโดยใช้สถิติเชิงอนุมาน (Inferential statistics)[16-18] จะได้ศึกษาชีวสถิติใน บทที่ 5 ถึง บทที่ 11

1.3 บทสรุป

วิทยาการระบาดเป็นการศึกษาที่เกี่ยวข้องกับการระบาดของโรคที่อธิบายถึง ความถี่ สถานที่เกิด เวลาที่เกิด ขั้นตอนต่อมาอาจทำการทดสอบปัจจัยที่มีความสัมพันธ์ต่อการเกิดโรค โดยใช้ความรู้ทางชีวสถิติมาเป็นเครื่องมือช่วยสรุปผลของการระบาดของโรค

1.4 แบบฝึกหัด

1. วิทยาการระบาดเป็นการศึกษาอะไร
2. การศึกษาเชิงสังเกตประกอบไปด้วยอะไรบ้าง
3. วิทยาการระบาดมุ่งเน้นอธิบายถึงอะไร
4. กระบวนการทางชีวสถิติถูกนำมาใช้ในงานวิจัยขั้นตอนไหนบ้าง
5. การอนุมานค่าสถิติที่ได้จากตัวอย่างไปสู่ข้อมูลประชากรที่ศึกษาเรียกว่าสถิติอะไร

1.5 เอกสารอ้างอิง

1. Manja V, Lakshminrusimha S. Epidemiology and Clinical Research Design, Part 1: Study Types. Neoreviews. 2014;15(12):e558-e69.
2. Bensyl DM, King ME, Greiner A. Applied Epidemiology Training Needs for the Modern Epidemiologist. Am J Epidemiol. 2019;188(5):830-5.
3. Brownson RC, Samet JM, Bensyl DM. Applied epidemiology and public health: are we training the future generations appropriately? Ann Epidemiol. 2017;27(2):77-82.
4. Bensyl DM, King ME, Greiner A. Applied Epidemiology Training Needs for the Modern Epidemiologist. Am J Epidemiol. 2019;188(5):830-5.
5. Levin KA. Study design III: Cross-sectional studies. Evid Based Dent. 2006;7(1):24-5.
6. Fosgate GT, Cohent ND. Epidemiological study design and the advancement of equine health. Equine Vet J. 2008;40(7):693-700.
7. Ranganathan P, Aggarwal R. Study designs: Part 3 - Analytical observational studies. Perspect Clin Res. 2019;10(2):91-4.
8. Källestål C, Stenlund H. Different analytical approaches in an experimental cohort study on preventive measures for caries in adolescents. A comparison between incidence density and increment analysis. Caries Res. 2003;37(1):44-50.
9. Lewallen S, Courtright P. Epidemiology in practice: case-control studies. Community Eye Health. 1998;11(28):57-8.
10. Song JW, Chung KC. Observational studies: cohort and case-control studies. Plast Reconstr Surg. 2010;126(6):2234-42.
11. Dunn WR, Lyman S, Marx R. Research methodology. Arthroscopy. 2003;19(8):870-3.
12. Grossoehme DH. Overview of qualitative research. J Health Care Chaplain. 2014;20(3):109-22.
13. Spriestersbach A, Röhrig B, du Prel JB, Gerhold-Ay A, Blettner M. Descriptive statistics: the specification of statistical measures and their presentation in tables and graphs. Part 7 of a series on evaluation of scientific publications. Dtsch Arztebl Int. 2009;106(36):578-83.

14. Nick TG. Descriptive statistics. Methods Mol Biol. 2007;404:33-52.
15. Kaliyadan F, Kulkarni V. Types of Variables, Descriptive Statistics, and Sample Size. Indian Dermatol Online J. 2019;10(1):82-6.
16. Allua S, Thompson CB. Inferential statistics. Air Med J. 2009;28(4):168-71.
17. Guetterman TC. Basics of statistics for primary care research. Fam Med Community Health. 2019;7(2).
18. Bradley MT, Brand A. Accuracy when inferential statistics are used as measurement tools. BMC Res Notes. 2016;9:241.

บทที่ 2

รูปแบบการศึกษาเชิงสังเกตการณ์

จากที่ได้กล่าวไว้ในบทที่ 1 รูปแบบงานวิจัยมี 2 แบบคือ การศึกษาเชิงสังเกตการณ์ (Observational study design) และการศึกษาเชิงทดลอง (Experimental study design) ในบทนี้เราจะได้ศึกษาถึงรูปแบบการศึกษาเชิงสังเกตการณ์ประกอบไปด้วย การศึกษาแบบตัดขวาง (Cross-sectional study) การศึกษาแบบย้อนหลัง (Retrospective study) และการศึกษาไปข้างหน้า (Cohort study)[1]

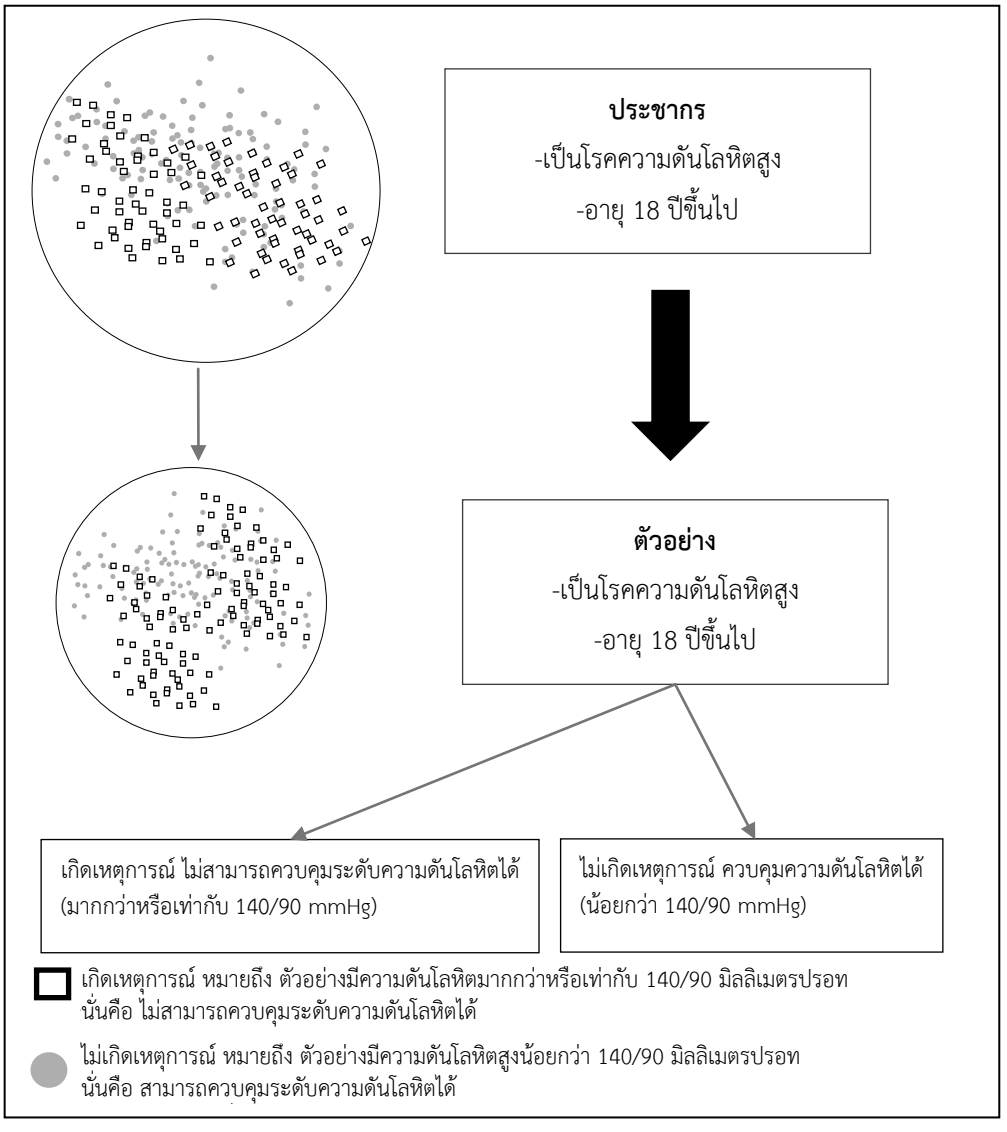
2.1 การศึกษาแบบตัดขวาง

การศึกษาแบบตัดขวาง[2] เป็นรูปแบบการศึกษาที่มีการเก็บข้อมูล ณ ปัจจุบัน โดยส่วนใหญ่รูปแบบการศึกษานี้จะใช้ในงานวิจัยเชิงสำรวจ เพื่อศึกษาถึงความชุกของการเกิดโรค อย่างไรก็ตามอาจจะเป็นการศึกษาเพื่อวิเคราะห์หาความเสี่ยงสัมพัทธ์ของการเกิดโรคได้ ภาพรวมของรูปแบบงานวิจัยนี้ดังรูปที่ 2.1

รูปที่ 2.1 รูปแบบการศึกษาภาคตัดขวาง

ตัวอย่างที่ 2.1 การศึกษาความชุกและปัจจัยที่มีความสัมพันธ์กับการไม่สามารถควบคุมความดันโลหิตได้ในกลุ่มที่เป็นโรคความดันโลหิตสูงที่อาศัยอยู่ในเขตชนบทภาคกลางของประเทศไทย[3] การศึกษานี้มีวัตถุประสงค์เพื่อศึกษาความชุกของการไม่สามารถควบคุมความดันโลหิตได้ในกลุ่มที่เป็นโรคความดันโลหิตสูง โดยมีรูปแบบงานวิจัยเป็นแบบภาคตัดขวาง ทำการ

เก็บข้อมูลในเขตชนบทจำนวน 2 อำเภอ โดยวิธีการสัมภาษณ์กลุ่มตัวอย่างที่มีอายุ 18 ปีขึ้นไป จำนวนทั้งหมด 406 ราย จากงานวิจัยเรื่องนี้พบว่ามีความชุกของการไม่สามารถควบคุมความดันโลหิตได้คิดเป็นร้อยละ 54.4 เมื่อทำการทดสอบปัจจัยเสี่ยงพบว่า คนที่มีพฤติกรรมเนือยนิ่งมากกว่า 5 ชั่วโมงต่อวันมีความเสี่ยงต่อการไม่สามารถควบคุมความดันโลหิตได้จำนวน 2 เท่าเมื่อเทียบกับคนที่มีพฤติกรรมเนือยนิ่งน้อยกว่า 5 ชั่วโมงต่อวัน จากตัวอย่างดังกล่าวจะเห็นได้ว่า วัตถุประสงค์หลักของการวิจัยเป็นการรายงานความชุกของการเกิดโรค และวัตถุประสงค์รองเป็นการทดสอบปัจจัยเสี่ยงของการเกิดโรค การเก็บข้อมูลเริ่มจากกำหนดกลุ่มประชากรและสุ่มตัวอย่างจากประชากรที่กำหนดดังรูปที่ 2.2



รูปที่ 2.2 ตัวอย่างการศึกษาแบบภาคตัดขวาง

ตัวอย่างที่ 2.2 การศึกษาเรื่องความสัมพันธ์ของการดื่มนมแม่กับการเป็นโรคอ้วนในเด็กอายุ 5 และ 6 ปี ในประเทศเยอรมัน[4] รูปแบบการศึกษาเป็นแบบภาคตัดขวาง มีเด็กที่เข้าร่วมการศึกษาทั้งหมดจำนวน 9,357 คน ผลการศึกษาพบว่าความชุกของการเป็นโรคอ้วน ในกลุ่มที่ไม่เคยดื่มนมแม่คิดเป็น 4.5% และกลุ่มที่ดื่มนมแม่เป็นโรคอ้วน 2.8% เมื่อทำการทดสอบปัจจัยเสี่ยงของการเป็นโรคอ้วนและควบคุมปัจจัยกวนต่างๆ แล้วพบว่า การดื่มนมแม่สามารถลดความเสี่ยงของการเป็นโรคอ้วนได้ (ค่าความเสี่ยงสัมพัทธ์ 0.75, ช่วงเชื่อมั่นที่ 95%: 0.57, 0.98) จากตัวอย่างนี้จะเห็นได้ว่า นักวิจัยทำการศึกษาความชุกของการเกิดโรคก่อน จากนั้นเป็นการศึกษาความสัมพันธ์ของปัจจัยที่สนใจ ลักษณะรูปแบบการศึกษาดังกล่าวการคำนวณขนาดตัวอย่างจะทำการคำนวณตามวัตถุประสงค์หลักนั้นคือ การศึกษาความชุกของการเกิดโรค ไม่ได้คำนวณเพื่อทดสอบทางสถิติในวัตถุประสงค์รอง (การทดสอบปัจจัยเสี่ยง) ดังนั้นบางตัวแปรที่ทดสอบอาจจะไม่สามารถสนับสนุนอำนาจการทดสอบทางสถิติ[5-7] (Power of analysis) ได้ทั้งหมด ทั้งนี้ทั้งนั้นรูปแบบการศึกษาภาคตัดขวางมีทั้งข้อดีและข้อจำกัด[8] ดังนี้

ข้อดี

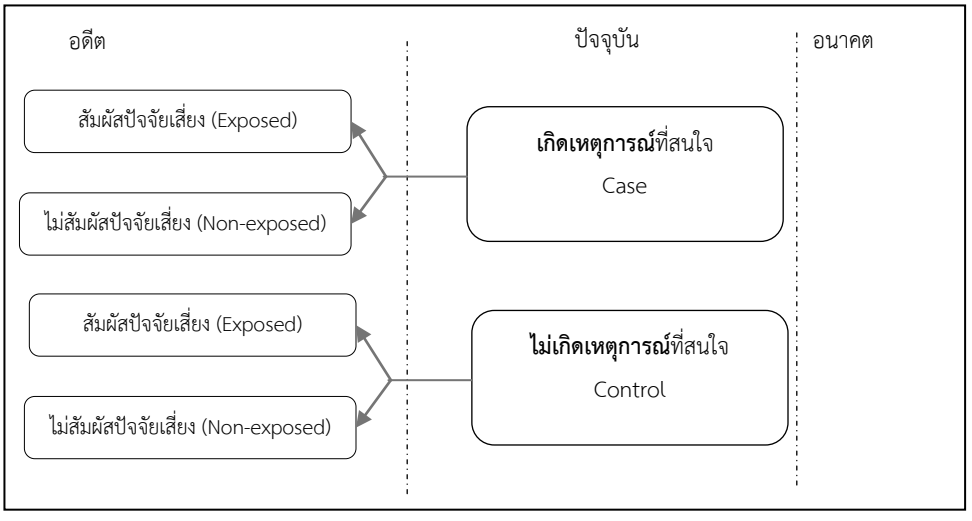
1. ใช้งบประมาณในการวิจัยน้อยและสามารถเก็บข้อมูลได้รวดเร็ว
2. สามารถประมาณค่าความชุกของประชากรได้
3. เหมาะกับการวางแผนทางด้านสาธารณสุขและการสร้างสมมุติฐานงานวิจัย
4. ไม่มีการสูญหายของข้อมูล (Loss to follow-up)

ข้อจำกัด

1. ในการทดสอบความสัมพันธ์อาจจะไม่สามารถสรุปความสัมพันธ์ในเชิงเหตุ-ผล (Causality)
2. เป็นการศึกษาเฉพาะช่วงเวลาหนึ่งเท่านั้นช่วงเวลาเปลี่ยนไปผลการศึกษาที่ได้อาจจะไม่เหมือนเดิม
3. อาจจะเกิดอคติของเนย์แมน (Neyman's bias) หรือความชุก-อุบัติการณ์ (Prevalence – incidence bias) ได้ [9-11]

2.2 การศึกษาแบบย้อนหลัง (Retrospective study)

การศึกษาแบบย้อนหลัง อาจจะคุ้นเคยกันในชื่อ Case-control study[12] หรือ การศึกษาย้อนหลังแบบมีกลุ่มควบคุม เป็นการศึกษาที่เริ่มจากการคนที่เกิดเหตุการณ์ (Cases) จากนั้นจึงค้นหาคนที่ไม่เกิดเหตุการณ์หรือเรียกว่า กลุ่มควบคุม (Controls) และทำการเปรียบเทียบหรือประเมินการสัมผัสปัจจัยเสี่ยงในอดีตที่ผ่านมา (History exposure) ของทั้งสองกลุ่ม[13] ทั้งนี้ การเลือกคนที่เกิดเหตุการณ์และกลุ่มควบคุมมีความสำคัญเป็นอย่างมากเพราะอาจจะเกิดอคติในงานวิจัยต่างๆ ขึ้นได้ (จะได้ศึกษาในบทที่ 4) ผังของการศึกษานี้แสดงในรูปที่ 2.3



รูปที่ 2.3 การศึกษาย้อนหลังแบบมีกลุ่มควบคุม

การเลือก คนที่เกิดเหตุการณ์หรือ Case โดยจะเลือกจากคนที่เกิดเหตุการณ์ที่สนใจศึกษา เช่น การศึกษาความสัมพันธ์ระหว่างการรับประทานอาหารเค็มกับการเป็นโรคความดันโลหิตสูง ในตัวอย่างนี้ การรับประทานอาหารเค็ม (รับประทาน, ไม่รับประทาน) เป็นปัจจัยที่ต้องการศึกษา และการเป็นโรคความดันโลหิตสูง (เป็น, ไม่เป็น) เป็นเหตุการณ์ที่สนใจศึกษา ดังนั้นการเลือกคนที่เกิดเหตุการณ์จะเริ่มจากเลือกคนที่เป็นโรคความดันโลหิตสูง จากนั้นจึงเริ่มเลือกคนที่ไม่เกิดเหตุการณ์

การเลือกคนที่ไม่เกิดเหตุการณ์ (Control) จะเลือกจากกลุ่มคนที่ไม่เกิดเหตุการณ์หรือไม่เกิดโรค ที่มีลักษณะใกล้เคียงหรือโอกาสการสัมผัสปัจจัยเสี่ยงพอๆ กับกลุ่มคนที่เกิดโรคทั้งนี้ การเลือกกลุ่มควบคุมสามารถเลือกได้จากแหล่งต่างๆ ดังนี้ [14]

เลือกจากในโรงพยาบาลหรือจากคลินิก (Hospital controls) เลือกคนที่ไม่เกิดโรคที่มาโรงพยาบาลเดียวกันคนที่เกิดโรค วิธีการนี้ง่ายต่อการเก็บข้อมูลและมีคุณภาพของข้อมูลค่อนข้างดีเพราะได้กลุ่มเปรียบเทียบที่มีลักษณะคล้ายๆ กับคนที่เกิดโรค

เลือกจากประชากรทั่วไป (Population controls) เลือกกลุ่มคนไม่เกิดโรคจากประชากรทั่วไป โดยมีเกณฑ์การนำเข้าเหมือนกับกลุ่มคนไม่เกิดโรคและพิจารณาถึงโอกาสการสัมผัสปัจจัยเสี่ยงพอๆ กันกับกลุ่มคนที่เกิดโรค

เลือกจากเพื่อนบ้าน (Neighborhoods controls) การเลือกดังกล่าวจะเลือกกลุ่มคนที่ไม่เกิดโรคที่เป็นเพื่อนบ้านหรือมีบ้านติดกันจะมีบ้านติดกันฝั่งซ้าย ฝั่งขวา หลังบ้านหรือหน้าบ้านของคนที่เกิดโรคก็ได้ แต่ต้องกำหนดนิยามการเลือกว่าจะให้เลือกบ้านที่อยู่ในตำแหน่งไหนและควรเป็นมาตรฐานเดียวกัน

เลือกจากคนที่เป็นเพื่อน (Friend controls) การเลือกคนที่ไม่เป็นโรคที่เป็นเพื่อนของคนที่เกิดโรคจะมีข้อดีในแง่มุมของการลักษณะทางด้านเศรษฐานะทางสังคม หรือมีพฤติกรรมการดำเนินชีวิตที่คล้ายๆ กัน

ตัวอย่างที่ 2.3 การศึกษาความสัมพันธ์ของการดื่มเครื่องดื่มแอลกอฮอล์และการเกิดกล้ามเนื้อหัวใจตาย (Myocardial infarction: MI) ในประเทศเซอเบีย[15] ทำการศึกษาโดย Milena Ilic และคณะในปี 2018 จากผลการศึกษาพบว่า การดื่มแอลกอฮอล์อย่างหนักภายในรวดเดียว (Binge drinking) มีความเสี่ยงต่อการเกิดโรคกล้ามเนื้อหัวใจตายได้ เมื่อเทียบกับคนที่ไม่ดื่ม ตัวอย่างการเลือกกลุ่มเป็นโรคและการเลือกกลุ่มควบคุมดังนี้

การเลือกกลุ่มคนที่เกิดโรคกล้ามเนื้อหัวใจตาย คือกลุ่มคนที่มารับการรักษาที่โรคที่โรงพยาบาลและได้รับการวินิจฉัยโรคใหม่ ตามเกณฑ์และแนวทางปฏิบัติของ the European Society of Cardiology/American College of Cardiology

การเลือกกลุ่มควบคุม (กลุ่มคนที่ไม่เกิดโรคกล้ามเนื้อหัวใจตาย) เลือกจากคนที่มาโรงพยาบาลเดียวกับคนที่เป็นโรคแต่ทั้งนี้ไม่มีประวัติเป็นโรคกล้ามเนื้อหัวใจตายมาก่อน ไม่มีประวัติเป็นโรคหัวใจและหลอดเลือด (Cardiovascular disease) เป็นผู้ป่วยที่มารักษาด้วยโรคที่ไม่เกี่ยวข้องกับโรคกล้ามเนื้อหัวใจตายเช่น มารักษาด้วยโรคกระดูกและข้อหักหรือเคลื่อน เป็นต้น

จากตัวอย่างที่ 2.3 ได้แสดงให้เห็นถึงการเลือกกลุ่มคนที่เกิดเหตุการณ์ก่อน จากนั้นจึงเลือกกลุ่มคนที่ไม่เกิดเหตุการณ์ รูปแบบการศึกษานี้มีข้อดีและข้อจำกัด[14,16] ดังนี้

ข้อดี

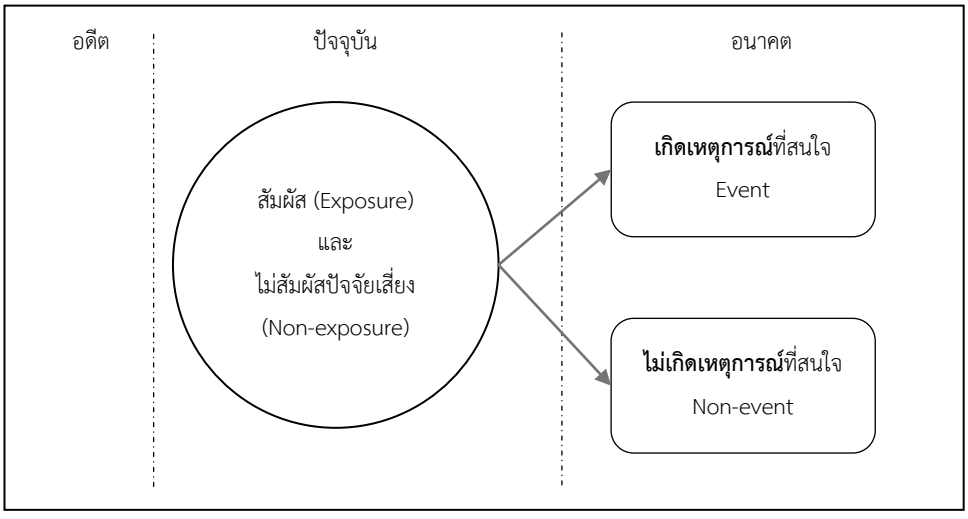
1. สามารถเก็บข้อมูลได้อย่างรวดเร็ว
2. สามารถใช้ในการศึกษากับโรคที่เกิดขึ้นน้อย (Rare disease) หรือโรคที่มีระยะเวลาการก่อโรคค่อนข้างนาน เช่น โรคมะเร็ง เป็นต้น
3. สามารถศึกษาหลายปัจจัยเสี่ยงได้ (Multiple exposures)
4. ในกรณีที่ทำการศึกษาโรคที่เกิดขึ้นน้อยนักวิจัยสามารถเพิ่มจำนวนกลุ่มควบคุม (Multiple control per case) ได้ เพื่อที่จะเพิ่มอำนาจการทดสอบทางสถิติให้มีมากขึ้นหรือเรียกว่า Matched case-control study เช่นการเลือกแบบ 1 ต่อ 4 โดยเลือกจับคู่ตามเพศ (Matching) หมายความว่า กลุ่มคนที่เป็นโรค 1 คน เป็นเพศชาย จากนั้นจึงเลือกกลุ่มคนที่ไม่เป็นโรคที่เป็นเพศชาย จำนวน 4 คน ทำลักษณะนี้ไปเรื่อยๆ จนได้จำนวนขนาดตัวอย่างที่ต้องการ

ข้อจำกัด

1. มีโอกาสเกิดอคติในการจำการสัมผัสปัจจัยเสี่ยงที่ผ่านมา (Recall bias)
2. เนื่องจากว่าการศึกษา Case-control study เป็นการทดสอบการสัมผัสปัจจัยเสี่ยงในอดีตโดยเริ่มเก็บข้อมูลจากคนที่เกิดโรคในปัจจุบัน ซึ่งไม่สามารถที่จะประเมินได้ว่าการสัมผัสปัจจัยเสี่ยงนั้น เกิดก่อน-เป็นโรค หรือ เกิดหลัง-เป็นโรค ทำให้นิยมใช้ผลการศึกษารูปแบบนี้สร้างสมมุติฐานงานวิจัย เพราะไม่สามารถทดสอบความสัมพันธ์เชิงเหตุ-เชิงผลได้ (Causation)
3. ไม่เหมาะกับการศึกษาที่ต้องการวัดผลลัพธ์ทางการศึกษาหลายผลลัพธ์ (Multiple outcomes) เพราะจะทำให้ค้นหาคนที่เกิดเหตุการณ์ยากและปัจจัยเสี่ยงอาจจะมีหลายตัวแปรทำให้เก็บข้อมูลได้ยาก

2.3 การศึกษาไปข้างหน้า

รูปแบบการศึกษาไปข้างหน้าหรือ Cohort study เป็นการติดตามกลุ่มตัวอย่างไปข้างหน้าตามระยะเวลาที่กำหนดแล้วดูว่ามีการเกิดเหตุการณ์ที่สนใจหรือไม่ ทั้งนี้กลุ่มตัวอย่างที่ติดตามในตอนเริ่มต้นมีทั้งคนที่สัมผัสปัจจัยเสี่ยงและไม่ได้สัมผัสปัจจัยเสี่ยงดังรูปที่ 2.4



รูปที่ 2.4 การศึกษาไปข้างหน้า

ตัวอย่างที่ 2.4 การศึกษาเพื่อหาความสัมพันธ์ของการดื่มกาแฟกับการเป็นโรคมะเร็งเยื่อบุโพรงมดลูกทำการศึกษาโดย Youjin Je และคณะ (2011)[17] ผู้วิจัยได้ทำการติดตามผู้หญิงที่ดื่มกาแฟอายุระหว่าง 34 ถึง 59 ปี จำนวนทั้งหมด 67,470 ราย โดยใช้เวลาติดตามผล 26 ปี พบว่า การดื่มกาแฟมากกว่าหรือเท่ากับ 4 แก้วต่อวัน ช่วยลดความเสี่ยงต่อการเป็นโรคมะเร็งเยื่อบุโพรงมดลูกได้ จากตัวอย่างนี้จะเห็นได้ว่าเป็นการศึกษากลุ่มตัวอย่างที่สัมผัสปัจจัยเสี่ยง คือ การดื่มกาแฟในระดับต่างๆ (น้อยกว่า 1 แก้วต่อวัน, 1 แก้วต่อวัน, 2-3 แก้วต่อวัน, มากกว่าหรือเท่ากับ 4 แก้วต่อวัน) ทั้งนี้การศึกษาไปข้างหน้ามีข้อดีและข้อจำกัด[14, 16, 18] ดังนี้

ข้อดี

1. สามารถระบุความเป็น เชิงเหตุ-เชิงผล ของการเกิดโรคได้ (Causation) เพราะสามารถระบุได้ว่า มีการสัมผัสปัจจัยเสี่ยงก่อนแล้วจึงทำให้เกิดโรค
2. สามารถคำนวณอุบัติการณ์ได้
3. สามารถทดสอบผลลัพธ์ได้หลายอย่างจากการสัมผัสปัจจัยเสี่ยงนั้น

ข้อจำกัด

1. ถ้าเหตุการณ์ที่ศึกษาเกิดขึ้นน้อยอาจจะต้องใช้ขนาดตัวอย่างที่มากในการติดตามผล
2. ใช้งบประมาณในการทำวิจัยมาก
3. ใช้ระยะเวลาในการติดตามผลที่นานและอาจจะเกิดการติดตามผลไม่ได้ (Loss to follow-up)

จากรูปแบบการศึกษาไปข้างหน้า กรณีนักวิจัยได้ทำการเก็บข้อมูลไปจนถึงระยะเวลาหนึ่ง นักวิจัยจะทำการศึกษาโดยการเริ่มจาก คนที่เกิดเหตุการณ์ จากนั้นจึงเลือก กลุ่มคนที่ไม่เกิดเหตุการณ์ ซึ่งข้อมูลเหล่านี้ได้เก็บมาและเก็บไว้ในฐานข้อมูล การศึกษาไปข้างหน้าดังกล่าว รูปแบบการศึกษาลักษณะนี้เรียกว่า Nested case-control study [19,20]

2.4 บทสรุป

รูปแบบงานวิจัยเชิงสังเกตการณ์เป็นรูปแบบที่นักวิจัยเองได้ทำการสังเกต เหตุการณ์การสัมผัสหรือการมีของปัจจัยเสี่ยงต่างๆ ของกลุ่มตัวอย่างแล้วทดสอบปัจจัยเสี่ยงเหล่านั้นมีความสัมพันธ์กับการเกิดเหตุการณ์ที่สนใจหรือไม่ ถ้าจุดเริ่มต้นของการศึกษาอยู่ ณ ปัจจุบัน แล้วทำการเก็บข้อมูล ณ ปัจจุบัน เรียกว่าเป็นการศึกษาแบบ Cross-sectional study ถ้าทำการเริ่มต้นจากการเกิดโรค ณ ปัจจุบัน แล้วสอบถามถึงเหตุการณ์การสัมผัสปัจจัยเสี่ยงในอดีตที่ผ่านมา ลักษณะนี้เรียกว่าเป็นการศึกษาแบบ Case-control study นอกจากนั้นแล้ว ถ้าเริ่มต้นติดตามอาสาสมัครงานวิจัยที่สัมผัสและไม่สัมผัสปัจจัยเสี่ยง แล้วทำการติดตามผลไปข้างหน้า (ไม่ว่าจะเป็นการเริ่มต้นติดตามการสัมผัสปัจจัยเสี่ยงจากอดีต หรือ ปัจจุบัน) จนถึงระยะเวลาที่กำหนด (เกิดโรค, ไม่เกิดโรค) เรียกว่า Cohort study หรือ Prospective cohort study ทั้งนี้รูปแบบการศึกษาแต่ละแบบมีข้อดีและข้อจำกัดที่แตกต่างกัน การจะเลือกใช้รูปแบบใดต้องคำนึงถึงการดำเนินของโรคด้วย

2.5 แบบฝึกหัด

1. ในกรณีที่ผลลัพธ์ที่ศึกษามีการเกิดเหตุการณ์น้อย รูปแบบการศึกษาใดเหมาะสมที่สุด
2. รูปแบบการศึกษาใดไม่สามารถสรุปการเป็นเชิงเหตุ-เชิงผลของการเกิดเหตุการณ์ได้ (Causality)
3. ในรูปแบบการศึกษา Case-control study ถ้ามีการจับคู่เช่น เพศ ในขั้นตอนออกแบบการศึกษาลักษณะดังกล่าวเรียกว่าอะไร
4. ถ้าต้องการสำรวจความชุกของการเกิดโรคความดันโลหิตสูง โดยมีระยะเวลาที่จำกัดจะใช้การศึกษารูปแบบใด
5. รูปแบบการศึกษาใดมีโอกาสขาดการติดต่อของอาสาสมัครงานวิจัยสูงที่สุด

2.6 เอกสารอ้างอิง

1. Gordis L. Epidemiology: Elsevier/Saunders; 2009.
2. Szklo M, Nieto FJ. Epidemiology: Beyond the Basics: Jones & Bartlett Learning; 2014.
3. Meelab S, Bunupuradah I, Suttiruang J, Sakulrojanawong S, Thongkua N, Chantawiboonchai C, et al. Prevalence and associated factors of uncontrolled blood pressure among hypertensive patients in the rural communities in the central areas in Thailand: A cross-sectional study. PLoS One. 2019;14(2):e0212572.
4. von Kries R, Koletzko B, Sauerwald T, von Mutius E, Barnert D, Grunert V, et al. Breast feeding and obesity: cross sectional study. BMJ. 1999;319(7203):147-50.
5. Hickey GL, Grant SW, Dunning J, Siepe M. Statistical primer: sample size and power calculations-why, when and how? Eur J Cardiothorac Surg. 2018;54(1):4-9.
6. Suresh K, Chandrashekara S. Sample size estimation and power analysis for clinical research studies. J Hum Reprod Sci. 2012;5(1):7-13.
7. Anderson SF. Best (but oft forgotten) practices: sample size planning for powerful studies. Am J Clin Nutr. 2019;110(2):280-95.
8. Levin KA. Study design III: Cross-sectional studies. Evid Based Dent. 2006;7(1):24-5.
9. Swanson DM, Anderson CD, Betensky RA. Hypothesis Tests for Neyman's Bias in Case-Control Studies. J Appl Stat. 2018;45(11):1956-77.
10. Jager KJ, Tripepi G, Chesnaye NC, Dekker FW, Zoccali C, Stel VS. Where to look for the most frequent biases? Nephrology (Carlton). 2020;25(6):435-41.
11. Tripepi G, Jager KJ, Dekker FW, Wanner C, Zoccali C. Bias in clinical research. Kidney Int. 2008;73(2):148-53.
12. Lewallen S, Courtright P. Epidemiology in practice: case-control studies. Community Eye Health. 1998;11(28):57-8.
13. Schlesselman J, editor Case-Control Studies: Design, Conduct, Analysis1982.
14. Song JW, Chung KC. Observational studies: cohort and case-control studies. Plast Reconstr Surg. 2010;126(6):2234-42.
15. Ilic M, Grujicic Sipetic S, Ristic B, Ilic I. Myocardial infarction and alcohol consumption: A case-control study. PLoS One. 2018;13(6):e0198129.

16. Mann CJ. Observational research methods. Research design II: cohort, cross sectional, and case-control studies. Emerg Med J. 2003;20(1):54-60.
17. Je Y, Hankinson SE, Tworoger SS, De Vivo I, Giovannucci E. A prospective cohort study of coffee consumption and risk of endometrial cancer over a 26-year follow-up. Cancer Epidemiol Biomarkers Prev. 2011;20(12):2487-95.
18. Whitney JD. Comparative, observational designs: case-control and cohort studies. J Wound Ostomy Continence Nurs. 2000;27(3):191-3.
19. Sedgwick P. Nested case-control studies: advantages and disadvantages. Bmj-British Medical Journal. 2014;348.
20. Ernster VL. Nested Case-Control Studies. Preventive Medicine. 1994;23(5):587-90.

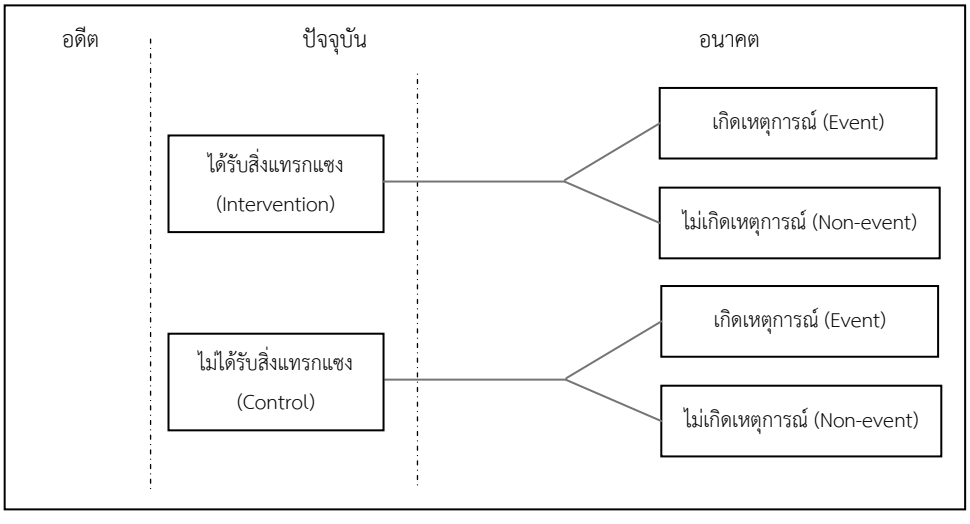
บทที่ 3

รูปแบบการศึกษาเชิงทดลอง

งานวิจัยที่ใช้รูปแบบการศึกษาเชิงทดลอง (Experimental study design) ในบทนี้ประกอบไปด้วยการทดลองเชิงสุ่มโดยมีกลุ่มเปรียบเทียบ (Randomized controlled trial) การทดลองแบบสองระยะไขว้กัน (Crossover design) และการทดลองแบบสี่กลุ่มโซโลมอน (Solomon four-group design)

3.1 การทดลองแบบสุ่มโดยมีกลุ่มเปรียบเทียบ

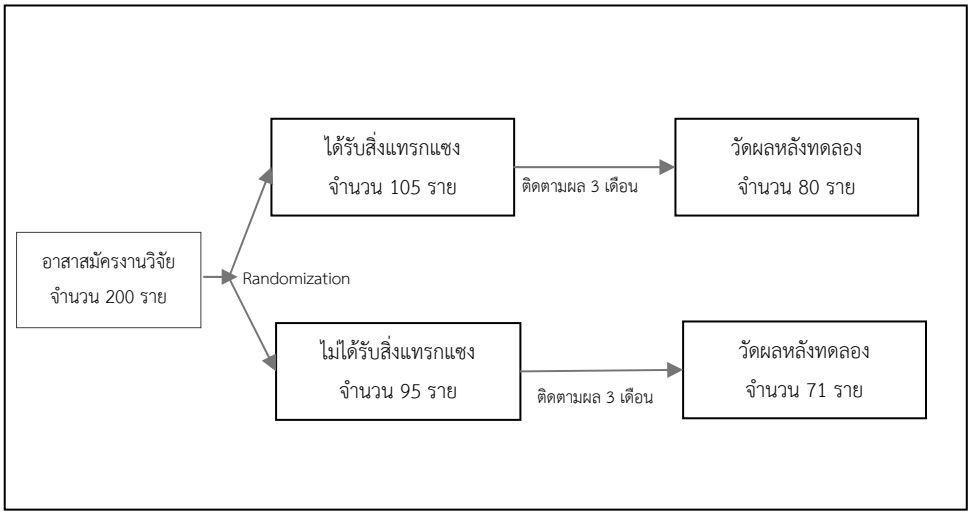
การทดลองแบบสุ่มที่มีกลุ่มเปรียบเทียบ (Randomized controlled trial) เริ่มต้นจากนักวิจัยต้องการศึกษาผลของสิ่งแทรกแซงว่ามีผลต่อผลลัพธ์ที่ต้องการศึกษาหรือไม่ โดยแบ่งออกเป็น 2 กลุ่มคือ กลุ่มที่ได้รับสิ่งแทรกแซง (Intervention group) และกลุ่มควบคุมซึ่งไม่ได้รับสิ่งแทรกแซง (Control group) แผนผังการทดลองดังรูปที่ 3.1



รูปที่ 3.1 การศึกษาแบบสุ่มมีกลุ่มควบคุม

ตัวอย่างที่ 3.1 การทดสอบโปรแกรมการอบรบฝึกสมรรถภาพทางกายในผู้ป่วยหลอดเลือดสมองกึ่งเฉียบพลัน (Subacute stroke) ทำการศึกษาโดย Nave และคณะ (2019)[1] ได้ทำการทดลองในผู้ป่วยจำนวน 200 ราย แบ่งเป็นกลุ่มได้รับสิ่งแทรกแซง (สิ่งแทรกแซงคือการออกกำลัง

กายแบบแอโรบิคมีระบบพยุงตัวขณะฝึกเดินและการฝึกเดินบนลู่จำนวน 20 ครั้งๆ ละ 25 นาที) จำนวน 105 ราย และกลุ่มควบคุม (ได้รับการผ่อนคลายกล้ามเนื้อจำนวน 20 ครั้งๆ ละ 25 นาที) จำนวน 95 ราย ในการศึกษานี้มีการจัดสรรตัวอย่างให้ได้รับสิ่งแทรกแซงและควบคุมเป็นแบบชั้นภูมิ (แบ่งตาม กลุ่มอายุ (≤ 65 ปี, > 65 ปี) คะแนนความสามารถในการเดิน (คะแนน ≤ 3, > 3) และสถานที่ในการทดลอง) โดยมีสัดส่วน 1 ต่อ 1 ผลการทดลองพบว่าค่าเฉลี่ยของความเร็วในการเดินและค่าเฉลี่ยของการดำเนินชีวิตประจำวัน (ดัชนีบาร์เธล: Barthel index) ระหว่างการออกกำลังกายแบบแอโรบิคกับกลุ่มควบคุมไม่มีความแตกต่างกัน ผังการทดลองอย่างง่ายดังรูปที่ 3.2



รูปที่ 3.2 รูปแบบการทดลองแบบสุ่มมีกลุ่มควบคุม

การจัดสรร (Random allocation) สิ่งแทรกแซงและสิ่งควบคุมให้กับอาสาสมัครงานวิจัย สำหรับงานวิจัยรูปแบบนี้ มีวัตถุประสงค์คือเพื่อให้อาสาสมัครงานวิจัยมีโอกาสได้รับสิ่งแทรกแซงหรือสิ่งควบคุมเท่าๆ กัน โดยการจัดสรรสิ่งแทรกแซงเป็น 3 ชนิด คือ 1. การจัดสรรอย่างง่าย (Simple allocation) 2. การจัดสรรแบบบล็อก (Block allocation) และ 3. การจัดสรรแบบชั้นภูมิ (Stratified allocation)

3.1.1 การจัดสรรอย่างง่าย

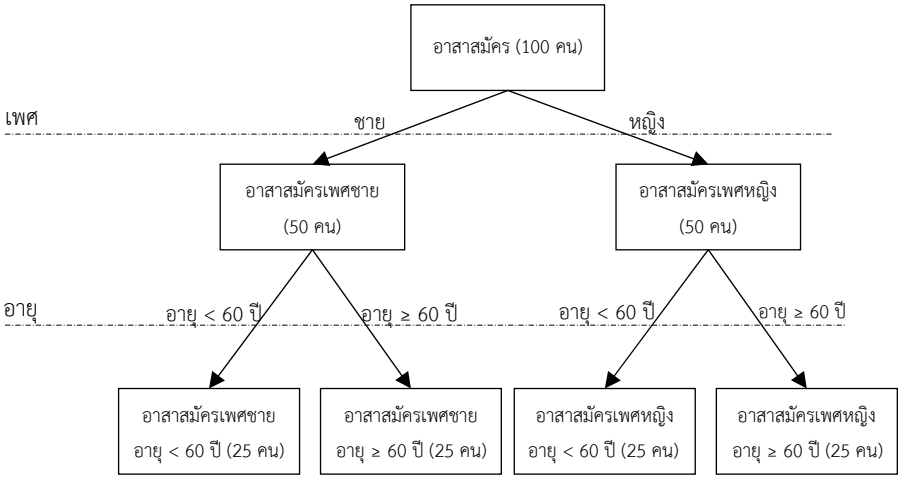
เป็นวิธีที่ทำได้ง่ายและโอกาสที่การจัดสรรสิ่งทดลองและสิ่งควบคุมให้อาสาสมัครมีค่าเท่ากัน เช่น ในกรณีที่เป็นการทดลอง 2 กลุ่ม การจัดสรรทำได้โดยการโยนเหรียญทั้งนี้กำหนดให้ถ้าออกหัว (Head) อาสาสมัครจะได้รับสิ่งแทรกแซง ถ้าออกก้อย (Tail) อาสาสมัครจะได้รับสิ่งควบคุม วิธีการนี้ทำให้เกิดความไม่สมดุลของตัวอย่างในแต่ละกลุ่ม (Imbalance)[2, 3]

3.1.2 การจัดสรรแบบบล็อก

การจัดสรรแบบบล็อกมีวัตถุประสงค์เพื่อให้ขนาดตัวอย่างในแต่ละกลุ่มมีจำนวนขนาดตัวอย่างเท่ากัน (Balance)[2, 4-6] ทั้งนี้การกำหนดรูปแบบบล็อกได้เป็น 2 (Block of 2), 4 (Block of 4) หรือ 8 (Block of 8) เช่น การทดลองจำนวน 2 กลุ่ม คือ A หมายถึงกลุ่มแทรกแซง และ B หมายถึงกลุ่มควบคุม สามารถกำหนดรูปแบบบล็อกได้เป็น Block of 4 ดังนี้ ABBA, AABB, BABA, หรือ ABAB เป็นต้น ถ้าจำนวนตัวอย่างทั้งหมด 100 คน กลุ่ม A จะถูกจัดสรรจำนวน 50 คน และกลุ่ม B จะถูกจัดสรรจำนวน 50 คน

3.1.3 การจัดสรรแบบชั้นภูมิ

การจัดสรรตัวอย่างแบบชั้นภูมิ มีวัตถุประสงค์เพื่อควบคุมปัจจัยที่มีผลต่อการเกิดผลลัพธ์ของการศึกษาและมีผลต่อสิ่งแทรกแซง[7-9] อย่าง เช่น เพศ และ อายุ อาจจะมีผลต่อการเกิดผลลัพธ์ของการศึกษา จึงทำการจัดสรรแบบชั้นภูมิตามเพศ (เพศชายและเพศหญิง) และอายุ (อายุน้อยกว่า 60 ปี และ อายุมากกว่าหรือเท่ากับ 50 ปี) ตัวอย่างรูปที่ 3.3



รูปที่ 3.3 ตัวอย่างการจัดสรรแบบชั้นภูมิ

จากตัวอย่างอาสาสมัครงานวิจัยจำนวน 100 คน จัดสรรตามชั้นภูมิจำนวน 2 ชั้นภูมิ คือ ชั้นภูมิที่ 1 จัดสรรไปอยู่ในกลุ่มเพศชาย 50 คนและเพศหญิง 50 คน จากนั้นอาสาสมัครในแต่ละกลุ่มถูกจัดสรรในชั้นภูมิที่ 2 คือกลุ่มที่มีอายุน้อยกว่า 60 ปี และ มากกว่าหรือเท่ากับ 50 ปี จำนวนกลุ่มละ 25 คน และในกลุ่มเพศหญิงก็จัดสรรในลักษณะเดียวกันดังรูปที่ 3.3

3.1.4 การปกปิด (Blinding)

การศึกษาแบบสุ่มมีกลุ่มควบคุมมีวิธีการลดอคติของงานวิจัยลงได้จากการใช้วิธีการจัดสรรตัวอย่างให้มีโอกาสได้รับสิ่งแทรกแซงหรือกลุ่มควบคุมเท่าๆ กัน อย่างไรก็ตามวิธีการดังกล่าวอาจจะไม่เพียงพอ จึงจะใช้วิธีการปกปิดผู้ที่มีส่วนเกี่ยวข้องในการทำวิจัยไม่ให้ทราบว่าอาสาสมัครงานวิจัยได้รับสิ่งควบคุมหรือสิ่งแทรกแซงการปกปิดประกอบไปด้วย อาสาสมัครงานวิจัย นักวิจัย ผู้เก็บข้อมูล ผู้ประเมินผลลัพธ์งานวิจัย และนักสถิติ[10, 11]

ข้อดี[12,13]

1. ความตรงภายในค่อนข้างสูง (Internal validity)
2. มีการใช้วิธีการจัดสรรแบบสุ่ม (Random allocation) ในการจัดสรรสิ่งแทรกแซงและกลุ่มควบคุมจึงทำให้เกิดอคติในการวิจัยค่อนข้างน้อย
3. ศึกษาผลลัพธ์รองได้หลายผลลัพธ์ (Secondary outcomes)
4. มีความสัมพันธ์เชิงสาเหตุและผล ทำให้สามารถอธิบายสิ่งแทรกแซงที่ทดสอบถึงความเป็นเหตุและผลทำให้เกิดผลลัพธ์

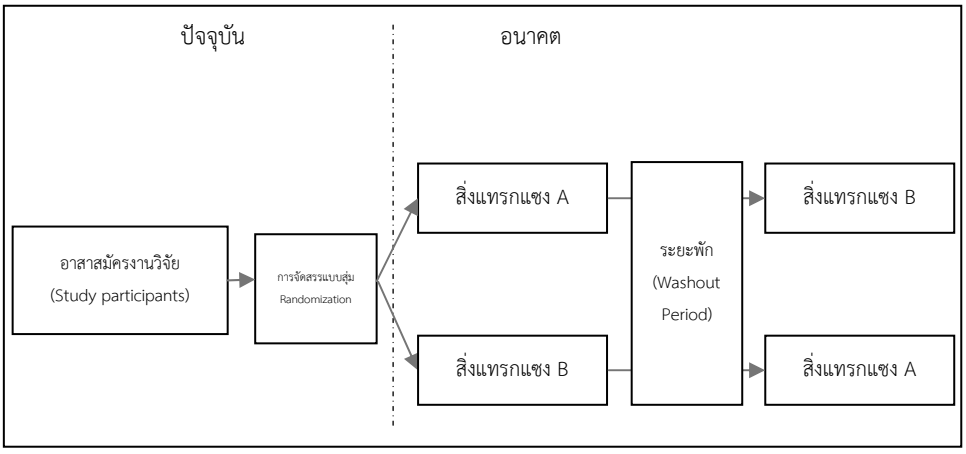
ข้อจำกัด

1. ความเที่ยงตรงภายนอกค่อนข้างน้อย (External validity) เนื่องจากว่าเป็นการทดลองซึ่งมีการกำหนดเกณฑ์การนำเข้าที่เฉพาะเจาะจงเฉพาะกลุ่มทำให้ไม่สามารถขยายออกไปสู่ประชากรได้
2. อาจเกิดการขาดการติดต่อหรือการติดตามอาสาสมัครงานวิจัยได้ (Loss to follow-up)
3. มีค่าใช้จ่ายที่สูง

3.2 การศึกษาแบบสองระยะไขว้กัน

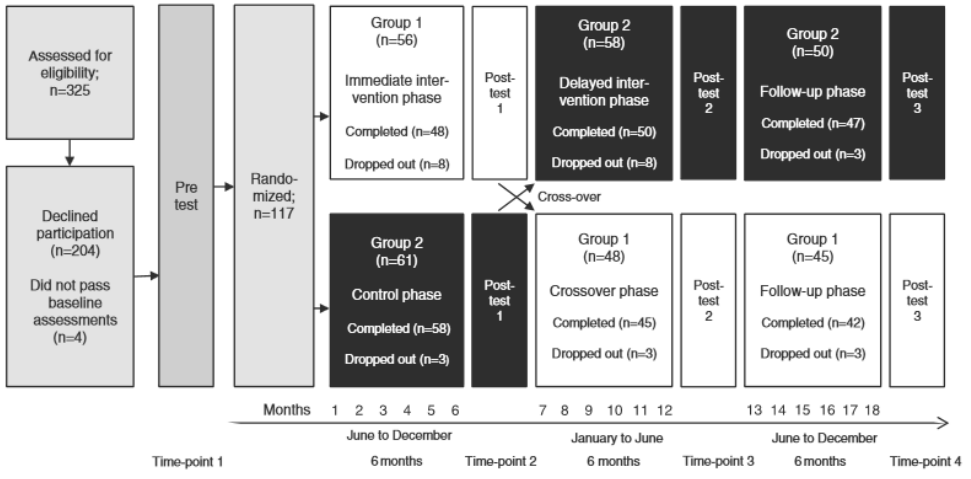
รูปแบบการศึกษาแบบสองระยะไขว้กัน[15-17] (Cross-over design) ส่วนใหญ่จะนิยมใช้กันในงานวิจัยทางคลินิก (Clinical trial) ที่เกี่ยวข้องกับผู้ป่วยที่เป็นโรคเรื้อรังเพราะอาสาสมัครจะได้รับทั้งสิ่งแทรกแซงและสิ่งควบคุมและใช้เวลาในการศึกษาที่นานขึ้น รูปแบบการศึกษานี้มีการจัดสรรสิ่งแทรกแซง (สมมุติเป็น A) และสิ่งควบคุม (สมมุติเป็น B) ให้อาสาสมัครงานวิจัยเป็นลำดับ การกำหนดลำดับ (Sequence) สามารถกำหนดรูปแบบได้ 2 รูปแบบคือ (กรณีมีสิ่งทดลองและสิ่งควบคุมอย่างละ 1) รูปแบบที่ 1 คือ AB และ BA (มี 2 แนวทาง) รูปแบบที่ 2 คือ AABB, BBAA, ABAB และ BABA (มี 4 แนวทาง) สมมุติว่าใช้รูปแบบที่ 1 ในการจัดสรรสิ่งแทรกแซงและสิ่งควบคุม สามารถทำได้ดังนี้ หนึ่งฉลากจะมีลำดับของการให้สิ่งแทรกแซงและสิ่งควบคุมที่ต่างกัน อาจจะได้รับ A ก่อนหรือได้รับ B ก่อน ขึ้นอยู่กับการสุ่มได้ฉลากอันไหน สมมุตินักวิจัยสุ่มได้ฉลาก AB เมื่อทำการ

เผยฉลากออกมาอาสาสมัครงานวิจัยจะได้รับ A ก่อน จากนั้นรอให้หมดฤทธิ์ผลของ A (Wash-out period) จึงจะได้รับ B ต่อ แผนผังการวิจัยดังรูปที่ 3.4



รูปที่ 3.4 รูปแบบการศึกษาแบบสองระยะไขว้กัน

ตัวอย่างที่ 3.2 การทดสอบการอบรมหลายรูปแบบ (Multimodal training) มีผลต่อการกลับมาใช้งานการทำหน้าที่ทางกายของผู้สูงอายุหรือไม่ (Retention of functional fitness)[17] ทำการศึกษาในผู้สูงอายุ 71 ถึง 90 ปี จำนวน 117 ราย สิ่งแทรกแซงคือการฝึกความแข็งแกร่งของร่างกายสองครั้งต่อสัปดาห์และฝึกความทนทานของร่างกายในแต่ละวัน (Daily endurance) ผลการศึกษาพบว่าการออกกำลังกายดังกล่าวสามารถทำให้การกลับมาใช้งานการทำหน้าที่ทางกายของผู้สูงอายุดีขึ้น ผังการทดลองการศึกษาดังกล่าวใน รูปที่ 3.5 การศึกษาเริ่มจากประเมินอาสาสมัครงานวิจัยเพื่อเข้าสู่งานวิจัยจำนวน 325 ราย และทำการจัดสรรอาสาสมัครงานวิจัยจำนวน 117 ราย โดยใช้วิธีบล๊อก กลุ่มแรกจะได้สิ่งทดลองก่อน (Group 1) ต่อมาจะได้กลุ่มควบคุม (Group 2) และติดตามประเมินผลอีก 1 ครั้ง ในขณะที่อีกกลุ่มหนึ่งจะได้กลุ่มควบคุม (Group 2) ก่อนและตามด้วยสิ่งทดลอง (Group 1) จากนั้นประเมินผลลัพธ์อีก 1 ครั้ง จากการศึกษานี้จะเห็นว่าไม่มีระยะเวลาหมดฤทธิ์ของสิ่งทดลองหรือควบคุมเป็นเพราะว่าสิ่งทดลองเป็นการออกกำลังกาย ดังนั้นสามารถสลับเพื่อรับสิ่งแทรกแซงหรือควบคุมในลำดับที่ 2 ได้เลย (Sequence)



รูปที่ 3.5 ตัวอย่าง ผังการทดลองแบบ 2x2 Crossover design
ที่มา: Gudlaugsson J, Gudnason V, Aspelund T, Siggeirsdottir K, Olafsdottir AS, Jonsson PV, et al. Effects of a 6-month multimodal training intervention on retention of functional fitness in older adults: a randomized-controlled cross-over design. Int J BehavNutrPhys Act. 2012;9:107.[17]

ข้อดี[16, 18]

1. สามารถลดอคติที่เกิดจากตัวแปรกวนได้อย่างสมบูรณ์ เพราะอาสาสมัครงานวิจัยได้รับทั้ง Intervention และ Control ในคนเดียวกัน
2. สามารถลดความไม่เท่าเทียมกันของตัวแปรข้อมูลพื้นฐาน (Baseline characteristics) ระหว่างกลุ่มได้
3. เหมาะกับโรคเรื้อรังเช่น โรคหืดหอบ (Asthma)

ข้อกำจัด

1. อาสาสมัครงานวิจัยอาจขอถอนตัวออกจากการศึกษาหลังจากได้รับ Intervention หรือ Control ในคาบแรก (1^{st} period)
2. มีโอกาสที่ฤทธิ์ของสิ่งทดลองในคาบแรกไม่หมด และส่งผลถึงสิ่งทดลองในคาบที่ 2 ได้ หรือเรียกว่า Carry-over effect

3.3 การทดลองแบบสี่กลุ่มโซโลมอน

รูปแบบการทดลองแบบสี่กลุ่มโซโลมอน[19, 20] (Solomon four-group design) มีลักษณะคล้ายกับรูปแบบการศึกษาแบบสุ่มมีกลุ่มควบคุม (Randomized controlled trial) โดยรูปแบบการทดลองแบบสี่กลุ่มโซโลมอนดังรูปที่ 3.6

กลุ่มที่	การวัดผลก่อนทดลอง (Pre-test)		การวัดผลหลังทดลอง (Post-test)
1	วัดผล	สิ่งแทรกแซง	วัดผล
2	วัดผล	กลุ่มควบคุม	วัดผล
3	ไม่วัดผล	สิ่งแทรกแซง	วัดผล
4	ไม่วัดผล	กลุ่มควบคุม	วัดผล

รูปที่ 3.6 ผังการทดลองแบบสี่กลุ่มโซโลมอน

รูปแบบการศึกษาจะแบ่งกลุ่มทดลองออกเป็น 4 กลุ่มโดยที่กลุ่มที่ 1 และ 2 มีการวัดผลก่อนการทดลอง (Pretesting) ส่วนกลุ่มที่ 3 และ 4 ไม่มีการวัดผลก่อนการทดลองและมีการวัดผลหลังการทดลอง (Post-test) ทั้ง 4 กลุ่ม สำหรับข้อดี ข้อด้อยมีดังนี้

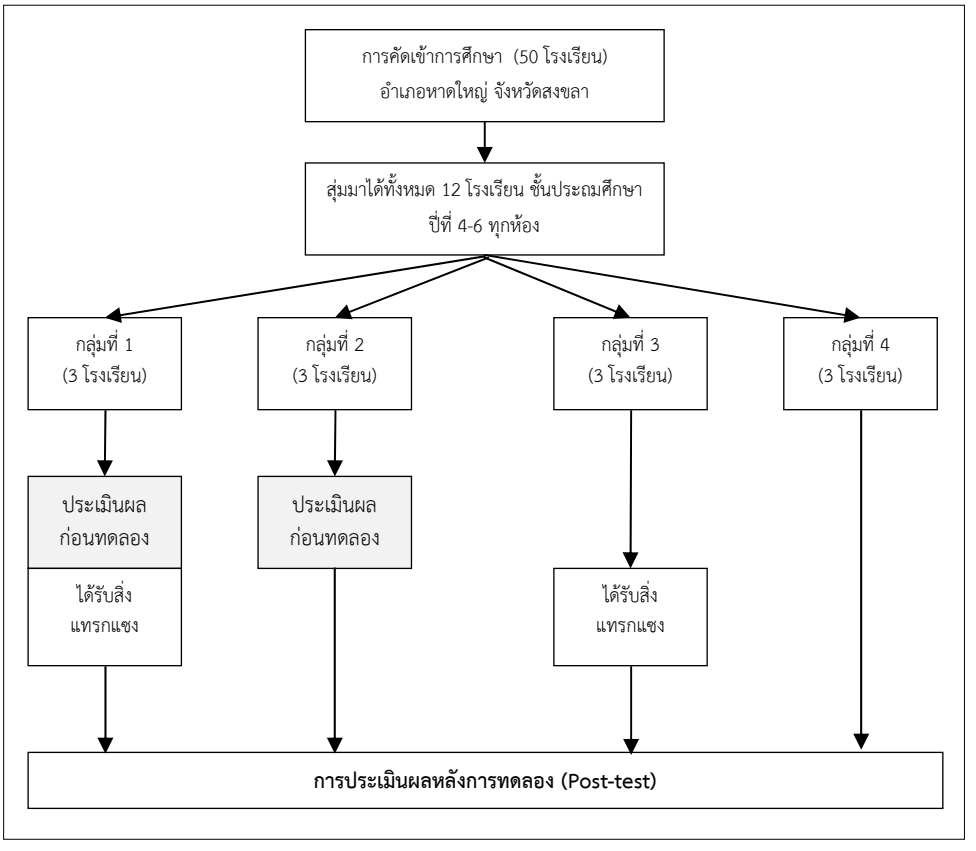
ข้อดี

1. เหมาะกับการทดลองการวัดผลก่อนการทดลอง ว่ามีผลต่อการวัดผลหลังทดลองหรือไม่ (Pretesting effect)
2. สามารถวัดผลของฤดูกาลหรือระยะเวลาที่ผ่านมา (Seasonal effect) ว่ามีผลต่อการวัดผลลัพธ์ทางการศึกษาหรือไม่

ข้อจำกัด

1. มีการใช้งบประมาณวิจัยที่สูงและใช้ทรัพยากรที่มาก เช่น ผู้ร่วมวิจัย และ ทีมวิจัยที่มาก เพราะจากรูปแบบดังกล่าวเป็นการเพิ่มจำนวนกลุ่มทดลองและกลุ่มควบคุมอีก 1 เท่า
2. ใช้สถิติขั้นสูงในการวิเคราะห์ผล

ตัวอย่างที่ 3.3 การศึกษาผลของโปรแกรมบ้านปลอดในนักเรียนชั้นประถมศึกษาปีที่ 4 ถึง 6 โดยทดลองโปรแกรมให้ความรู้ (Educational Intervention) บ้านปลอดบุหรี่แก่เด็กนักเรียน[21] มีแผนผังการทดลองดังรูปที่ 3.7



รูปที่ 3.7 ตัวอย่างผังการทดลองแบบสี่กลุ่มโซโลมอน
ที่มา: ดัดแปลงจาก Intarut N, Chongsuvivatwong V, McNeil E. Effects of a School-based Intervention Program on Attitude and Knowledge of Household Members Towards a Smoke-free Home: a Cluster Controlled Trial. Asian Pac J Cancer Prev. 2016;17(3):1235-42. Figure 2. Study Flow Diagram.[21]

จากตัวอย่างที่ 3.3 เป็นการทดสอบประสิทธิผลของโปรแกรมบ้านปลอดบุหรี่ โดยเด็กนักเรียนชั้นประถมศึกษาปีที่ 4-6 (Change agent) ในกรณีนี้อาจจะใช้รูปแบบการศึกษาแบบสุ่มมีกลุ่มควบคุมได้ แต่จากการทบทวนวรรณกรรมของนักวิจัย พบว่าโปรแกรมหรือสิ่งแทรกแซงที่มีลักษณะการให้ความรู้ สอน หรือ อบรม (Educational or Behavioral intervention) มักจะเกิดอคติในตอนวัดผลลัพธ์ครั้งแรก ที่เรียกว่า Pretesting effects[22-24] และการศึกษานี้มีการวัดความรู้และทัศนคติเป็นผลลัพธ์รอง ทำให้นักวิจัยคาดว่าน่าจะเกิด Pretesting effect ขึ้นได้ จึงออกแบบการวิจัยโดยใช้รูปแบบการทดลองแบบสี่กลุ่มของโซโลมอนและจากผลการทดสอบทางสถิติพบว่าไม่มีผลของ Pretesting effect ในการวัดผลหลังทดลอง

3.4 บทสรุป

จากรูปแบบการทดลองทั้ง 3 รูปแบบจะเห็นว่าแต่ละรูปแบบมีข้อดี และข้อจำกัดที่แตกต่างกัน การจะเลือกใช้รูปแบบการทดลองแบบใดขึ้นอยู่กับวัตถุประสงค์ของงานวิจัย การดำเนินของโรคเช่นโรคที่สามารถรักษาหายขาดได้อาจจะใช้รูปแบบการทดลองแบบสุ่มมีกลุ่มควบคุม (Randomized controlled trial) หรือ โรคเรื้อรังอาจใช้การศึกษาแบบไขว้ (Cross-over design) หรือถ้าเป็นงานวิจัยที่มีสิ่งแทรกแซงเป็นลักษณะการให้ความรู้ อบรม หรือ การปรับเปลี่ยนพฤติกรรม ควรทบทวนวรรณกรรมที่เกี่ยวข้องถึงการเกิดขึ้นของ Pretesting effect ว่าจะเกิดขึ้นหรือไม่ ถ้ามีก็ใช้รูปแบบงานวิจัยแบบสี่กลุ่มโซโลมอน

3.5 แบบฝึกหัด

1. งานวิจัยรูปแบบสุ่มที่มีกลุ่มควบคุมการจัดสรรแบบสุ่มมีวัตถุประสงค์เพื่ออะไร
2. รูปแบบการศึกษาแบบสุ่มมีกลุ่มควบคุมควรจะมีการติดตามผลการทดลองยาวนานเท่าใด
3. รูปแบบการศึกษาแบบไขว้กันเหมาะกับโรคใด
4. ระยะเวลาที่รอให้ฤทธิ์ของสิ่งทดลองหมด (Wash-out period) ขึ้นอยู่กับสิ่งใด
5. วัตถุประสงค์หลักของการใช้รูปแบบงานวิจัยแบบสี่กลุ่มโซโลมอนคืออะไร

3.6 เอกสารอ้างอิง

1. Nave AH, Rackoll T, Grittner U, Blasing H, Gorsler A, Nabavi DG, et al. Physical Fitness Training in Patients with Subacute Stroke (PHYS-STROKE): multicentre, randomised controlled, endpoint blinded trial. BMJ. 2019;366:l5101.
2. Efird J. Blocked randomization with randomly selected block sizes. Int J Environ Res Public Health. 2011;8(1):15-20.
3. Vickers AJ. How to randomize. J Soc Integr Oncol. 2006;4(4):194-8.
4. Suresh K. An overview of randomization techniques: An unbiased assessment of outcome in clinical research. J Hum Reprod Sci. 2011;4(1):8-11.
5. Lim CY, In J. Randomization in clinical studies. Korean J Anesthesiol. 2019;72(3):221-32.
6. Yang HL, Wu XB, Mao C. [Block randomization in clinical trials]. Zhonghua Yu Fang Yi Xue Za Zhi. 2019;53(4):437-40.

7. Karanicolas PJ, Farrokhyar F, Bhandari M. Practical tips for surgical research: blinding: who, what, when, why, how? Can J Surg. 2010;53(5):345-8.
8. Miller LE, Stewart ME. The blind leading the blind: use and misuse of blinding in randomized controlled trials. Contemp Clin Trials. 2011;32(2):240-3.
9. Stanley K. Design of randomized controlled trials. Circulation. 2007;115(9):1164-9.
10. Booth CM, Tannock IF. Randomised controlled trials and population-based observational research: partners in the evolution of medical evidence. Br J Cancer. 2014;110(3):551-5.
11. Spieth PM, Kubasch AS, Penzlin AI, Illigens BM, Barlinn K, Siepmann T. Randomized controlled trials - a matter of design. Neuropsychiatr Dis Treat. 2016;12:1341-9.
12. Tripepi G, Jager KJ, Dekker FW, Zoccali C. Selection bias and information bias in clinical research. Nephron Clin Pract. 2010;115(2):c94-9.
13. Mezei G, Kheifets L. Selection bias and its implications for case-control studies: a case study of magnetic field exposure and childhood leukaemia. Int J Epidemiol. 2006;35(2):397-406.
14. Wellek S, Blettner M. On the proper use of the crossover design in clinical trials: part 18 of a series on evaluation of scientific publications. Dtsch Arztebl Int. 2012;109(15):276-81.
15. Mills EJ, Chan AW, Wu P, Vail A, Guyatt GH, Altman DG. Design, analysis, and presentation of crossover trials. Trials. 2009;10:27.
16. Sibbald B, Roberts C. Understanding controlled trials. Crossover trials. Bmj. 1998;316(7146):1719.
17. Gudlaugsson J, Gudnason V, Aspelund T, Siggeirsdottir K, Olafsdottir AS, Jonsson PV, et al. Effects of a 6-month multimodal training intervention on retention of functional fitness in older adults: a randomized-controlled cross-over design. Int J Behav Nutr Phys Act. 2012;9:107.
18. Krogh HB, Storebo OJ, Faltinsen E, Todorovac A, Ydedahl-Jensen E, Magnusson FL, et al. Methodological advantages and disadvantages of parallel and crossover randomised clinical trials on methylphenidate for attention deficit hyperactivity disorder: a systematic review and meta-analyses. BMJ Open. 2019;9(3):e026478.

19. McGahee T. The Use Of The Solomon Four-Group Design In Nursing Research. Southern Online Journal of Nursing Research. 2009;9.
20. Huck SW, Sandler HM. A Note on the Solomon 4-Group Design. The Journal of Experimental Education. 1973;42(2):54-5.
21. Intarut N, Chongsuvivatwong V, McNeil E. Effects of a School-based Intervention Program on Attitude and Knowledge of Household Members Towards a Smoke-free Home: a Cluster Controlled Trial. Asian Pac J Cancer Prev. 2016;17(3):1235-42.
22. Richland LE, Kornell N, Kao LS. The pretesting effect: do unsuccessful retrieval attempts enhance learning? J Exp Psychol Appl. 2009 Sep;15(3):243-57.
23. Latimier A, Riegert A, Peyre H, Ly ST, Casati R, Ramus F. Does pre-testing promote better retention than post-testing?. NPJ Sci Learn. 2019;4:15. Published 2019 Sep 24. doi:10.1038/s41539-019-0053-1.
24. Hallberg K, Cook TD, Steiner PM, Clark MH. Pretest Measures of the Study Outcome and the Elimination of Selection Bias: Evidence from Three Within Study Comparisons. Prev Sci. 2018;19(3):274-283.

บทที่ 4
อคติในงานวิจัย

จากรูปแบบการศึกษาในบทที่ 2 และ 3 เราได้เรียนรู้ถึงลักษณะของรูปแบบการศึกษาเชิงสังเกตการณ์และเชิงทดลอง ในบทนี้จะได้ศึกษาเกี่ยวกับอคติ (Bias) ที่เกิดขึ้นกับงานวิจัย สามารถแบ่งออกได้เป็น 2 แบบ คือ ความคลาดเคลื่อนเชิงระบบ (Systematic error) และความคลาดเคลื่อนแบบสุ่ม (Random error) ความคลาดเคลื่อนแบบสุ่มเกิดจากการประมาณค่าปัจจัยเสี่ยง (Exposure factor) หรือผลลัพธ์ทางการศึกษา (Outcome) หรือขนาดตัวอย่างที่น้อยเกินไป ส่วนความคลาดเคลื่อนแบบระบบเกิดขึ้นในขั้นตอนการทำวิจัยประกอบไปด้วย อคติเกี่ยวกับการเลือกตัวอย่าง (Selection bias) อคติของข้อมูล (Information bias) และอคติจากตัวแปรกวน (Confounding)

4.1 อคติจากการเลือกตัวอย่าง

อคติจากการเลือกตัวอย่าง[1-3] (Selection bias) สามารถเกิดขึ้นได้ในทุกรูปแบบการศึกษา เป็นอคติที่เกิดขึ้นจากการเลือกตัวอย่างเข้ามาในงานวิจัยสามารถลดอคติดังกล่าวได้ โดยการกำหนดเกณฑ์ในการคัดเข้าและคัดออกของอาสาสมัครงานวิจัยให้รัดกลุ่ม สามารถแบ่งอคติดังกล่าวดังนี้ คือ อคติจากการไม่ตอบรับ[4, 5] (Non-response bias) อคติจากการเกิดอุบัติการณ์-ความชุก[6] (Incidence-Prevalence bias) อคติจากการสูญหายในระหว่างการติดตาม[7-9] (Loss to follow-up bias) และ อคติโดยอาสาสมัคร[10-12] (Volunteer bias)

4.1.1 อคติจากการไม่ตอบรับ

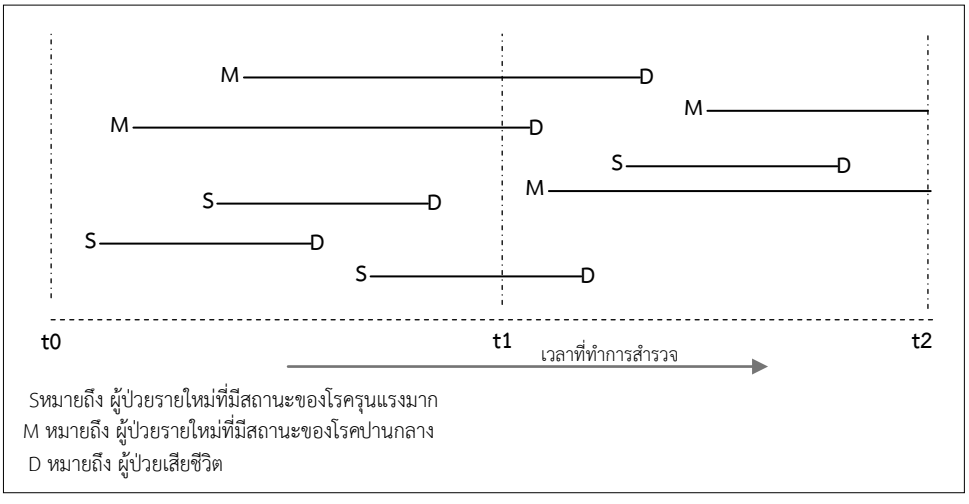
อคติจากการไม่ตอบรับ (Non-response Bias) เกิดขึ้นโดยเมื่อเราได้เชิญอาสาสมัครเข้าร่วมงานวิจัยแล้วผู้ร่วมวิจัยไม่ได้ตอบข้อคำถามที่เราได้ถามหรือเกิดจากการปฏิเสธเข้าร่วมวิจัย หรือ เกิดจากการที่ไม่สามารถเข้าถึงอาสาสมัครงานวิจัยที่กำหนดไว้ได้อคติเหล่านี้อาจจะเกิดขึ้นจากการตอบแบบสอบถามที่ข้อคำถามที่มีความอ่อนไหว (Sensitive question) หรือ มาตรการเชิญอาสาสมัครงานวิจัยอาจไม่ตรงกับความสะดวกของอาสาสมัครงานวิจัย

แนวทางลดอคติโดยการทดลองการเก็บข้อมูลก่อนการเก็บข้อมูลจริง หลีกเลี่ยงคำถามที่ใช้เวลานานเกินไปหรือสั้นเกินไป ถ้ามีการนัดหมายในการติดตามผลการทดลองควรส่งข้อความแจ้งเตือนให้อาสาสมัครงานวิจัยทราบล่วงหน้า ในการสัมภาษณ์ควรจัดห้องให้มีลักษณะปิดเป็น

ส่วนตัว และในบางครั้งอาสาสมัครงานวิจัยอาจจะไม่สะดวกอยู่ให้นักวิจัยสัมภาษณ์ อันเนื่องมาจากการต้องไปประกอบอาชีพ หรือ การไปทำธุระ ต่างๆ ดังนั้นอาจจ่ายค่าตอบแทนเพื่อเป็นแรงจูงใจ (Incentive) ในการเพิ่มโอกาสการเข้าร่วมวิจัยของอาสาสมัคร

4.1.2 อคติจากการเกิดอุบัติการณ์-ความชุก

อคติจากการเกิดอุบัติการณ์-ความชุก (Incidence-Prevalence bias หรือ Neyman'sbias หรือ Survival bias) เกิดจากการที่เลือกเกณฑ์การนำเข้าหรือคัดออก Case หรือ Control ที่มีการสัมผัสปัจจัยเสี่ยงซึ่งเกิดก่อนการเกิดเหตุการณ์ หรือเกิดหลังการเกิดเหตุการณ์



รูปที่ 4.1 อคติจากการเกิดอุบัติการณ์-ความชุก

ที่มา: ดัดแปลงมาจาก Tripepi G, Jager KJ, Dekker FW, Wanner C, Zoccali C. Bias in clinical research. Kidney Int. 2008;73(2):148-53.Figure 1. A hypothetical example of incidence-prevalence bias.[13]

ตัวอย่างดังรูปที่ 4.1 ในการติดตามอาสาสมัครเริ่มต้นด้วยการเป็นโรคที่มีระดับรุนแรงมาก (S) จำนวน 4 ราย ระดับรุนแรงปานกลาง (M) จำนวน 4 ราย เวลาติดตามเริ่มต้น (t0) ถึงเวลาสุดท้าย (t2) เมื่อคำนวณความเสี่ยงในการเสียชีวิตแล้วพบว่าคนที่เป็นโรคที่มีระดับรุนแรงมากมีความเสี่ยงทำให้เสียชีวิตเป็น 1 เท่า (S/M: 4/4 =1) เมื่อเทียบกับคนที่เป็นโรคระดับรุนแรงปานกลาง (เสียชีวิตเท่ากัน) อย่างไรก็ตามถ้าทำการศึกษาแบบภาคตัดขวาง ณ เวลา t1 จะได้ว่าคนที่เป็นโรคที่มีระดับรุนแรงมากมีความเสี่ยงทำให้เสียชีวิตเป็น 0.5 เท่า (S/M: 1/2 = 0.5)

เมื่อเทียบกับคนที่เป็นโรคระดับรุนแรงปานกลาง ซึ่งจะเห็นว่าคนที่เสี่ยงระดับรุนแรงเสียชีวิตน้อยกว่าคนที่เสี่ยงระดับปานกลาง จากข้อมูลดังกล่าวถ้าเหตุการณ์เกิดก่อนหรือหลังการศึกษาผลการศึกษาที่ได้อาจจะแตกต่างจากความเป็นจริง

4.1.3 อคติจากการสูญหายในระหว่างการติดตาม

อคติจากการสูญหายในระหว่างการติดตามผลหรือติดต่ออาสาสมัครงานวิจัยไม่ได้ (Loss to follow-up bias) ในระหว่างการทำวิจัย ทำให้ไม่สามารถทราบสถานะของอาสาสมัครที่หายไปมีลักษณะการสัมผัสปัจจัยเสี่ยงหรือเกิดผลลัพธ์ทางการศึกษาเป็นอย่างไรบ้าง ทั้งนี้สามารถลดอคตินี้ได้ในขั้นตอนของการวิเคราะห์ข้อมูลได้โดยใช้การวิเคราะห์ทางสถิติที่ชื่อว่า Intention-To-Treat (ITT) [14-16]

4.1.4 อคติโดยอาสาสมัคร

อคติโดยอาสาสมัคร (Volunteer bias) เกิดจากการที่อาสาสมัครงานวิจัยสมัครเข้าร่วมโครงการ อาจเพราะสนใจในสุขภาพมากขึ้นหรือคิดว่าได้ประโยชน์เฉพาะตน ทำให้ไม่สามารถได้อาสาสมัครงานวิจัยที่เข้าร่วมเป็นประชากรทั้งหมด[17] ตัวอย่างเช่นการศึกษาของ Beer, C., และคณะ (2010) ได้ทำการศึกษาความสัมพันธ์คะแนนคุณภาพชีวิตของผู้ป่วยโรคสมองเสื่อม แบบกรอกข้อมูลเองและแบบสัมภาษณ์โดยผู้ช่วยงานวิจัยเป็นผู้กรอก[18] พบว่าผู้ป่วยสามารถให้คะแนนคุณภาพชีวิตได้ดีทั้งสองวิธี อย่างไรก็ตามเนื่องจากการศึกษานี้ได้เชิญชวนอาสาสมัครงานวิจัยที่อยู่ในสถานรักษาผู้ป่วย (Residential care facilities) ทำให้คะแนนคุณภาพชีวิตมีระดับที่ดีอาจเป็นเพราะกลุ่มตัวอย่างที่อาศัยอยู่ในสถานรักษา เป็นคนที่ได้รับการรักษาหรือฟื้นฟูได้ดีกว่าประชากรผู้ป่วยโรคสมองเสื่อมที่อาศัยอยู่ในที่อื่น

4.2 อคติของข้อมูล

อคติของข้อมูล (Information biasหรือ Observation bias หรือ Measurement bias) เกิดจากการได้ข้อมูลจากแหล่งของข้อมูล (อาสาสมัครงานวิจัย ฐานข้อมูลวิจัย จากฐานข้อมูลโรงพยาบาล ข้อมูลจาก 43 แฟ้มของกระทรวงสาธารณสุข) อาจจะมีข้อมูลบางข้อไม่ตรงกับความเป็นจริงเรียกว่าการจำแนกกลุ่มผิดประเภท (Misclassification) สามารถแบ่งได้ 2 แบบ คือ Non-differential misclassificationและ Differential misclassification[19-21]

4.2.1 Non-differential misclassification

ลักษณะของ Non-differential misclassification คือ การจัดกลุ่มการสัมผัสปัจจัยเสี่ยงหรือผลลัพธ์ทางการศึกษาผิดเพี้ยนไปจากเดิม ทำให้ค่าความเสี่ยงสัมพัทธ์ (Odds ratio) มีค่าเข้าใกล้ 1 ตัวอย่างการทดสอบความสัมพันธ์ของการได้รับควันบุหรี่มือสองต่อการเกิดโรคมะเร็งปอดได้ข้อมูลที่เป็นจริงดังข้อมูลตารางที่ 4.1

ตารางที่ 4.1 การสัมผัสควันบุหรี่มือสองต่อการเกิดโรคมะเร็งปอดกรณีไม่มีการเกิดอคติ

การได้รับควันบุหรี่มือสอง	เป็นมะเร็งปอด	ไม่เป็นมะเร็งปอด	Odds ratio
ได้รับ	2,000 (a)	1,000 (b)	6.0
ไม่ได้รับ	1,000 (c)	3,000 (d)	
รวม	3,000	4,000	

ตารางที่ 4.2 การสัมผัสควันบุหรี่มือสองต่อการเกิดโรคมะเร็งปอดกรณีมีการเกิดอคติ 40% จากค่าจริง

การได้รับควันบุหรี่มือสอง	เป็นมะเร็งปอด	ไม่เป็นมะเร็งปอด	Odds ratio
ได้รับ	1,200 (a)	600 (b)	3.7
ไม่ได้รับ	1,800 (c)	3,400 (d)	
รวม	3,000	4,000	

จากตารางที่ 4.1 สมมุติเป็นข้อมูลที่เกิดขึ้นจริงค่าความเสี่ยงสัมพัทธ์มีค่าเท่ากับ 6.0 ในขณะที่ถ้าในงานวิจัยมีการเกิดอคติแบบ Non-differential misclassification ขึ้นดังตารางที่ 4.2 การผิดเพี้ยนของตัวอย่างในการจัดกลุ่มการสัมผัสปัจจัยเสี่ยง ถูกจัดไปอยู่ในกลุ่มไม่สัมผัสปัจจัยเสี่ยงจำนวน 40% ทำให้ค่าความเสี่ยงสัมพัทธ์เท่ากับ 3.7 ซึ่งลดลงจากความเป็นจริงซึ่งเท่ากับ 6.0 (ลดลง 38.3%)

ตัวอย่างที่ 4.1 CharuluxananS. Intarut N และคณะ (2007) ได้ทำการศึกษาความสัมพันธ์ปัจจัยเสี่ยงที่ทำให้เกิดภาวะออกซิเจนในเลือดต่ำ (Oxygen desaturation) ในผู้ป่วยระหว่างการผ่าตัด นักวิจัยได้ทำการทดสอบปัจจัยทางด้านการจำแนกภาวะความแข็งแรงของผู้ป่วยและความเสี่ยงต่อการให้ยาระงับความรู้สึก (ASA physical status) กับการเกิดภาวะออกซิเจนในเลือดต่ำผลการวิเคราะห์พบว่ามีโอกาสที่จะมีการจัดกลุ่มของ ASA physical status ผิดเพี้ยนไป

(Non-differential misclassification) ทำให้ค่าความเสี่ยงสัมพัทธ์ของความสัมพันธ์ระหว่าง ASA physical status กับการเกิดภาวะออกซิเจนในเลือดต่ำ อาจจะมีค่าที่ต่ำกว่าความเป็นจริง[22]

4.2.2 Differential misclassification

ลักษณะของ Differential misclassification จะมีลักษณะที่ตรงกันข้ามกับ Non-differential misclassification คือทิศทางของค่าความเสี่ยงสัมพัทธ์ออกห่างจาก 1 หรือเข้าใกล้ 1 (ค่อนข้างกำหนดทิศทางยาก) โดยมีการจัดกลุ่มการสัมผัสปัจจัยเสี่ยงที่มากกว่าปกติ (Misclassification of exposure) หรือมีการจัดกลุ่มผลลัพธ์ทางการศึกษาที่ผิดเพี้ยนไป (Misclassification of outcome) สาเหตุอาจเกิดจากการเกิดอคติในการจดจำการสัมผัสปัจจัยเสี่ยงในอดีต (Recall bias) แตกต่างกันหรือเกิดจากนักวิจัยทราบว่าอาสาสมัครรายนี้เป็นคนที่เกิดโรค (Case) จึงเฝ้าถามหรือซักไซ้ถึงการสัมผัสปัจจัยเสี่ยง (Interviewer bias) ส่วนใหญ่จะเจอในการศึกษาแบบ Case-control study, Hospital-based case-control study ตัวอย่างในตารางที่ 4.3 สมมุติว่า ความสัมพันธ์ของจำนวนมวนบุหรี่ที่สูบต่อการเกิดโรคมะเร็งปอดได้ข้อมูลที่เป็นจริงดังตาราง

ตารางที่ 4.3 ความถี่ของความสัมพันธ์ระหว่างจำนวนมวนที่สูบบุหรี่กับการเป็นโรคมะเร็งปอด

จำนวนมวนบุหรี่ที่สูบต่อวัน	เป็นมะเร็งปอด (Case)	ไม่เป็นมะเร็งปอด (Control)	Odds ratio
มากกว่า 25 (Exposure)	2,000 (a)	1,000 (b)	6.0
น้อยกว่า 25 (Non-exposure)	1,000 (c)	3,000 (d)	
รวม	3,000	4,000	

ตารางที่ 4.4 ความถี่ของความสัมพันธ์ระหว่างจำนวนมวนที่สูบบุหรี่กับการเป็นโรคมะเร็งปอด มีเกิด Misclassification กลุ่ม Case-non-exposure เป็น Case-exposure จำนวน 40%

จำนวนมวนบุหรี่ที่สูบต่อวัน	เป็นมะเร็งปอด (Case)	ไม่เป็นมะเร็งปอด (Control)	Odds ratio
มากกว่า 25 (Exposure)	2,800 (a) (2,000+2,000x0.4)	1,000 (b)	42
น้อยกว่า 25 (Non-exposure)	200 (c) (1,000-2,000x0.4)	3,000 (d)	
รวม	3,000	4,000	

ตารางที่ 4.5 ความถี่ของความสัมพันธ์ระหว่างจำนวนมวนที่สูบบุหรี่กับการเป็นโรคมะเร็งปอดมี Misclassification กลุ่ม Case-exposure เป็น Control-non-exposure จำนวน 40%

จำนวนมวนบุหรี่ที่สูบต่อวัน	เป็นมะเร็งปอด (Case)	ไม่เป็นมะเร็งปอด (Control)	Odds ratio
มากกว่า 25 (Exposure)	1,200 (a) (2,000-2,000x0.4)	1,000 (b)	2
น้อยกว่า 25 (Non-exposure)	1,800 (c) (1,000+2,000x0.4)	3,000 (d)	
รวม	3,000	4,000	

จากตารางที่ 4.3 สมมุติข้อมูลที่ได้มีความเป็นจริงและเกิดอคติน้อยที่สุดจึงนำมาเป็นข้อมูลที่ใช้อ้างอิงโดยมีค่า Odds ratio เท่ากับ 6 สมมุติในการศึกษามีอคติเกิดขึ้นโดยกลุ่มสัมผัสปัจจัยเสี่ยง (Exposure) มีค่าเพิ่มขึ้นจากความเป็นจริง 40% ของค่าจริง (2000 ราย) ทำให้ได้ข้อมูลดังตารางที่ 4.4 เมื่อคำนวณค่า Odds ratio แล้วเท่ากับ 42 เกิดการประมาณค่าที่สูงเกินความเป็นจริง (Overestimation) อีกหนึ่งกรณีศึกษาสมมุติการศึกษาเกิดอคติโดยกลุ่มสัมผัสปัจจัยเสี่ยง (Exposure) มีค่าลดลงจากความเป็นจริง 40% ข้อมูลดังตารางที่ 4.5 คำนวณค่า Odds ratio เท่ากับ 2 ทำให้ได้ค่าการประมาณค่าต่ำกว่าความเป็นจริง (Underestimation)

การลดอคติ Misclassification สามารถทำได้ดังนี้[23-25]

1. แบบสอบถามสำหรับการเก็บข้อมูลควรมีมาตรฐานเดียวกันและจัดทำคู่มือการเก็บข้อมูลลักษณะข้อคำถามควรเป็นปลายปิด (Closed-end question) (ลด Interviewer bias และ Recall bias)
2. ควรทดลองเก็บข้อมูลก่อนนำไปใช้จริงเพื่อนำปัญหาอุปสรรคที่ได้มาปรับปรุงแก้ไข (ลด Interviewer bias และ Recall bias)
3. อบรมวิธีการเก็บข้อมูลแก่ทีมวิจัยก่อนการเก็บข้อมูลเพื่อให้มีมาตรฐานเดียวกันในการเก็บข้อมูล (ลด Interviewer bias)
4. ปกปิด (Blinding) อาสาสมัครงานวิจัย นักวิจัยผู้เก็บข้อมูลหรือนักชีวสถิติสถิติ (ลด Interviewer bias)

4.3 อคติจากตัวแปรกวน

ตัวแปรกวน (Confounding variable) เป็นอคติที่เกิดจากมีปัจจัยตัวที่ 3 เข้ามาบิดเบือนค่าความเสี่ยงสัมพัทธ์ของความสัมพันธ์ระหว่างสองตัวแปรที่ทดสอบให้มีค่ามากกว่า (Overestimation) หรือมีค่าน้อยกว่าความเป็นจริง (Underestimation) [26-28] ตัวอย่างการศึกษาของ JameeShahwan, A และคณะในปี 2019 ทำการศึกษาทางระบาดวิทยาเพื่อทดสอบปัจจัยเสี่ยงต่อการเกิดโรคหลอดเลือดหัวใจและโรคหลอดเลือดสมอง[29] การศึกษาพบว่าการเป็นโรคความดันโลหิตสูงมีโอกาสเสี่ยงทำให้เกิดโรคหลอดเลือดหัวใจสูงเมื่อเทียบกับคนที่ไม่เป็นโรคความดันโลหิตสูง โดยมีการควบคุมตัวแปร อายุ จากผลการศึกษาดังกล่าวเมื่อทำการแจกแจงความถี่ข้อมูลความสัมพันธ์ระหว่างโรคความดันโลหิตสูงกับโรคหลอดเลือดหัวใจสูงแบ่งตามกลุ่มอายุได้ดังตารางที่ 4.6

ตารางที่ 4.6 ค่าความเสี่ยงสัมพัทธ์ระหว่างความสัมพันธ์ระหว่างการเป็นโรคความดันโลหิตสูงกับการเป็นโรคหลอดเลือดหัวใจสูงแบ่งตามกลุ่มอายุ

กลุ่มอายุ (ปี)	ค่าความเสี่ยงสัมพัทธ์	ช่วงเชื่อมั่นที่ 95%
25-34	3.89	0.39, 19.73
35- 44	3.12	0.90, 9.79
45-54	2.25	1.12, 4.45
55-64	4.76	2.59, 9.07
≥65	1.59	0.89, 2.88
Crude odds ratio	4.72	3.50, 6.39
M-H combined (Adjusted)	2.66	1.95, 3.64

ที่มา: ดัดแปลงมาจากJameeShahwan A, Abed Y, Desormais I, Magne J, Preux PM, Aboyans V, et al. Epidemiology of coronary artery disease and stroke and associated risk factors in Gaza community -Palestine. PLoS One. 2019;14(1):e0211131. S1Dataset. Databaseofstudy participants[30]

จากตารางที่ 4.6 พบว่าค่า Crude odds ratio (ค่าความเสี่ยงสัมพัทธ์ที่ยังไม่ได้ควบคุมตัวแปรใด) มีค่าเท่ากับ 4.72 เมื่อทำการควบคุมตัวแปรกลุ่มอายุแล้ว พบว่า Adjusted odds ratio (ค่าความเสี่ยงสัมพัทธ์ที่มีการควบคุมตัวแปรอื่นร่วมด้วย) มีค่าเท่ากับ 2.66 เห็นได้ว่าเมื่อทำการควบคุมตัวแปรกลุ่มอายุแล้วค่า Adjusted odds ratio เปลี่ยนไป 43.6% (4.72-2.66 / 4.72) ของค่า Crude odds ratio ลักษณะเช่นนี้เรียกว่าเป็นลักษณะของ Confounding effect

การลดหรือควบคุมอคติตัวแปรกวนสามารถทำได้ใน 2 ขั้นตอน คือ

1. ขั้นตอน Design งานวิจัย
   - 1.1 ใช้รูปแบบการศึกษาแบบ Match case-control study
   - 1.2 Stratifiedrandom sampling
2. ขั้นตอนการวิเคราะห์
   - 2.1 ใช้ Statistical modeling ในการควบคุมตัวแปรที่คาดว่าจะเป็น confounder เช่น Multiple logistic regression หรือ Multiple linear regression เป็นต้น

4.4 บทสรุป

อคติในงานวิจัยมี 3 ประเภท คือ Selection bias, Information bias และ Confounding effect ทั้งนี้นักวิจัยสามารถลดหรือควบคุมอคติต่างๆ โดยเริ่มตั้งแต่การวางแผนออกแบบงานวิจัย การเก็บข้อมูล และการวิเคราะห์ข้อมูล เพื่อให้ได้ผลลัพธ์ของงานวิจัยมีอคติเกิดขึ้นน้อยที่สุด

4.5 แบบฝึกหัด

1. อคติที่เกิดขึ้นในงานวิจัยสามารถกำจัดได้หรือไม่
2. การควบคุมปัจจัยกวนสามารถทำได้ ในขั้นตอนไหนบ้าง
3. การลดอคติ Misclassification สามารถทำได้อย่างไร
4. ถ้ามีปัจจัยที่เป็น Confounding factors หลายตัวแปรสามารถควบคุมตัวแปรเหล่านั้นได้อย่างไร
5. การลดอคติจากการลืม การสัมผัสปัจจัยเสี่ยงในอดีต การสัมภาษณ์หรือตอบแบบสอบถามด้วยตนเองสามารถทำได้อย่างไร

4.6 เอกสารอ้างอิง

1. Tripepi G, Jager KJ, Dekker FW, Zoccali C. Selection bias and information bias in clinical research. Nephron Clin Pract. 2010;115(2):c94-9.
2. Haneuse S. Distinguishing Selection Bias and Confounding Bias in Comparative Effectiveness Research. Med Care. 2016;54(4):e23-9.
3. Mark DH. Interpreting the term selection bias in medical research. Fam Med. 1997;29(2):132-6.
4. Cheung KL, Ten Klooster PM, Smit C, de Vries H, Pieterse ME. The impact of non-response bias due to sampling in public health studies: A comparison of voluntary versus mandatory recruitment in a Dutch national survey on adolescent health. BMC Public Health. 2017;17(1):276.
5. Maitland A, Lin A, Cantor D, Jones M, Moser RP, Hesse BW, et al. A Nonresponse Bias Analysis of the Health Information National Trends Survey (HINTS). J Health Commun. 2017;22(7):545-53.
6. Jager KJ, Tripepi G, Chesnaye NC, Dekker FW, Zoccali C, Stel VS. Where to look for the most frequent biases? Nephrology (Carlton). 2020;25(6):435-41.

7. Howe CJ, Cole SR, Lau B, Napravnik S, Eron JJ, Jr. Selection Bias Due to Loss to Follow Up in Cohort Studies. Epidemiology. 2016;27(1):91-7.
8. Dettori JR. Loss to follow-up. Evid Based Spine Care J. 2011;2(1):7-10.
9. Kristman V, Manno M, Côté P. Loss to follow-up in cohort studies: how much is too much? Eur J Epidemiol. 2004;19(8):751-60.
10. Jordan S, Watkins A, Storey M, Allen SJ, Brooks CJ, Garaiova I, et al. Volunteer bias in recruitment, retention, and blood sample donation in a randomised controlled trial involving mothers and their children at six months and two years: a longitudinal analysis. PLoS One. 2013;8(7):e67912.
11. Callahan CA, Hojat M, Gonnella JS. Volunteer bias in medical education research: an empirical study of over three decades of longitudinal data. Med Educ. 2007;41(8):746-53.
12. Amori G, Lenox RH. Do volunteer subjects bias clinical trials? J Clin Psychopharmacol. 1989;9(5):321-7.
13. Tripepi G, Jager KJ, Dekker FW, Wanner C, Zoccali C. Bias in clinical research. Kidney Int. 2008;73(2):148-53.
14. Mumford SL, Schisterman EF, Cole SR, Westreich D, Platt RW. Time at Risk and Intention-to-treat Analyses Parallels and Implications for Inference. Epidemiology. 2015;26(1):112-8.
15. Gravel J, Opatrny L, Shapiro S. The intention-to-treat approach in randomized controlled trials: Are authors saying what they do and doing what they say? Clinical Trials. 2007;4(4):350-6.
16. Abraha I, Cozzolino F, Orso M, Marchesi M, Germani A, Lombardo G, et al. A systematic review found that deviations from intention-to-treat are common in randomized trials and systematic reviews. J Clin Epidemiol. 2017;84:37-46.
17. Jordan S, Watkins A, Storey M, Allen SJ, Brooks CJ, Garaiova I, et al. Volunteer Bias in Recruitment, Retention, and Blood Sample Donation in a Randomised Controlled Trial Involving Mothers and Their Children at Six Months and Two Years: A Longitudinal Analysis. Plos One. 2013;8(7).

18. Beer C, Flicker L, Horner B, Bretland N, Scherer S, Lautenschlager NT, et al. Factors Associated with Self and Informant Ratings of the Quality of Life of People with Dementia Living in Care Facilities: A Cross Sectional Study. Plos One. 2010;5(12).
19. Kesmodel US. Information bias in epidemiological studies with a special focus on obstetrics and gynecology. Acta Obstet Gynecol Scand. 2018;97(4):417-23.
20. Szklo M, Nieto FJ. Epidemiology: Beyond the Basics: Jones & Bartlett Learning; 2014.
21. Gordis L. Epidemiology: Elsevier/Saunders; 2009.
22. Charuluxananan S, Suraseranivongse S, Punjasawadwong Y, Somboonviboon W, Sriswasdi S, Pranootnarabhal T, et al. Risk factors of intraoperative oxygen desaturation: a case-control study of 152,314 anesthetics. J Med Assoc Thai. 2007;90(11):2359-65.
23. Höfler M. The effect of misclassification on the estimation of association: a review. Int J Methods Psychiatr Res. 2005;14(2):92-101.
24. Brenner H, Blettner M. Misclassification bias arising from random error in exposure measurement: implications for dual measurement strategies. Am J Epidemiol. 1993;138(6):453-61.
25. White E. The effect of misclassification of disease status in follow-up studies: implications for selecting disease classification criteria. Am J Epidemiol. 1986;124(5):816-25.
26. Jager KJ, Zoccali C, Macleod A, Dekker FW. Confounding: what it is and how to deal with it. Kidney Int. 2008;73(3):256-60.
27. Kahlert J, Gribsholt SB, Gammelager H, Dekkers OM, Luta G. Control of confounding in the analysis phase - an overview for clinicians. Clin Epidemiol. 2017;9:195-204.
28. Nørgaard M, Ehrenstein V, Vandenbroucke JP. Confounding in observational studies based on large health care databases: problems and potential solutions - a primer for the clinician. Clin Epidemiol. 2017;9:185-93.
29. Guetterman TC. Basics of statistics for primary care research. Fam Med Community Health. 2019;7(2).

30. Kaliyadan F, Kulkarni V. Types of Variables, Descriptive Statistics, and Sample Size. Indian Dermatol Online J. 2019;10(1):82-6.

บทที่ 5
ชีวสถิติเบื้องต้น

จากบทที่ 2-4 ได้กล่าวถึงรูปแบบงานวิจัยและอคติที่อาจเกิดขึ้นในงานวิจัย ในบทนี้จะได้กล่าวถึงการประยุกต์ใช้ชีวสถิติในงานวิจัยด้านวิทยาศาสตร์สุขภาพสถิติแบ่งออกเป็น 2 ประเภทหลักๆ คือสถิติเชิงพรรณนา (Descriptive statistics) และสถิติเชิงอนุมาน (Inferential statistics)[1-5]

สถิติเชิงพรรณนา (Descriptive statistics) หมายถึงวิธีการทางสถิติที่ใช้พรรณนาลักษณะของข้อมูลที่ต้องการศึกษาเช่น อายุของคนที่อาศัยอยู่ในจังหวัดมหาสารคาม เป็นเท่าใด จึงหาค่ามาเพื่อพรรณนาข้อมูลอายุของคนที่อาศัยอยู่ในจังหวัดมหาสารคามสถิติเชิงพรรณนาจะประกอบไปด้วย ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ค่าความถี่ ร้อยละ เป็นต้น ทั้งนี้การพรรณนาด้วยค่าเฉลี่ยหรือค่าความถี่ ขึ้นอยู่กับมาตราวัดและการแจกแจงของข้อมูล ในกรณีอายุจะใช้ค่า ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน (กรณีข้อมูลมีการแจกแจงปกติ) มาพรรณนาข้อมูลดังนี้

สถิติเชิงอนุมาน (Inferential statistics) คือ วิธีการทางสถิติที่ใช้ทฤษฎีความน่าจะเป็นในการประมาณค่าของข้อมูลที่เก็บมาได้ (ข้อมูลที่เก็บมาได้เรียกว่า ตัวอย่าง) ในการประมาณค่าไปสู่ประชากร เช่น อายุของคนที่อาศัยอยู่ในจังหวัดมหาสารคามโดยเฉลี่ยแล้วมีอายุเท่าไหร่ ถ้าเราทำการเก็บข้อมูลทุกคนในจังหวัดจะเป็นการสิ้นเปลืองทั้งงบประมาณ คน และเวลา จึงทำการสุ่มตัวอย่างมาจำนวนหนึ่งแล้วทำการประมาณค่าเฉลี่ยของอายุ (คำนวณจากตัวอย่างที่เก็บมาได้) เพื่อเป็นตัวแทนของประชากรของจังหวัด

ในงานวิจัยทางด้านวิทยาศาสตร์สุขภาพมีตัวแปรต่างๆในการเก็บข้อมูล โดยสามารถแบ่งออกโดยละเอียด 4 กลุ่มคือ[6-8]

5.1 มาตรานามบัญญัติ

มาตราวัดแบบนามบัญญัติ (Nominal scale) ลักษณะข้อมูลเป็นกลุ่มเช่น ตัวแปร เพศ แบ่งออกเป็น เพศชาย และเพศหญิง ลักษณะดังกล่าวไม่สามารถนำข้อมูลมา บวก ลบ คูณ หารกันได้ และไม่สามารถบอกได้ถึงระดับความแตกต่างของข้อมูลเมื่อสลับตำแหน่งกัน

5.2 มาตราอันดับ

มาตราวัดแบบอันดับ (Ordinal scale) เป็นตัวแปรที่มีข้อมูลลักษณะเป็นกลุ่มเรียงอันดับหรือระดับได้ สามารถบอกถึงระดับความรุนแรงของโรคหรือบอกถึงความมาก น้อยของมาตราวัดเมื่อสลับตำแหน่งกันได้ เช่นระดับความรุนแรงของความเจ็บปวด โดยแบ่งสเกลการวัดตั้งแต่ 0, 1, 2, 3, 4, 5, 6,7, 8, 9, 10 โดยที่ หมายเลข 0 หมายถึงไม่เจ็บปวดเลย หมายเลข 5 หมายถึง ปวด ระดับปานกลาง หมายเลข 10 หมายถึง ปวดมากที่สุด[9-11] เป็นต้น จากตัวอย่างดังกล่าวเห็นได้ว่าระดับของข้อมูล สามารถระบุได้จำนวน น้อย หรือ มากได้ ถ้ามีการสลับตำแหน่งกัน เช่น หมายเลข 10 เทียบกับ หมายเลข 2 ก็สรุปได้ว่าระดับความปวด ที่ระดับ 10 มีค่ามากกว่า ระดับ 2

5.3 มาตราอันตรภาค

มาตราวัดแบบอันตรภาค (Interval scale) มาตราวัดนี้ไม่มีศูนย์แท้ (หมายถึงค่าของข้อมูลมีค่าเป็น 0 แต่ยังมีความหมายของข้อมูล) และมีหน่วยการวัดที่เท่ากัน ข้อมูลชนิดนี้สามารถนำไปบวก ลบ คูณ หารได้

5.4 มาตราอัตราส่วน

ลักษณะของข้อมูลการวัดแบบอัตราส่วน (Ratio scale) สามารถเรียงระดับโดยการแบ่งเป็นช่วงเท่าๆ กัน เช่น ระยะทาง เวลา น้ำหนัก ส่วนสูง อายุ เป็นต้น

อย่างไรก็ตามเราสามารถแบ่งตามลักษณะข้อมูลดังกล่าวได้เป็น 2 ประเภทใหญ่ๆ ดังนี้

1. ข้อมูลจำแนกประเภท (Categorical data)[12-14] เป็นข้อมูลที่มีประเภทตัวแปรกลุ่ม เช่น เพศ ระดับการศึกษา การเป็นโรค (เป็นโรค และ ไม่เป็นโรค) การติดเชื้อ (ติดเชื้อ และ ไม่ติดเชื้อ) ระดับความคิดเห็น (ไม่เห็นด้วย ปานกลาง เห็นด้วย) เป็นต้น ทั้งนี้ลักษณะของตัวแปรกลุ่มหมายถึงข้อมูล มาตรานามบัญญัติ และ มาตราอันดับ
2. ข้อมูลแบบต่อเนื่อง (Continuous data)[2, 15] เป็นข้อมูลที่มีค่าหรือลักษณะแบบต่อเนื่อง เช่น รายได้ อายุ ความสูง อายุครรภ์ เป็นต้น โดยข้อมูลแบบต่อเนื่องหมายถึงข้อมูลที่มีมาตราวัดแบบอันตรภาคและมาตราอัตราส่วนตัวอย่างลักษณะข้อมูลประเภทต่างๆ ดังแสดงใน ตารางที่ 5.1

ตารางที่ 5.1 ตัวอย่างการจำแนกตัวแปรตามมาตราต่างๆ

ชื่อตัวแปร	ความหมาย	ระดับ	มาตราวัด
age	อายุ	ตัวแปรต่อเนื่อง	มาตราอัตราส่วน
sex	เพศ	1=ผู้ชาย 2=ผู้หญิง	มาตรานามบัญญัติ
cvd	โรงหลอดเลือดและหัวใจ	0=ไม่เป็น 1=เป็น	มาตรานามบัญญัติ
bmi	ดัชนีมวลกาย	0= ดัชนีมวลกายน้อยกว่าหรือเท่ากับ 30 1=ดัชนีมวลกายมากกว่า 30	มาตราอันดับ
prevchd	ประวัติการเป็นโรคโรงหลอดเลือดและหัวใจ	0=ไม่เป็น 1=เป็น	มาตรานามบัญญัติ
cursmoke	สถานะการสูบบุหรี่	0=ไม่สูบ 1=สูบ	มาตรานามบัญญัติ

5.5 การตั้งสมมุติฐานทางสถิติ

การตั้งสมมุติฐานทางสถิติเริ่มจากคำถามงานวิจัยที่ต้องการศึกษา แล้วจึงนำข้อสงสัยดังกล่าวมาพิสูจน์ในวิชาชีวสถิติเรียกว่า สมมุติฐานทางสถิติ และการพิสูจน์ข้อสงสัยดังกล่าวเรียกว่าการทดสอบสมมุติฐาน ในการทดสอบสมมุติฐานประกอบไปด้วยการคำนวณขนาดตัวอย่าง วางแผนเก็บข้อมูล การเก็บข้อมูลอย่างมีระเบียบแบบแผน และการวิเคราะห์ข้อมูลเพื่อประมาณค่าพารามิเตอร์และทดสอบสมมุติฐานทางสถิติ

ในการตั้งสมมุติฐานงานวิจัยไม่สามารถพิสูจน์ให้แน่ชัดหรือยืนยันได้ว่าการตั้งสมมุติฐานงานวิจัยที่ตั้งขึ้นมาถูกต้องหรือไม่ เป็นข้อจริงหรือข้อเท็จ[16-19] ดังนั้นสามารถเขียนเป็นข้อเท็จจริงได้ดังนี้

ตารางที่ 5.2 ข้อเท็จและข้อจริงของการตั้งสมมุติฐานและการตัดสินใจทางด้านสถิติ

ความจริง	การตัดสินใจ	
	ยอมรับ H_0	ปฏิเสธ H_0
H_0เป็นจริง	ตัดสินใจถูก	เกิดความผิดพลาดชนิดที่ 1 (Type I error)
H_0ไม่จริง	เกิดความผิดพลาดชนิดที่ 2 (Type II error)	ตัดสินใจถูก

จากตารางที่ 5.1 สามารถอธิบายความหมายความเป็นไปได้ 4 กรณี คือ

กรณีที่ 1 ถ้าความจริง H_0 เป็นจริง แล้วผลการศึกษาวิจัยเราตัดสินใจ ยอมรับ H_0 เหตุการณ์นี้สามารถกล่าวได้ว่าการตัดสินใจในงานวิจัยของเรา เป็นการตัดสินใจที่ถูกต้อง

กรณีที่ 2. ถ้าความจริง H_0 เป็นจริง แล้วผลการศึกษาวิจัยเราตัดสินใจ ปฏิเสธ H_0 เหตุการณ์นี้สามารถระบุได้ว่า การตัดสินใจในงานวิจัยของเรา เป็นการตัดสินใจผิดพลาด ซึ่งเหตุการณ์นี้เรียกความผิดพลาดนี้ว่าเป็นความผิดพลาดแบบชนิดที่ 1 (Type I error) หรือเรียกว่า α error

กรณีที่ 3. ถ้าความจริง H_0 เป็นเท็จ แล้วผลการศึกษาวิจัยเราตัดสินใจ ยอมรับ H_0 เหตุการณ์นี้สามารถระบุได้ว่า เป็นการตัดสินใจผิดพลาด ซึ่งเหตุการณ์นี้เรียกความผิดพลาดนี้ว่าเป็นความผิดพลาดแบบชนิดที่ 2 (Type II error) หรือเรียกว่า β error

กรณีที่ 4. ถ้าความจริง H_0 เป็นเท็จ แล้วผลการศึกษาวิจัยเราตัดสินใจ ปฏิเสธ H_0 เหตุการณ์นี้สามารถระบุได้ว่า เป็นการตัดสินใจที่ถูกต้อง หรือเรียกว่าเป็นอำนาจของการทดสอบสมมุติฐานที่ 1- β เช่น ถ้ากำหนดเกณฑ์การเกิดความผิดพลาดชนิดที่ 2 ไว้ไม่เกิน 20% หรือ เท่ากับ 0.2 ดังนั้นในการศึกษานี้จะมีอำนาจของการทดสอบสมมุติฐานเท่ากับ 1-0.2=0.8 หรือเท่ากับ 80%

ในการตั้งสมมุติฐานทางสถิติสามารถตั้งสมมุติฐานได้ทั้งแบบด้านเดียว (One-tailed) และแบบสองด้าน (Two-tailed)[20, 21]

5.5.1 การตั้งสมมุติฐานแบบสองด้าน

ตัวอย่างเช่นต้องการเปรียบเทียบค่าเฉลี่ยของระดับคอเลสเตอรอลระหว่างเพศชายและเพศหญิงว่ามีความแตกต่างกันหรือไม่ สามารถเขียนคำถามงานวิจัยและสมมุติฐานได้ดังนี้

คำถามงานวิจัย: ค่าเฉลี่ยของระดับคอเลสเตอรอลระหว่างเพศชาย (μ_1) กับเพศหญิง (μ_2) มีค่าต่างกันหรือไม่

สมมุติฐานงานวิจัย: ค่าเฉลี่ยของคอเลสเตอรอลระหว่างเพศชายกับหญิงมีค่าแตกต่างกัน

สมมุติฐานงานทางสถิติ:

H_0: ค่าเฉลี่ยของระดับคอเลสเตอรอลระหว่างเพศชายกับหญิงมีค่าไม่แตกต่างกัน

H_a: ค่าเฉลี่ยของระดับคอเลสเตอรอลระหว่างเพศชายกับหญิงมีค่าแตกต่างกัน

หรือ

H_0: $\mu_1 = \mu_2$

H_a: $\mu_1 \neq \mu_2$

โดยที่ μ_1 หมายถึง ค่าเฉลี่ยของระดับคอเลสเตอรอลในกลุ่มประชากรเพศชาย

μ_2 หมายถึง ค่าเฉลี่ยของระดับคอเลสเตอรอลในกลุ่มประชากรเพศหญิง

5.5.2 การตั้งสมมุติฐานแบบด้านเดียว

การตั้งสมมุติฐานงานวิจัยแบบด้านเดียวสามารถตั้งได้ 2 แบบ คือ

5.5.2.1. การตั้งสมมุติฐานงานวิจัยแบบมากกว่า

ตัวอย่าง

คำถามงานวิจัย: ค่าเฉลี่ยของระดับคอเลสเตอรอลของ**เพศชายมีค่ามากกว่าเพศหญิงหรือไม่**

สมมุติฐานงานวิจัย: ค่าเฉลี่ยของระดับคอเลสเตอรอลของ**เพศชายมีค่ามากกว่าเพศหญิง**

สมมุติฐานทางสถิติ:

H_0: ค่าเฉลี่ยของระดับคอเลสเตอรอลเพศชาย มีค่า น้อยกว่าหรือเท่ากับ ($\leq$) เพศหญิง

H_a: ค่าเฉลี่ยของระดับคอเลสเตอรอลเพศชาย มีค่ามากกว่า เพศหญิง

หรือ

$H_0: \mu_1 \leq \mu_2$

$H_a: \mu_1 > \mu_2$

โดยที่ μ_1 หมายถึง ค่าเฉลี่ยของระดับคอเลสเตอรอลในกลุ่มประชากรเพศชาย

μ_2 หมายถึง ค่าเฉลี่ยของระดับคอเลสเตอรอลในกลุ่มประชากรเพศหญิง

5.5.2.2. การตั้งสมมุติฐานงานวิจัยแบบน้อยกว่า

ตัวอย่าง

คำถามงานวิจัย: ค่าเฉลี่ยของระดับคอเลสเตอรอลของเพศ**ชายมีค่าน้อยกว่าเพศหญิงหรือไม่**

สมมุติฐานงานวิจัย: ค่าเฉลี่ยของระดับคอเลสเตอรอลของ**เพศชายมีค่าน้อยกว่าเพศหญิง**

สมมุติฐานงานทางสถิติ:

H_0: ค่าเฉลี่ยของระดับคอเลสเตอรอลเพศชาย มีค่า มากกว่าหรือเท่ากับ (≥) เพศหญิง

H_a: ค่าเฉลี่ยของคอเลสเตอรอลของเพศชาย มีค่าน้อยกว่า เพศหญิง

หรือ

$H_0: \mu_1 \geq \mu_2$

$H_a: \mu_1 < \mu_2$

โดยที่ μ_1 หมายถึง ค่าเฉลี่ยของระดับคอเลสเตอรอลในกลุ่มประชากรเพศชาย

μ_2 หมายถึง ค่าเฉลี่ยของระดับคอเลสเตอรอลในกลุ่มประชากรเพศหญิง

5.6 บทสรุป

ชีวสถิติเป็นศาสตร์ของสถิติที่นำมาประยุกต์ใช้ในงานวิจัยด้านชีวภาพ สุขภาพ และการแพทย์ในงานวิจัยจะมีการระบุตัวแปรที่ต้องการศึกษาประกอบไปด้วยมาตราวัดแบบต่าง ๆ ในบทต่อไปเป็นการแนะนำโปรแกรมทางสถิติ ที่ได้นำมาใช้วิเคราะห์ข้อมูลในตำราเล่มนี้

5.7 แบบฝึกหัด

1. ชีวสถิติหมายถึงอะไร
2. ข้อมูลที่มีมาตราวัดอันตรภาคชั้นหรืออัตราส่วนสามารถนำข้อมูลเป็นแบบนามบัญญัติหรืออันดับได้หรือไม่
3. ในงานวิจัยทำไมจึงต้องทำการสุ่มตัวอย่าง
4. สมมุติว่า ถ้าเรามีข้อมูลประชากรที่เราต้องการศึกษาทั้งหมด จะสามารถใช้ค่าสถิติในการรายงานผลเหมาะสมหรือไม่
5. ระดับความเจ็บปวด (Visual analog scale) มีค่าตั้งแต่ 0 ถึง 10 โดยที่ 0 หมายถึงไม่เจ็บปวดเลย และ 10 เจ็บปวดมากที่สุด มีมาตราวัดเป็นแบบใด

5.8 เอกสารอ้างอิง

1. Dettori JR, Norvell DC. The Anatomy of Data. Global Spine J. 2018;8(3):311-3.
2. Hazra A, Gogtay N. Biostatistics Series Module 2: Overview of Hypothesis Testing. Indian J Dermatol. 2016;61(2):137-45.
3. Hazra A, Gogtay N. Biostatistics Series Module 1: Basics of Biostatistics. Indian J Dermatol. 2016;61(1):10-20.
4. Manja V, Lakshminrusimha S. Principles of Use of Biostatistics in Research. Neoreviews. 2014;15(4):e133-e50.
5. Rosner B. Fundamentals of biostatistics: Seventh edition. Boston : Brooks/Cole, Cengage Learning; 2011.
6. Mishra P, Pandey CM, Singh U, Gupta A. Scales of measurement and presentation of statistical data. Ann Card Anaesth. 2018;21(4):419-22.
7. Sawamura J, Morishita S, Ishigooka J. Interpretation for scales of measurement linking with abstract algebra. J Clin Bioinforma. 2014;4:9.

8. Simundić AM. [Types of variables and distributions]. Acta Med Croatica. 2006;60 Suppl 1:17-35.
9. Haefeli M, Elfering A. Pain assessment. Eur Spine J. 2006;15 Suppl 1(Suppl 1):S17-24.
10. Karcioglu O, Topacoglu H, Dikme O, Dikme O. A systematic review of the pain scales in adults: Which to use? Am J Emerg Med. 2018;36(4):707-14.
11. van Zanden JE, Wagenaar S, Ter Maaten JM, Ter Maaten JC, Ligtenberg JJM. Pain score, desire for pain treatment and effect on pain satisfaction in the emergency department: a prospective, observational study. BMC Emerg Med. 2018;18(1):40.
12. Scott M, Flaherty D, Currall J. Statistics: dealing with categorical data. J Small Anim Pract. 2013;54(1):3-8.
13. Greenfield ML, Wojtys EM, Kuhn JE. A statistics primer. Tests for continuous data. Am J Sports Med. 1997;25(6):882-4.
14. Bentley TG, Weinstein MC, Kuntz KM. Effects of categorizing continuous variables in decision-analytic models. Med Decis Making. 2009;29(5):549-56.
15. Biau DJ, Jolles BM, Porcher R. P value and the theory of hypothesis testing: an explanation for new researchers. Clin Orthop Relat Res. 2010;468(3):885-92.
16. Sedgwick P. Pitfalls of statistical hypothesis testing: type I and type II errors. BMJ. 2014;349:g4287.
17. Banerjee A, Chitnis UB, Jadhav SL, Bhawalkar JS, Chaudhury S. Hypothesis testing, type I and type II errors. Ind Psychiatry J. 2009;18(2):127-31.
18. Shreffler J, Huecker MR. Hypothesis Testing, P Values, Confidence Intervals, and Significance. StatPearls. Treasure Island (FL)2020.
19. Ludbrook J. Should we use one-sided or two-sided P values in tests of significance? Clin Exp Pharmacol Physiol. 2013;40(6):357-61.
20. Wolterbeek R. One and two sided tests of significance. Statistical hypothesis should be brought into line with clinical hypothesis. BMJ. 1994;309(6958):873-4.
21. Bland JM, Altman DG. One and two sided tests of significance. BMJ. 1994;309(6949):248.

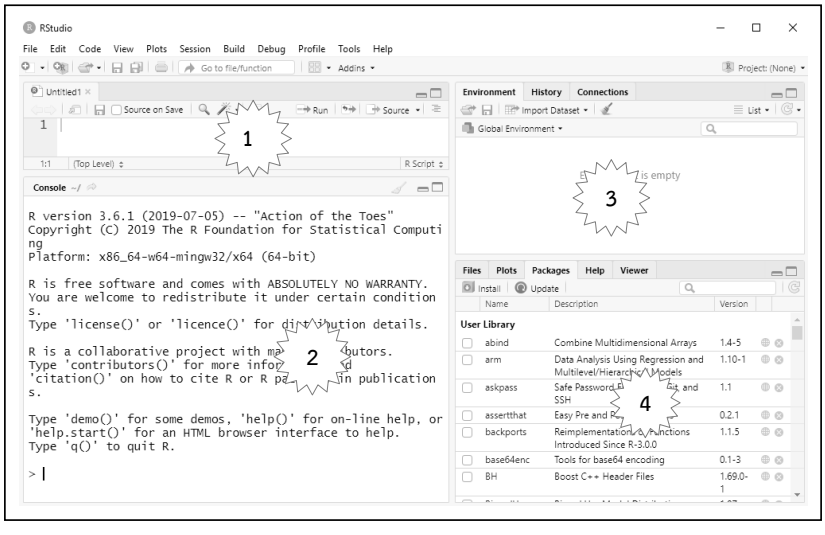
บทที่ 6

การใช้โปรแกรม R และชุดคำสั่งเบื้องต้น

จากบทที่ 5 เราได้ทราบถึงชีวสถิติเบื้องต้น มาตราวัดของตัวแปร การตั้งสมมุติฐานทางสถิติ ในบทนี้จะได้ศึกษาเกี่ยวกับโปรแกรมทางสถิติ ซึ่งจะถูกนำมาใช้ตลอดในตำราเล่มนี้ โปรแกรม R[1] เป็นโปรแกรมหนึ่งที่ใช้ในการวิเคราะห์ข้อมูลทางสถิติซึ่งจะมีชุดคำสั่งพื้นฐานสำหรับการวิเคราะห์ข้อมูลเป็นโปรแกรมที่เปิดให้ใช้งานฟรีทั้งนี้ยังเปิดโอกาสให้นักพัฒนาโปรแกรมได้พัฒนาชุดคำสั่งเพื่อนำมาใช้งานในการวิเคราะห์ข้อมูลเพื่อให้มีความง่ายขึ้นตามลักษณะการใช้งานต่างๆ เช่น ชุดคำสั่งชื่อ epiDisplay[2] พัฒนามาเพื่อการวิเคราะห์ข้อมูลทางวิทยาการระบาดหรือชุดคำสั่งชื่อ ggplot2 มีวัตถุประสงค์ในการวาดกราฟ[3] ในบทนี้จะได้ศึกษาถึงการติดตั้งโปรแกรมและการจัดการข้อมูลเบื้องต้นซึ่งเป็นส่วนสำคัญในการเริ่มใช้งานวิเคราะห์ข้อมูล

6.1 การติดตั้งโปรแกรม

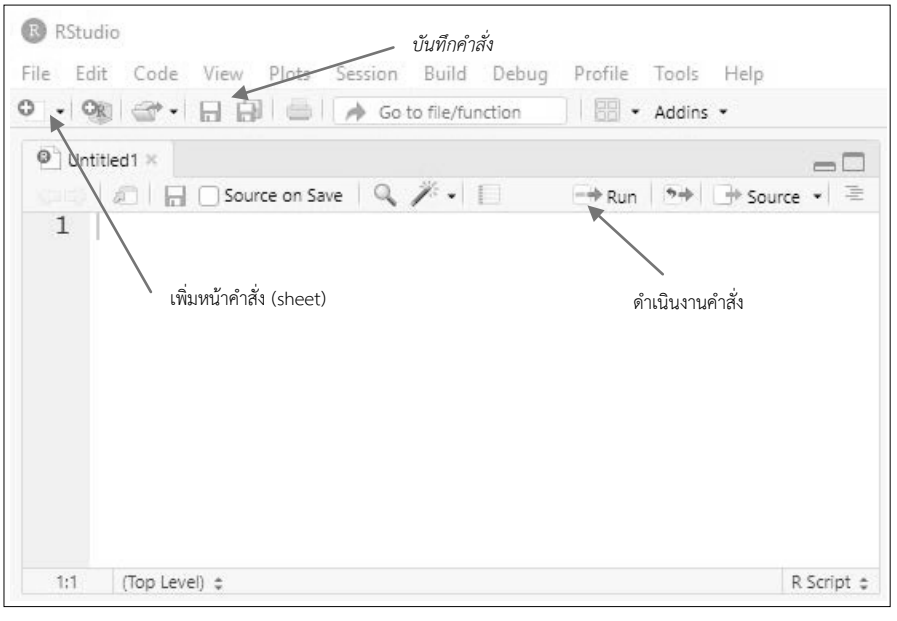
การติดตั้งโปรแกรม R สามารถดาวน์โหลดโปรแกรมได้จากเว็บไซต์ https://cran.r-project.org/bin/windows/base/ และดาวน์โหลดโปรแกรม RStudio[4] ได้จากเว็บไซต์ https://rstudio.com/products/rstudio/download/จากนั้นให้ทำการติดตั้งโปรแกรม R และ โปรแกรม RStudio เมื่อทำการติดตั้งโปรแกรมทั้งสองแล้วให้เปิดโปรแกรม RStudio ขึ้นมาเพื่อเริ่มใช้งานในการวิเคราะห์ข้อมูล



รูปที่ 6.1 หน้าต่างของโปรแกรม RStudio

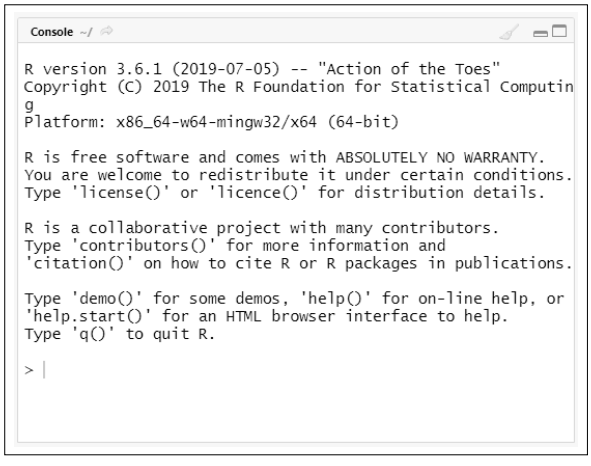
จากรูปที่ 6.1 มีหน้าต่างย่อยๆ 4 หน้าต่าง ได้แก่ Source (หมายเลข 1) Console (หมายเลข 2) Environment (หมายเลข 3) และ File (หมายเลข 4)

หมายเลข 1 Source เป็นหน้าต่างที่ใช้สำหรับการเขียนคำสั่ง (Command) และสามารถดำเนินงาน (run) คำสั่งได้จากหน้าต่างนี้ การดำเนินงานคำสั่งสามารถทำได้ทีละบรรทัดหรือหลายบรรทัด การดำเนินงานทีละบรรทัดทำได้โดยการทำการเน้นคำสั่งที่จะดำเนินงาน (Highlight)หรือ ให้ตัวชี้ตำแหน่ง (cursor) กระพริบอยู่ในบรรทัดที่จะดำเนินงานคำสั่ง จากนั้นให้คลิกตรงคำว่า “RUN” (หรือ Ctrl+Enter) โปรแกรมก็จะรันคำสั่งนั้นให้เราและเราสามารถทำการเพิ่มหน้าคำสั่งและบันทึกคำสั่งในหน้านั้นๆ ได้ดังรูปที่ 6.2



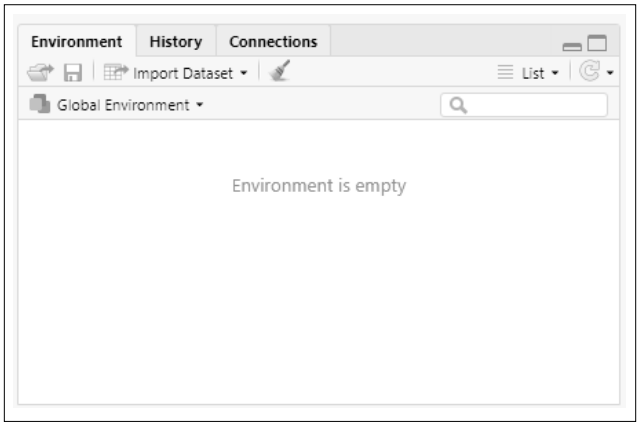
รูปที่ 6.2 หน้าต่าง Source

หมายเลข 2 Console หน้าต่าง Console เป็นหน้าต่างที่แสดงผลการดำเนินงานคำสั่งของโปรแกรมคล้ายๆ กับหน้าต่างของโปรแกรม R ในที่นี้ได้ใช้ Rstudio เพื่อให้การใช้งานโปรแกรมง่ายขึ้น ในหน้าต่างนี้จะบอกเวอร์ชั่นของโปรแกรม R ที่ใช้อยู่ ดังรูปที่ 6.3



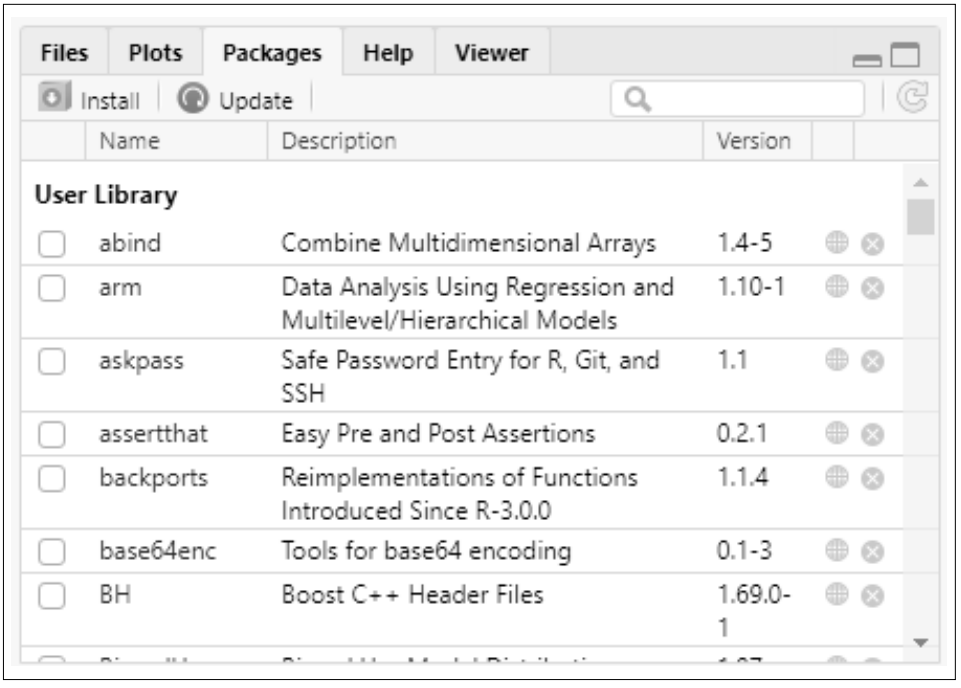
รูปที่ 6.3 หน้าต่าง Console

หมายเลข 3 Environment เป็นหน้าต่างที่แสดง Environment history และ Connection ทั้งนี้ หน้าต่างนี้ แสดงข้อมูล (Dataframe) และ Values ต่างๆ ดังรูปที่ 6.4



รูปที่ 6.4 หน้าต่าง Environment history และ connection

หมายเลข 4 File เป็นหน้าต่างที่แสดง Files, Plot (แสดงกราฟ), Packages (ใช้สำหรับติดตั้งชุดคำสั่งหรือ package), Help (สำหรับดูการใช้คำสั่ง) และ Viewers ดังรูปที่ 6.5



รูปที่ 6.5 Files, Plot, Packages, Help และ Viewers

6.2 การนำเข้าข้อมูล

การนำเข้าข้อมูลมีหลายรูปแบบในตำรานี้แบ่งออกเป็น 2 แบบดังนี้

6.2.1 การนำเข้าข้อมูลโดยการสร้าง Data จากค่าของข้อมูล (Values)

การสร้าง Data (Dataframe) จาก Values วิธีนี้เหมาะกับข้อมูลที่มีไม่มากนัก เช่น การเปรียบเทียบค่าดัชนีมวลกายระหว่างกลุ่มที่ดื่มเครื่องดื่มแอลกอฮอล์มีข้อมูลดัง ตารางที่ 6.1

ตารางที่ 6.1 ข้อมูลตัวอย่างดัชนีมวลกาย

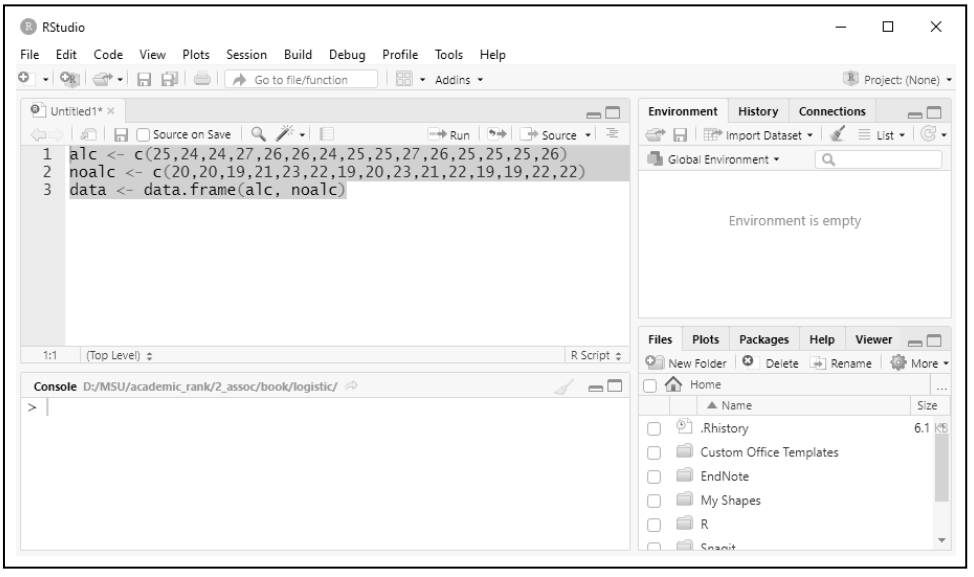
กลุ่มดื่ม	25	24	24	27	26	26	24	25	25	27	26	25	25	25	26
กลุ่มไม่ดื่ม	20	20	19	21	23	22	19	20	23	21	22	19	19	22	22

จากข้อมูลในตารางที่ 6.1 สามารถเขียนเป็นคำสั่งในโปรแกรม R ดังนี้

```
alc<- c(25,24,24,27,26,26,24,25,25,27,26,25,25,25,26)
noalc<- c(20,20,19,21,23,22,19,20,23,21,22,19,19,22,22)
data <- data.frame(alc, noalc)
```

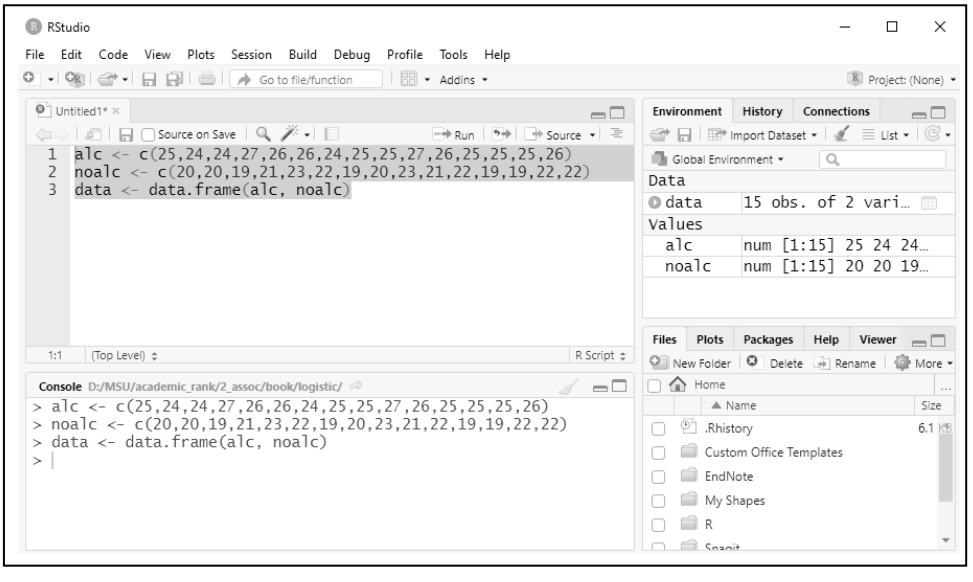
รูปที่ 6.6 ข้อมูลตัวอย่างดัชนีมวลกายแยกตามกลุ่มดื่มกับไม่ดื่มเครื่องดื่มแอลกอฮอล์

จากนั้นทำการเขียนคำสั่งใน Source ดังรูปที่ 6.7



รูปที่ 6.7 การเขียน Command ในหน้าต่าง Source

เมื่อเขียนคำสั่งแล้วทำการดำเนินงานคำสั่งจะได้ผลลัพธ์ดังในรูปที่ 6.8



รูปที่ 6.8 ผลลัพธ์จากการดำเนินงานคำสั่งใน Source

6.2.2 การนำเข้าข้อมูล csv

การนำเข้าข้อมูลที่อยู่ในรูปแบบ[5,6] csv (Comma delimited)โดยส่วนใหญ่จะรู้จักกันในรูปแบบ Excel อย่างไรก็ตามสามารถทำให้อยู่ในรูปแบบ csv ได้โดยการ “Save as” และเลือกรูปแบบ “Save as type” เป็นแบบ csv (Comma delimited) ตัวอย่างในตารางที่ 6.1 สามารถจัดการข้อมูลได้ดังรูปที่ 6.9

	A	B
1	alc	noalc
2	25	20
3	24	20
4	24	19
5	27	21
6	26	23
7	26	22
8	24	19
9	25	20
10	25	23
11	27	21
12	26	22
13	25	19
14	25	19
15	25	22
16	26	22

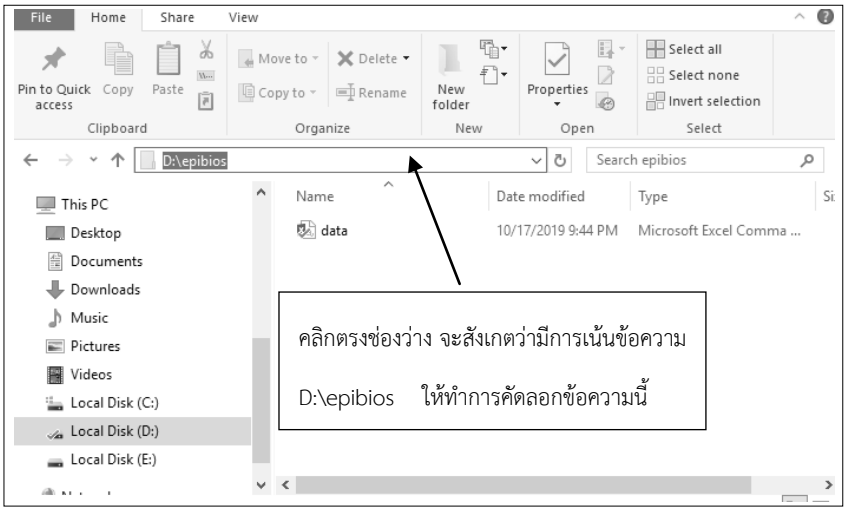
รูปที่ 6.9 ชุดข้อมูลตัวอย่างที่อยู่ในรูปแบบ csv

เมื่อได้ข้อมูลที่อยู่ในรูปแบบดังรูปที่ 6.9 ให้ทำการเขียนคำสั่งสำหรับนำเข้าข้อมูลคำสั่งแรกเป็นคำสั่งในการกำหนดตำแหน่งของโฟลเดอร์สำหรับการทำงาน *setwd()* และคำสั่งที่สองสำหรับการนำเข้าข้อมูล[7]

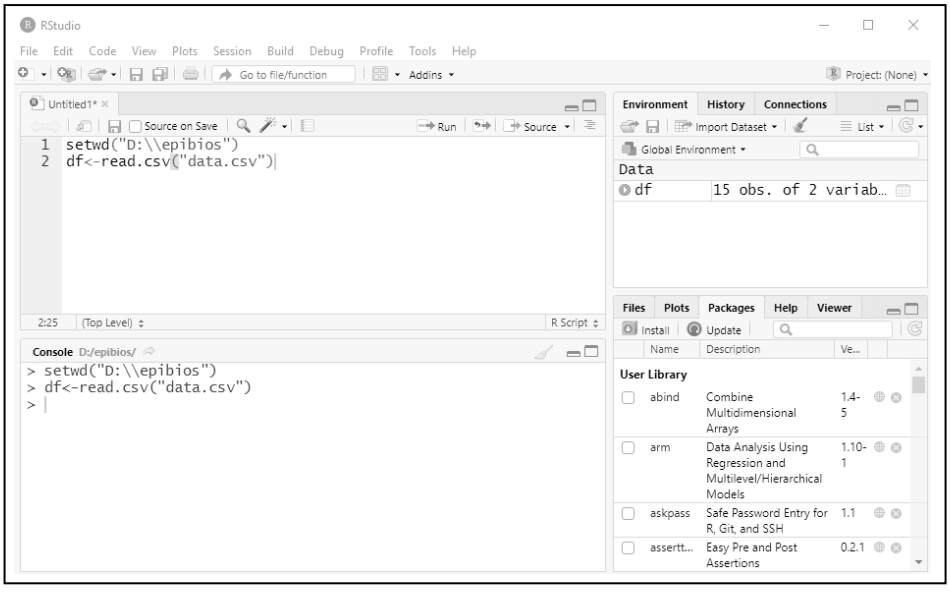
คำสั่งกำหนดเส้นทางโฟลเดอร์สำหรับการทำงาน (Set working directory)

รูปแบบคำสั่ง setwd(“path”)

โดยที่ path หมายถึง เส้นทางของโฟลเดอร์ทำงานสามารถค้นหาได้โดยการเปิดโฟลเดอร์ที่ทำงานจากนั้นให้ทำการคลิกตรงช่องว่างดังรูปที่ 6.10 ในตัวอย่างนี้โฟลเดอร์ที่ใช้ทำงานชื่อ epibios



รูปที่ 6.10 การเปิดดู path ของตำแหน่งโฟลเดอร์ทำงาน (Working directory)



รูปที่ 6.11 การนำเข้าข้อมูลในรูปแบบ csv

รูปแบบคำสั่ง

setwd("D:\\epibios")

df<-read.csv("data.csv")

ในรูปที่ 6.11 หลังจากที่ดำเนินงาน (run) คำสั่งใน Source แล้วพบว่าที่หน้าต่าง Environment จะแสดง Data มาให้ชื่อว่า df ซึ่งได้จากดำเนินงานคำสั่ง *df<-read.csv("data.csv")* นอกจากนั้นแล้วจะแสดงผลลัพธ์การดำเนินงานคำสั่งในหน้าต่าง Console ด้วย (ในหน้าต่าง Console จะมีเครื่องหมาย > นำหน้าทุกบรรทัด เป็นเครื่องหมายที่แสดงผลลัพธ์ในบรรทัดนั้นๆ)

จากการนำเข้าข้อมูลทั้ง 2 วิธี ข้อมูลที่นำเข้าสู่โปรแกรม R เพื่อวิเคราะห์ข้อมูลเรียกว่าดาต้าเฟรม (dataFrame) ซึ่งในการอ้างอิงดาต้าเฟรมทำได้ 2 วิธีคือใช้คำสั่ง *attach (ชื่อดาต้าเฟรม)* หรือ *ชื่อดาต้าเฟรม**$**ชื่อตัวแปร* ในตำราเล่มนี้จะใช้วิธีอ้างอิงชื่อดาต้าเฟรมคือ *ชื่อดาต้าเฟรม**$**ชื่อตัวแปร* เช่น *df$alc* เป็นต้น

6.3 การจัดกลุ่มข้อมูล[8-11]

การในเก็บข้อมูลงานวิจัยจะมีตัวแปรต่อเนื่องซึ่งในบางครั้งต้องการจัดกลุ่มข้อมูลจากตัวแปรต่อเนื่อง (Continuous data) เป็นตัวแปรจำแนกประเภท (Categorical data) ในตัวอย่างนี้ทำการจัดกลุ่มอายุ ให้เป็นตัวแปรจำแนกประเภท โดยกลุ่มที่ 1 อายุน้อยกว่าหรือเท่ากับ 50 ปี และ กลุ่มที่ 2 อายุมากกว่า 50 ปี โดยใช้คำสั่งพื้นฐานของโปรแกรม R คือ *cut()* ในบางครั้งไม่สามารถจำรูปแบบคำสั่งการใช้งานของชุดข้อมูลนั้นได้ จำเป็นต้องดูรูปแบบการใช้งานของชุกคำสั่งนั้นๆ ให้พิมพ์ ? หน้าคำสั่ง เช่น *?cut* จากนั้นสั่งดำเนินงานคำสั่งได้ผลลัพธ์ รูปแบบการใช้งานของชุดคำสั่งนั้นในหน้าต่าง Help (หน้าต่างหมายเลข 4)

```
> age<-c(53,52,53,51,54,59,47,57,48,46,58,49,49,46)
> gender<-c(0,1,1,1,1,0,1,1,0,0,1,0,1,0)
>df<-data.frame(age,gender)
> ?cut
>df$age_gr<-cut(df$age,breaks=c(min(df$age,na.rm =
TRUE),50,max(df$age,na.rm=TRUE)),include.lowest = TRUE)
> table(df$age_gr)

[46,50] (50,59]
      6       8
```

รูปที่ 6.12 ผลการจัดกลุ่มข้อมูลโดยใช้คำสั่ง cut()

ในตัวอย่างนี้ใช้คำสั่ง

```
df$age_gr<-cut(df$age,
        breaks=c(min(df$age,na.rm =TRUE),50,max(df$age,na.rm=TRUE)),
        include.lowest = TRUE)
```

เป็นการจัดกลุ่มอายุใหม่โดยสร้างตัวแปรใหม่คือ *age_gr* จากตัวแปรเดิมชื่อ *df$age* มีการจัดกลุ่มใหม่ตั้งแต่ค่าต่ำสุดไปจนถึง 50 และ มากกว่าหรือเท่ากับ 50 จนถึงค่าสูงสุดสามารถใช้คำสั่ง *breaks=c(min(df$age,na.rm =TRUE), 50, max(df$age,na.rm=TRUE))* เมื่อทำการดำเนินงานคำสั่ง ให้ใช้คำสั่ง *table(ตัวแปร)* หรือ *table(df$age_gr)* เพื่อตรวจสอบดูผลลัพธ์ของการแบ่งกลุ่มอายุใหม่ว่าถูกต้องตรงตามที่กำหนดไว้หรือไม่ นอกจากนั้นแล้วยังสามารถทำการระบุค่าในข้อมูลนั้นได้ โดยเพิ่มคำสั่งดังนี้

```
df$age_gr<-cut(df$age,
        breaks=c(min(df$age,na.rm =TRUE),50,max(df$age,na.rm=TRUE)),
        labels=c("<=50",">50"),
        include.lowest = TRUE)
```

	A	B
1	age	gender
2	49	0
3	54	0
4	60	1
5	59	1
6	48	0
7	51	1
8	51	0
9	58	1
10	57	0
11	58	0
12	48	0
13	57	0
14	54	1
15	53	0

รูปที่ 6.13 ตัวอย่างการจัดการข้อมูลตัวแปรเพศ

ตัวแปรเพศ (ชื่อตัวแปร *gender*) ประกอบไปด้วยหมายเลข 1 หมายถึง เพศหญิง (Female) และ หมายเลข 2 หมายถึงเพศชาย (Male) ดังนั้นเราสามารถจัดการข้อมูลได้ดังนี้

```
>setwd("D:\\epibios")
>df<-read.csv("factornumeric.csv")
> summary(df)
      age            gender
 Min.   :40.00   Min.   :0.0000
 1st Qu.:41.25   1st Qu.:0.0000
 Median :46.00   Median :0.0000
 Mean   :47.36   Mean   :0.2857
 3rd Qu.:52.75   3rd Qu.:0.7500
 Max.   :58.00   Max.   :1.0000
> table(df$gender)

0  1
10  4
>df$gender<- factor(df$gender,
+                    levels = c(0,1),
+                    labels = c("female","male"))
> table(df$gender)

female   male
    10      4
```

รูปที่ 6.14 ผลการนำเข้าและจัดการข้อมูลตัวแปรเพศ

6.3 บทสรุป

ในบทนี้ได้เรียนรู้วิธีการใช้งานเบื้องต้นของโปรแกรม R และชุดคำสั่งต่างๆ ในบทต่อไปเป็นการวิเคราะห์ข้อมูลซึ่งต้องใช้พื้นฐานของการใช้โปรแกรม R ในการวิเคราะห์ข้อมูล ทั้งนี้เราสามารถอ่านเพิ่มเติมวิธีการใช้โปรแกรม R ได้จากเอกสารอ้างอิงท้ายบท

6.4 แบบฝึกหัด

1. โปรแกรม R และ โปรแกรม RStudio แตกต่างกันอย่างไร
2. หน้าต่างที่ใช้สำหรับเขียนคำสั่งในการวิเคราะห์ข้อมูลเรียกว่าอะไร
3. คำสั่งในการกำหนดเส้นทางโฟลเดอร์สำหรับการทำงานคืออะไร
4. ชุดคำสั่ง epiDisplay พัฒนาเพื่อใช้ในงานวิจัยทางด้านใด
5. การจัดกลุ่มข้อมูลจากข้อมูลต่อเนื่องเป็นข้อมูลกลุ่มใช้คำสั่งอะไร

6.5 เอกสารอ้างอิง

1. R Development Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2020.
2. Chongsuvivatwong V. epiDisplay: Epidemiological Data Display Package. 2018.
3. Wickham H. ggplot2: Elegant Graphics for Data Analysis: Springer-Verlag New York; 2016.
4. RStudio Team. RStudio: Integrated Development Environment for R. Boston, MA: RStudio, PBC; 2020.
5. Elff M. Data Management in R: A Guide for Social Scientists: SAGE Publications; 2020.
6. Sünsüli M. Introduction to Spatial Data Management with R: “Raster, Vector, Graphics and Visualization”: Saygın Kantara; 2019.
7. Venables WN, Smith DM, Team RDC. An introduction to R. Citeseer; 2009.
8. Chen DG, Peace KE. Clinical Trial Data Analysis Using R: CRC Press; 2010.
9. Chernick MR, Friis RH. Introductory Biostatistics for the Health Sciences: Modern Applications Including Bootstrap: Wiley; 2003.
10. Shintani A. Basic Concepts of Biostatistics With R Commander: Taylor & Francis; 2018.
11. Shahbaba B. Biostatistics with R: An Introduction to Statistics Through Biological Data: Springer New York; 2011.
12. Spector P. Data Manipulation with R: Springer; 2008.
13. Hothorn T, Everitt BS. A Handbook of Statistical Analyses using R: CRC Press; 2014.

บทที่ 7
การวิเคราะห์ข้อมูลต่อเนื่อง

จากบทที่ 5-6 ได้ศึกษามาตราวัดของตัวแปรการตั้งสมมุติฐานงานวิจัย และการใช้โปรแกรม R สำหรับการวิเคราะห์ข้อมูลทางสถิติ ในบทนี้จะได้ศึกษาการวิเคราะห์ข้อมูลที่ตัวแปรผลลัพธ์ทางการศึกษา (Outcome) มีลักษณะเป็นตัวแปรต่อเนื่อง (Continuous data) โดยใช้สถิติ Unpaired t-test, Mann-Whitney's U test, Paired t-test, Wilcoxon signed rank test, One-way ANOVA และ Kruskal Wallis test ได้

7.1 การเปรียบเทียบค่าเฉลี่ยกลุ่มเดียว

การประมาณค่าเฉลี่ยของประชากรกลุ่มเดียวเป็นการประมาณค่าข้อมูลที่มีการวัดแบบต่อเนื่อง[1] (Continuous data) เช่น อายุ รายได้ ระดับไขมันชนิด Low-Density Lipoprotein (LDL) อัตราการกรองของไต (Glomerular filtration rate : GFR) เป็นต้น

ตัวอย่างที่ 7.1 การเปรียบเทียบค่าเฉลี่ยของประชากรกลุ่มเดียว

การศึกษาน้ำหนักแรกเกิดของเด็กในจังหวัดมหาสารคาม ทำการสุ่มตัวอย่างและเก็บข้อมูลทั้งหมด 500 คน ซึ่งมีค่าเฉลี่ยเท่ากับ 2900 กรัม ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 400 ต้องการเปรียบเทียบค่าเฉลี่ยที่ได้กับค่าเฉลี่ยของประเทศมีค่าเท่ากับ 2600 กรัมมีค่าแตกต่างกันหรือไม่

วิธีทำ กำหนดสมมุติทางสถิติ (เป็นการทดสอบแบบสองทาง)

H_0 : ค่าเฉลี่ยน้ำหนักแรกเกิดของเด็กในจังหวัดมหาสารคามมีค่าเท่ากับ 2600 กรัม

H_a : ค่าเฉลี่ยน้ำหนักแรกเกิดของเด็กในจังหวัดมหาสารคามมีค่าไม่เท่ากับ 2600 กรัม

กำหนดความผิดพลาดชนิดที่ 1 (Type I error) นิยมกำหนดที่ 0.05 (α = 0.05)

ขั้นตอนทดสอบทางสถิติ

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

$$t = \frac{2900 - 2600}{\frac{400}{\sqrt{500}}} = 16.77051$$

จากค่า t ดังกล่าวสามารถนำไปเปิดตารางทางสถิติเพื่อคำนวณค่า t และค่าความน่าจะเป็นได้อย่างไรก็ตามสามารถใช้โปรแกรม R ช่วยในการคำนวณได้ดังรูปที่ 7.1

```
> mean=2900
>   sd=400
>   n =500
> test=2600
> t.value = (mean - test) / (sd/ sqrt(n))
> p.value = 2*pt(-abs(t.value), df=n-1)
> p.value
[1] 2.170086e-50
```

ค่าความน่าจะเป็น (p-value)

รูปที่ 7.1 ผลการวิเคราะห์ข้อมูลค่าเฉลี่ยกลุ่มเดียว

ผลการวิเคราะห์ข้อมูลดังกล่าวได้ค่าความน่าจะเป็นเท่ากับ 2.170086e-50 หรือ p-value < 0.001 โดยที่ถ้าค่าความน่าจะเป็นที่คำนวณได้ (Probability value: p-value) น้อยกว่า 0.05 จะปฏิเสธ สมมุติฐานว่าง สรุปได้ว่ามีหลักฐานเพียงพอที่จะปฏิเสธสมมุติฐานว่างหรือสมมุติฐาน H_0 แปลผลได้คือ จากผลการวิเคราะห์พบว่าค่าเฉลี่ยน้ำหนักแรกเกิดของเด็กในจังหวัดมหาสารคามมีค่าแตกต่างกับเกณฑ์ของประชากรเด็กในประเทศไทย 2600 กรัมอย่างมีนัยสำคัญทางสถิติ (p-value < 0.001)

7.2 การวิเคราะห์ข้อมูลเพื่อเปรียบเทียบค่าเฉลี่ยสองกลุ่มที่อิสระต่อกัน

การวิเคราะห์ข้อมูลเพื่อเปรียบเทียบค่าเฉลี่ยของประชากรสองกลุ่มที่อิสระต่อกัน (Unpaired t-test หรือ Independent t-test)[2] มีข้อตกลงเบื้องต้นของการใช้สถิติ ดังนี้

1. ประชากรมี 2 กลุ่ม
2. ข้อมูลมีความเป็นอิสระต่อกัน
3. มีการแจกแจงของข้อมูลแบบปกติทั้ง 2 กลุ่ม

ตัวอย่างที่ 7.2 การศึกษาเพื่อเปรียบเทียบค่าเฉลี่ยของดัชนีมวลกายระหว่างกลุ่มที่ดื่มและไม่ดื่ม เครื่องดื่มแอลกอฮอล์ได้ข้อมูลในดังแสดงในตารางที่ 7.1

ตารางที่ 7.1 ค่าดัชนีมวลกายระหว่างกลุ่มที่ดื่มเครื่องดื่มแอลกอฮอล์

กลุ่มดื่ม	25	24	24	27	26	26	24	25	25	27	26	25	25	25	26
กลุ่มไม่ดื่ม	20	20	19	21	23	22	19	20	23	21	22	19	19	22	22

วิธีทำ

ขั้นตอนที่ 1 ทดสอบการแจกแจงของข้อมูลโดยใช้สถิติ Sharpiro's wilk test[3-5]

```
>alc <-c(25,24,24,27,26,26,24,25,25,27,26,25,25,25,26)
>noalc<-c(20,20,19,21,23,22,19,20,23,21,22,19,19,22,22)
> data <- data.frame(alc, noalc)
> shapiro.test(data$alc)

      Shapiro-Wilk normality test

data:  data$alc
W = 0.8912, p-value = 0.06992

> shapiro.test(data$noalc)

      Shapiro-Wilk normality test

data:  data$noalc
W = 0.88429, p-value = 0.05501
```

ค่าความน่าจะเป็น (p-value)

ค่าความน่าจะเป็น (p-value)

รูปที่ 7.2 ผลการวิเคราะห์การแจกแจงของข้อมูล

จากการทดสอบทางสถิติพบว่าข้อมูลมีการแจกแจงปกติทั้งสองกลุ่ม (ความน่าจะเป็น Shapiro-Wilk test ของค่าเฉลี่ยน้ำหนักแรกเกิดในกลุ่มดื่มเครื่องดื่มแอลกอฮอล์เท่ากับ 0.699 และกลุ่มไม่ดื่มเครื่องดื่มแอลกอฮอล์เท่ากับ 0.055) ดังนั้น ใช้สถิติ Unpaired t-test สำหรับเปรียบเทียบค่าเฉลี่ยสองกลุ่มที่อิสระต่อกัน

ขั้นตอนที่ 2 ทดสอบความแปรปรวน

```
> var.test(alc,noalc, data = data)

        F test to compare two variances

data:  alc and noalc
F = 0.4386, num df = 14, denom df = 14, p-value = 0.1351
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.1472498 1.3063980
sample estimates:
ratio of variances
         0.4385965
```

รูปที่ 7.3 การทดสอบความแปรปรวนของค่าเฉลี่ยสองกลุ่มอิสระต่อกัน

การทดสอบความแปรปรวนพบว่าข้อมูลทั้งสองกลุ่มมีความแปรปรวนเท่ากัน (p-value = 0.135) ดังนั้นในการเปรียบเทียบค่าเฉลี่ยระหว่างสองกลุ่มใช้สถิติ Independent t-test หรือ Unpaired t-test แบบความแปรปรวนเท่ากัน[6-8]

ขั้นตอนที่ 3 ทดสอบสถิติ independent t-test หรือ Unpaired t-test

```
> t.test(data$alc,data$noalc)

        Welch Two Sample t-test

data:  data$alc and data$noalc
t = 9.9339, df = 24.299, p-value = 4.921e-10
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
3.592090 5.474577
sample estimates:
mean of x mean of y
 25.33333  20.80000
```

(3.59, 5.47) ช่วงเชื่อมั่นที่ 95% ความแตกต่างของค่าเฉลี่ยระหว่างสองกลุ่ม

รูปที่ 7.4 ผลการวิเคราะห์ข้อมูล independent t-test แบบความแปรปรวนเท่ากัน

ถ้าความแปรปรวนไม่เท่ากัน สามารถใช้คำสั่งในโปรแกรมได้ดังนี้

t.test(data$alc,data$noalc, var.equal = FALSE)

จากการทดสอบทางสถิติพบว่าได้ค่าความน่าจะเป็นเท่ากับ 4.921e-10 สามารถเขียนได้เป็น < 0.001 สรุปผลการวิเคราะห์ได้ดังนี้ ค่าเฉลี่ยของดัชนีมวลกายของประชากรกลุ่มที่ดื่มเครื่องดื่มแอลกอฮอล์มีค่าแตกต่างกับกลุ่มที่ไม่ดื่มเครื่องดื่มแอลกอฮอล์อย่างมีนัยสำคัญทางสถิติ (p-value < 0.001)

7.3 การวิเคราะห์ข้อมูลโดยใช้ Mann-Whitney's U test

จากการทดสอบแบบ Unpaired t-test มีการทดสอบการแจกแจงของข้อมูลก่อน ถ้าข้อมูลมีการแจกแจงไม่ปกติ การใช้สถิติ Unpaired t-test อาจไม่เหมาะสม ดังนั้นจึงใช้วิธีการวิเคราะห์แบบ Mann-Whitney's U test[9]

ตัวอย่างที่ 7.3 การศึกษาเพื่อเปรียบเทียบสัดส่วนเส้นรอบเอวต่อรอบสะโพก (Waist to Hip Ratio: WHR) ระหว่างเพศชายและเพศหญิง ได้ข้อมูลดังนี้

ตารางที่ 7.2 ข้อมูลสัดส่วนรอบเอวต่อรอบสะโพกกลุ่มตัวอย่างของเพศชายและหญิง

เพศชาย	0.99	0.98	0.98	0.91	0.55	0.96	0.75	0.89	0.86	0.89
เพศหญิง	0.59	0.75	0.76	0.83	0.86	0.86	0.89	0.89	0.96	0.98

จากข้อมูลดังกล่าวทำการทดสอบทางสถิติได้ดังนี้

วิธีทำ

ขั้นตอนที่ 1 ทดสอบการแจกแจงข้อมูล (ใช้คำสั่ง shapiro.test)

```
> whrmale<-c(0.99,0.98,0.98,0.91,0.55,0.96,0.75,0.89,0.86,0.89)
> whrfemale<-c(0.59,0.75,0.76,0.83,0.86,0.86,0.89,0.89,0.96,0.98)
> data <- data.frame(whrmale, whrfemale)
> shapiro.test(data$whrmale)

        Shapiro-Wilk normality test

data:  data$whrmale
W = 0.79646, p-value = 0.01313

> shapiro.test(data$whrfemale)

        Shapiro-Wilk normality test

data:  data$whrfemale
W = 0.92138, p-value = 0.3686
```

รูปที่ 7.5 ผลการทดสอบการแจงแจงข้อมูลของสัดส่วนเส้นรอบเอวต่อรอบสะโพก

ผลการวิเคราะห์พบว่าค่าสัดส่วนเส้นรอบเอวต่อรอบสะโพกของเพศชายมีการแจกแจงไม่ปกติ (p-value = 0.013) และเพศหญิงมีการแจกแจงปกติ (p-value = 0.369) ดังนั้นในการทดสอบทางสถิตินี้จะใช้ Mann-Whitney's U test

ขั้นตอนที่2 ทดสอบสถิติ Mann-Whitney's U test (ใช้คำสั่ง wilcox.test)

```
> wilcox.test(data$whrmale,data$whrfemale)

        Wilcoxon rank sum test with continuity correction

data:  data$whrmale and data$whrfemale
W = 66, p-value = 0.2378
alternative hypothesis: true location shift is not equal to 0
```

รูปที่ 7.6 ผลการทดสอบความแตกต่างของสัดส่วนเส้นรอบเอวต่อรอบสะโพก

ผลการทดสอบทางสถิติพบว่า ค่ากลางของสัดส่วนเส้นรอบเอวต่อรอบสะโพกระหว่างเพศชายและเพศหญิงมีค่าไม่แตกต่างกัน (p-value = 0.238)

7.4 การเปรียบเทียบค่าเฉลี่ยสองกลุ่มที่ไม่อิสระต่อกัน

การทดสอบทางสถิติสำหรับการเปรียบเทียบค่าเฉลี่ยสองกลุ่มที่ไม่อิสระต่อกันใช้สถิติ Paired t-test[10] เช่นการศึกษาที่มีรูปแบบ ก่อน-หลัง การทดลอง (Pre-post study) โดยมีข้อตกลงเบื้องต้นดังนี้

1. ข้อมูลมีความไม่อิสระต่อกัน
2. ข้อมูลต่อเนื่องและมีการแจกแจงแบบปกติ

ตัวอย่างที่ 7.4 การทดสอบโปรแกรมการออกกำลังกายเพื่อลดน้ำหนัก (ดัชนีมวลกาย) ในนักเรียนชั้นมัธยมศึกษาตอนปลาย โดยการเข้าร่วมออกกำลังกายเป็นเวลา 45 วันๆ ละ 30 นาที เก็บข้อมูลก่อน-หลังการเข้าร่วมโปรแกรม ได้ข้อมูลดังนี้

ตารางที่ 7.3 ข้อมูลก่อน-หลังการเข้าร่วมโปรแกรมออกกำลังกาย

ลำดับที่	ก่อนทดลอง	หลังทดลอง
1	30	24
2	28	23
3	33	24
4	28	25
5	30	24
6	29	23
7	30	24
8	31	23
9	32	23
10	29	25

ขั้นตอนการวิเคราะห์ข้อมูล

ขั้นตอนที่ 1 ทดสอบการแจกแจงของข้อมูล

```
> before<-c(30,28,33,28,30,29,30,31,32,29)
> after<-c(24,23,24,25,24,23,24,23,23,25)
> data <- data.frame(before, after)
> shapiro.test(data$before-data$after)

        Shapiro-Wilk normality test

data:  data$before - data$after
W = 0.92233, p-value = 0.3768
```

รูปที่ 7.7 การสร้างข้อมูลและทดสอบการแจกแจง

การทดสอบการแจกแจงของข้อมูลในกรณีที่ข้อมูลไม่อิสระต่อกัน ทำการทดสอบจากค่าก่อนการทดลอง – ค่าหลังการทดลอง ในโปรแกรม R สามารถนำข้อมูลลบกันได้โดยตรง (shapiro.test(data$before-data$after)) การทดสอบนี้พบว่าข้อมูลมีการแจกแจงแบบปกติ (p-value = 0.377) จึงใช้สถิติ Paired t-test ได้

ขั้นตอนที่ 2 ทดสอบสถิติโดยใช้ Paired t-test

```
> t.test(data$before,data$after,paired = TRUE)

        Paired t-test

data:  data$before and data$after
t = 9.858, df = 9, p-value = 4.03e-06
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 4.777257 7.622743
sample estimates:
mean of the differences
                    6.2
```

รูปที่ 7.8 ผลการวิเคราะห์ข้อมูล Paired t-test

ผลการทดสอบทางสถิติโดยใช้ Paired t-test พบว่า ค่าเฉลี่ยของดัชนีมวลกายก่อนและหลังการทดลองมีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติ (p-value=4.03e-06 หรือ p-value <0.001)

7.5 การเปรียบเทียบค่าเฉลี่ยสองกลุ่มที่ไม่อิสระต่อกันกรณีข้อมูลมีการแจกแจงไม่ปกติ

ในกรณีที่ข้อมูลมีการแจกแจงไม่ปกติ ใช้สถิติ Wilcoxon signed rank test[11,12] เพื่อทำการเปรียบเทียบค่ากลางของข้อมูลก่อน-หลังทดลอง

ตัวอย่างที่ 7.5 การทดสอบผลของการรับประทานอาหารที่มีกรดไขมันไลโนเลอิก (Linoleic acid) 25 กรัมต่อวันเป็นเวลา 30 วันต่อเนื่องกัน มีการวัดระดับโคเลสเตอรอลในเลือดก่อน-หลังการทดลองได้ผลดังนี้

ตารางที่ 7.4 ระดับโคเลสเตอรอลในเลือดก่อน-หลังการทดลอง

ลำดับที่	ก่อนทดลอง	หลังทดลอง
1	241	215
2	243	201
3	248	220
4	248	200
5	253	226
6	257	213
7	259	221
8	260	225
9	263	283
10	272	282

ขั้นตอนที่ 1 ทดสอบการแจกแจงของข้อมูล

```
> before<-c(241,243,248,248,253,257,259,260,272,263)
> after<-c(215,201,220,200,226,213,221,225,282,283)
> data <- data.frame(before, after)
> shapiro.test(data$before-data$after)

        Shapiro-Wilk normality test

data:  data$before - data$after
W = 0.81263, p-value = 0.02064
```

รูปที่ 7.9 การสร้างชุดข้อมูลและทดสอบการแจกแจง

จากผลการทดสอบการแจกแจงพบว่าข้อมูลมีการแจกแจงไม่ปกติ (p-value = 0.021) ดังนั้นจะใช้สถิตินอนพารามิติคือ Wilcoxon signed rank test

ขั้นตอนที่ 2 ทดสอบสถิติ Wilcoxon signed rank test

```
> wilcox.test(data$before, data$after, paired=TRUE)

        Wilcoxon signed rank test

data:  data$before and data$after
V = 52, p-value = 0.009766
alternative hypothesis: true location shift is not equal to 0
```

รูปที่ 7.10 ผลการวิเคราะห์ข้อมูลโดยใช้ Wilcoxon signed rank test

ผลการทดสอบทางสถิติพบว่า ค่ากลางของระดับโคเลสเตอรอลในเลือดก่อน-หลังได้รับอาหารที่มีกรดไลโนเลอิกจำนวน 25กรัมต่อวันเป็นเวลา 30 วันต่อเนื่องกันมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ (p-value = 0.009766 หรือ p-value < 0.001)

7.6 การเปรียบเทียบค่าเฉลี่ยมากกว่าสองกลุ่มที่อิสระต่อกัน

การเปรียบเทียบค่าเฉลี่ยมากกว่าสองกลุ่มใช้สถิติ One-way ANOVA มีข้อตกลงเบื้องต้นดังนี้[13,14]

1. ตัวอย่างแต่ละกลุ่มต้องสุ่มมาอย่างอิสระ
2. ประชากรแต่ละกลุ่มมีการแจกแจงปกติ

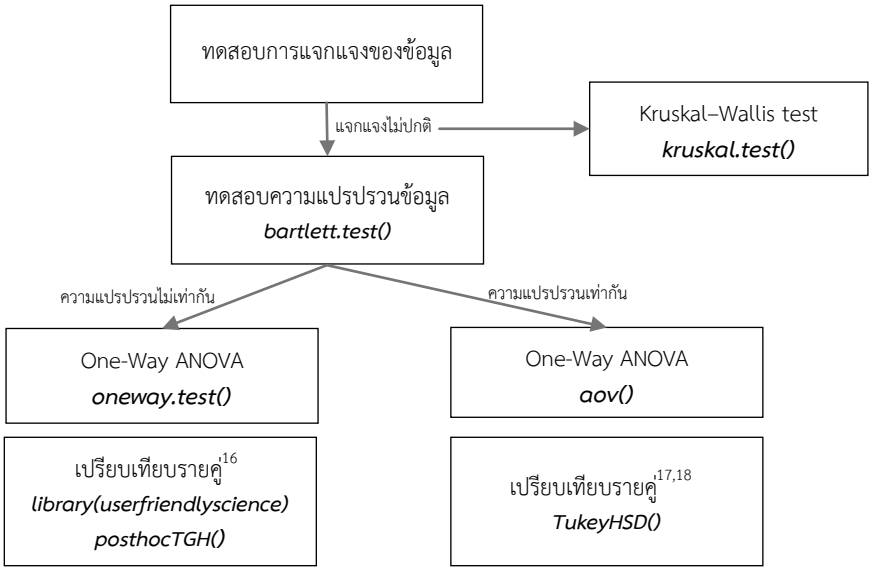
โดยการทดสอบสมมุติฐานทางสถิติของ One-way ANOVA สามารถกำหนดได้ดังนี้

H_0: ค่าเฉลี่ยของแต่ละกลุ่มไม่มีค่าแตกต่างกัน

H_a: ค่าเฉลี่ยของแต่ละกลุ่มมีค่าแตกต่างกัน อย่างน้อย 1 คู่

ถ้าการทดสอบสมมุติฐานของ One-way ANOVA พบว่าค่าเฉลี่ยของแต่ละกลุ่มมีค่าแตกต่างกันอย่างน้อยจำนวน 1 คู่ (p-value < 0.05)ให้ทำการทดสอบแบบรายคู่เพิ่มเติม เพื่อที่ทำการทดสอบมีคู่ไหนบ้างที่มีค่าเฉลี่ยแตกต่างกัน การเลือกใช้สถิติและคำสั่งในโปรแกรม R ดังรูปที่ 7.11 ขั้นตอนแรกทำการทดสอบการแจกแจงของข้อมูลในแต่ละกลุ่ม ถ้าพบว่ามีการแจกแจงไม่ปกติแค่กลุ่มเดียว จะใช้สถิติ Kruskal–Wallis test[15] (คำสั่งใน R: *kruskal.test()*) ในกรณีที่มีการแจกแจงปกติทุกกลุ่ม ให้ทำการทดสอบความเท่ากันของความแปรปรวนในแต่ละกลุ่ม ถ้าเท่ากันใช้การวิเคราะห์ One-way ANOVA แบบความแปรปรวนเท่ากัน (คำสั่งใน R: *aov()*) ถ้าความแปรปรวนไม่เท่ากันจะใช้การวิเคราะห์ One-way ANOVA แบบความแปรปรวนไม่เท่ากัน

ถ้าผลลัพธ์ของ One-way ANOVA มีค่า p-value < 0.05 ทำการทดสอบรายคู่ต่อไป (คำสั่งใน R: *posthocTGH()*)



รูปที่ 7.11 การทดสอบเพื่อเปรียบเทียบค่าเฉลี่ยมากกว่า 2 กลุ่ม

ตัวอย่างที่ 7.6 การเปรียบเทียบเส้นรอบเอว (Waist circumference เซนติเมตร) ระหว่างประชากร 3 กลุ่มอาชีพ คือ กลุ่มพนักงานออฟฟิศ (กลุ่มที่ 1) กลุ่มพนักงานร้านอาหาร (กลุ่มที่ 2) และกลุ่มเจ้าหน้าที่บันทึกข้อมูล (กลุ่มที่ 3) โดยมีข้อมูลดังนี้

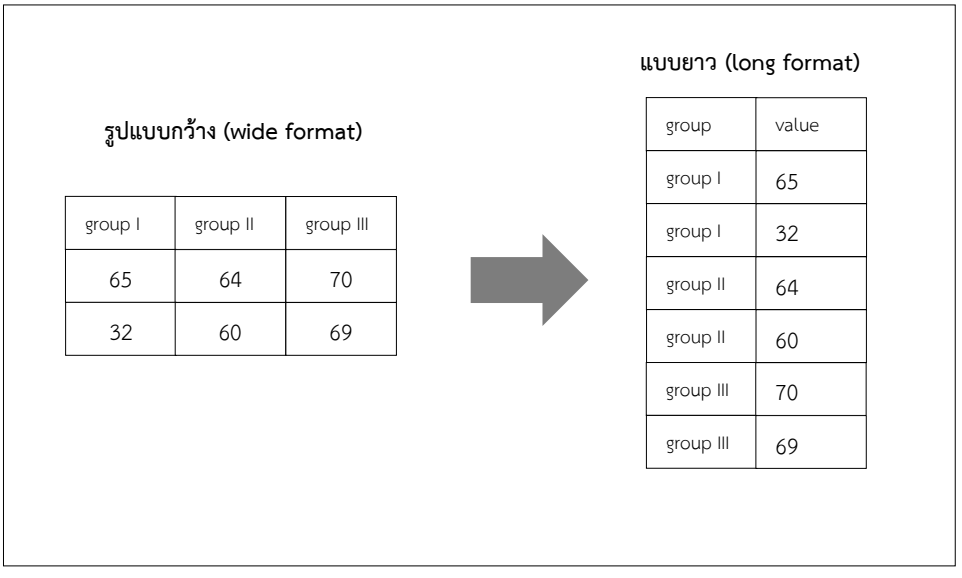
ตารางที่ 7.5 เส้นรอบเอวระหว่าง 3 กลุ่มตัวอย่าง

กลุ่มที่ 1	กลุ่มที่ 2	กลุ่มที่ 3
146	124	119
132	126	131
186	118	122
168	127	122
129	111	126
143	127	108
112	114	104
124	86	124
150	93	125
124	109	111

วิธีทำ

ขั้นตอนที่ 1 ทดสอบการแจกแจงของข้อมูล

ขั้นตอนการวิเคราะห์ข้อมูล ทำการสร้างข้อมูลให้อยู่ในรูป data.frame จากนั้นทำการเปลี่ยนแปลงรูปร่างของข้อมูล (Reshape) จากรูปแบบกว้าง (Wide format) ให้อยู่เป็นแบบยาว (Long format) ดังรูปที่ 7.12



รูปที่ 7.12 รูปแบบของข้อมูลแบบกว้าง (Wide format) และแบบยาว (Long format)

```
#One Way ANOVA
gr1<-c(146,132,186,168,129,143,112,124,150,124)
gr2<-c(124,126,118,127,111,127,114,86,93,109)
gr3<-c(119,131,122,122,126,108,104,124,125,111)
data <- data.frame(gr1, gr2,gr3)

#transform wide to long
anovadta<- reshape(data,
                   varying = c("gr1", "gr2", "gr3"),
                   v.names = "score",
                   timevar = "group",
                   times = c("group I", "group II", "group III"),
                   direction = "long")
```

รูปที่ 7.13 การเปลี่ยนรูปร่าง (Reshape) ของข้อมูล

จากคำสั่งดังกล่าวเป็นการสร้าง data.frame ชื่อ data เป็นรูปแบบกว้าง (wide format) จากนั้นใช้คำสั่งในโปรแกรม R ชื่อ *reshape* สร้าง data.frame ให้ชื่อ aovadta ซึ่งเป็นชุดข้อมูลที่ใช้สำหรับวิเคราะห์ One-way ANOVA

```
> #test normality
> shapiro.test(anovadta$score[anovadta$group=="group I"])

        Shapiro-Wilk normality test

data:  anovadta$score[anovadta$group == "group I"]
W = 0.93845, p-value = 0.5359

> shapiro.test(anovadta$score[anovadta$group=="group II"])

        Shapiro-Wilk normality test

data:  anovadta$score[anovadta$group == "group II"]
W = 0.86743, p-value = 0.09328

> shapiro.test(anovadta$score[anovadta$group=="group III"])

        Shapiro-Wilk normality test

data:  anovadta$score[anovadta$group == "group III"]
W = 0.91891, p-value = 0.348
```

รูปที่ 7.14 ผลการทดสอบการแจกแจงของข้อมูล

ผลการทดสอบพบว่า ข้อมูลมีการแจกแจงแบบปกติทั้ง 3 กลุ่ม (p-value > 0.05)

ขั้นตอนที่ 2 ทดสอบความแปรปรวน

```
> bartlett.test(score~group, data=anovadta)

        Bartlett test of homogeneity of variances

data:  score by group
Bartlett's K-squared = 7.0349, df = 2, p-value = 0.02968
```

รูปที่ 7.15 ผลการวิเคราะห์ความแปรปรวน

จากผลการวิเคราะห์ข้อมูลพบว่า ข้อมูลมีความแปรปรวนไม่เท่ากัน (p-value = 0.03) ลำดับต่อไปทำการทดสอบสถิติ One-way ANOVA แบบความแปรปรวนไม่เท่ากัน

ขั้นตอนที่ 3 การทดสอบ One-way ANOVA แบบความแปรปรวนไม่เท่ากัน

```
> anvun<-oneway.test(anovadta$score~anovadta$group)
> anvun

        One-way analysis of means (not assuming equal variances)

data:  anovadta$score and anovadta$group
F = 5.4146, num df = 2.000, denom df = 15.979, p-value = 0.01602
```

รูปที่ 7.16 ผลการวิเคราะห์ข้อมูล One-way ANOVA

ผลการวิเคราะห์พบว่า ค่าเฉลี่ยของเส้นรอบเอวระหว่าง 3 อาชีพมีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติ (p-value = 0.016) ทั้งนี้มีอย่างน้อย 1 คู่ ที่มีความแตกต่างกันจึงทำการทดสอบรายคู่ต่อ

ขั้นตอนที่ 4 การทดสอบรายคู่ (Multiple comparison) แบบความแปรปรวนไม่เท่ากัน

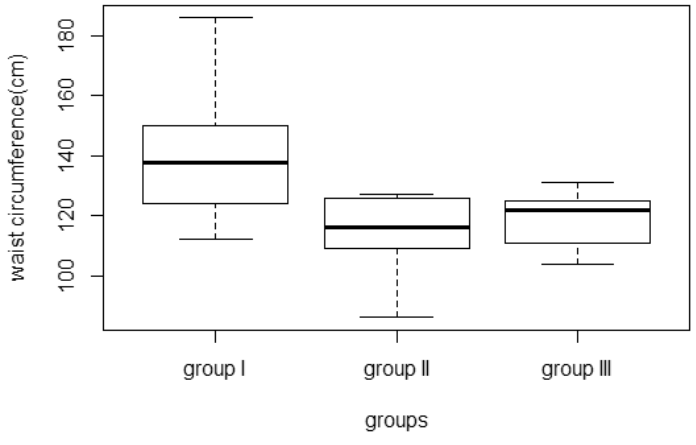
สถิติที่ใช้ทดสอบรายคู่มีหลายการทดสอบ ในที่นี้ใช้การปรับค่าของ Games-Howell ซึ่งได้นำวิธีการคำนวณของ Tukey ปรับปรุงซึ่งเหมาะกับกรณีที่ข้อมูลมีความแปรปรวนไม่เท่ากัน[19] การใช้งานในโปรแกรม R พบว่ามีอยู่ในชุดคำสั่งชื่อว่า **userfriendlyscience**[16] ดังนั้นจึงต้องติดตั้งชุดติดตั้ง (package) นี้ก่อนการใช้งาน

```
> library(userfriendlyscience)
> posthocTGH(anovadta$score,anovadta$group)
           n means variances
group I   10   141       501
group II  10   114       206
group III 10   119        76

                     diff ci.lo ci.hi   t df   p
group II-group I   -27.9 -49.7  -6.1 3.3 15 .01
group III-group I  -22.2 -42.5  -1.9 2.9 12 .03
group III-group II   5.7  -8.1  19.5 1.1 15 .54
```

รูปที่ 7.17 ผลการทดสอบรายคู่

หมายเหตุ; diff หมายถึง ความความแตกต่างระหว่างกลุ่ม, ci.lo หมายถึง ช่วงเชื่อมั่นที่ 95% ขอบเขตล่าง, ci.hi หมายถึง ช่วงเชื่อมั่นที่ 95% ขอบเขตบน, t หมายถึง ค่าสถิติ t, df หมายถึง องศาความเป็นอิสระ, p หมายถึง ค่าความน่าจะเป็น (p-value)

ผลการวิเคราะห์พบว่ากลุ่มที่ 2-1 และ 3-1 มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ สรุปผลการศึกษา จากการทดสอบค่าเฉลี่ยเส้นรอบเอวทั้งสามกลุ่มมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ (0.016) ทั้งนี้เมื่อทำการทดสอบรายคู่แล้วพบว่า กลุ่มที่ 2 (กลุ่มพนักงานร้านอาหาร) กับกลุ่มที่ 1 (กลุ่มพนักงานออฟฟิศ) มีค่าแตกต่างกัน 27.9 (95% CI: 6.1, 49.7) และ กลุ่มที่ 3 (กลุ่มเจ้าหน้าที่บันทึกข้อมูล) กับกลุ่มที่ 1 (กลุ่มพนักงานออฟฟิศ) มีค่าแตกต่างกัน 22.2 (95% CI: 1.9, 42.5) นอกจากนั้นแล้วเรายังสามารถแสดงผลด้วยกราฟ Box and Whisker Plots ดังรูปที่ 7.18



รูปที่ 7.18 เปรียบเทียบค่าเฉลี่ยเส้นรอบเอวระหว่าง 3 กลุ่ม

7.7 การเปรียบเทียบค่าเฉลี่ยมากกว่าสองกลุ่มที่อิสระต่อกัน กรณีข้อมูลมีการแจกแจงไม่ปกติ

ในกรณีที่ข้อมูลมีการแจกแจงไม่ปกติ สถิติที่ใช้ทดสอบ คือ Kruskal Wallis test[12] โดยการทดสอบค่ากลางของข้อมูล เพื่อทดสอบความแตกต่างกัน จากนั้นจึงทำการทดสอบรายคู่ต่อไป

ตัวอย่างที่ 7.7 การเปรียบเทียบค่าเฉลี่ยคะแนนความเครียดระหว่างกลุ่มการออกกำลังกาย (Physical activity) ระดับ สูง (High) ปานกลาง (Medium) และน้อย (Low) ทั้งนี้โดยใช้แบบสอบถามการวัดระดับความเครียด PHQ-9 (Patient Health Questionnaire-9) ซึ่งมีค่าอยู่ระหว่าง 0-27 ทั้งนี้ค่ายิ่งสูงยิ่งมีความเครียดมาก[20]

ตารางที่ 7.6 คะแนนความเครียดระหว่างกลุ่มออกกำลังกายระดับต่างๆ

High	Medium	Low
2	16	18
5	14	13
12	16	9
13	14	13
13	12	13
13	13	17
15	0	12
15	13	9
16	13	11
17	12	18

วิธีทำ

ขั้นตอนที่ 1 ทดสอบการแจกแจง

```
> shapiro.test(anovadta$score[anovadta$group=="high"])

        Shapiro-Wilk normality test

data:  anovadta$score[anovadta$group == "high"]
W = 0.81532, p-value = 0.02224

> shapiro.test(anovadta$score[anovadta$group=="medium"])

        Shapiro-Wilk normality test

data:  anovadta$score[anovadta$group == "medium"]
W = 0.65573, p-value = 0.0002572

> shapiro.test(anovadta$score[anovadta$group=="low"])

        Shapiro-Wilk normality test

data:  anovadta$score[anovadta$group == "low"]
W = 0.8901, p-value = 0.17
```

รูปที่ 7.19 ผลการทดสอบการแจกแจงของข้อมูล

ผลทดสอบการแจกแจงข้อมูลพบว่าในกลุ่ม High และ Medium มีการแจกแจงไม่ปกติ ดังนั้นในการเปรียบเทียบค่าเฉลี่ยระหว่าง 3 กลุ่มนี้ใช้สถิติ Kruskal Wallis test

ขั้นตอนที่ 2 ทดสอบทางสถิติโดยใช้ Kruskal Wallis test

```
> kruskal.test(anovadta$score~anovadta$group)

        Kruskal-Wallis rank sum test

data:  anovadta$score by anovadta$group
Kruskal-Wallis chi-squared = 0.002663, df = 2, p-value = 0.9987
```

รูปที่ 7.20 ผลการวิเคราะห์ข้อมูลสถิติKruskal Wallis Test

ผลการวิเคราะห์พบว่าค่ากลางของวิตามินทั้ง 4 กลุ่มไม่มีความแตกต่างกัน (p-value = 0.999)

7.8 บทสรุป

ในบทนี้ได้ศึกษาการวิเคราะห์ข้อมูลในกรณีที่ผลลัพธ์การศึกษาเป็นข้อมูลต่อเนื่อง ซึ่งการเลือกใช้สถิติต้องมีการทดสอบการแจกแจงของข้อมูลก่อน จากนั้นจึงเลือกใช้สถิติในการทดสอบความแตกต่างของค่าเฉลี่ยของข้อมูลทั้งนี้สามารถสรุปการเลือกใช้สถิติและการใช้คำสั่งในโปรแกรม R ได้ดังในตารางที่ 7.7

ตารางที่ 7.7 การเลือกใช้สถิติและคำสั่งการวิเคราะห์ข้อมูลในโปรแกรม R

การทดสอบ	**สถิติ**	**คำสั่งในโปรแกรม R**
การวิเคราะห์ข้อมูลเพื่อเปรียบเทียบค่าเฉลี่ยระหว่างสองกลุ่มที่อิสระต่อกัน	Unpaired t-test	ทดสอบการแจกแจงข้อมูล shapiro.test (ชื่อตัวแปร) ทดสอบการแปรปรวนของข้อมูล var.test(ชื่อตัวแปรที่ 1, ชื่อตัวแปรที่ 2, data = ชื่อดาต้าเฟรม) ทดสอบความแตกต่างของค่าเฉลี่ย t.test (ชื่อตัวแปรที่ 1, ชื่อตัวแปรที่ 2)
	ถ้าข้อมูลมีการแจกแจงไม่ปกติ Mann-Whitney's U test	wilcox.test(ชื่อตัวแปรที่ 1, ชื่อตัวแปรที่ 2)
การเปรียบเทียบค่าเฉลี่ยระหว่างสองกลุ่มที่ไม่อิสระต่อกัน	Paired t-test	t.test(ตัวแปรก่อนทดลอง, ตัวแปรหลังทดลอง ,paired = TRUE)
	ถ้าข้อมูลมีการแจกแจงไม่ปกติ	wilcox.test(ตัวแปรก่อนทดลอง, ตัวแปรหลังทดลอง, paired=TRUE)
การเปรียบเทียบค่าเฉลี่ยมากกว่าสองกลุ่มที่อิสระต่อกัน	One-way ANOVA	ทดสอบความแปรปรวน bartlett.test(ตัวแปรต่อเนื่อง~ตัวแปรกลุ่ม, data=ชื่อดาต้าเฟรม) oneway.test(ตัวแปรต่อเนื่อง~ตัวแปรกลุ่ม) library(userfriendlyscience) posthocTGH(ตัวแปรต่อเนื่อง, ตัวแปรกลุ่ม)
	ถ้าข้อมูลมีการแจกแจงไม่ปกติ	kruskal.test(ตัวแปรต่อเนื่อง, ตัวแปรกลุ่ม)

7.9 แบบฝึกหัด

1. การทดลองเพื่อเปรียบเทียบระดับไขมันดี (High-density lipoprotein cholesterol: HDL-C) ระหว่างกลุ่มรับประทานอาหารที่มีไขมันเป็นปกติ (Regular-fat) กับ กลุ่มที่มีการรับประทานอาหารไขมันต่ำ (Low-fat) จำนวนกลุ่มละ 11 คน ได้ผลการทดลองดังนี้

Regular-fat (G1)	50	40	51	53	54	42	43	59	53	57	59
Low-fat(G2)	54	60	51	53	59	58	55	56	54	50	52

จากข้อมูลดังกล่าวจงวิเคราะห์ข้อมูลและสรุปผลการทดลอง

2. การเปรียบเทียบคะแนนคุณภาพชีวิต (0 ถึง 100) ระหว่างกลุ่มที่มีดัชนีมวลกายปกติ ต่ำกว่าเกณฑ์และอ้วนข้อมูลดังตาราง

ปกติ	ต่ำกว่าเกณฑ์	อ้วน
81	74	75
87	83	75
96	83	74
87	85	76
81	72	83
85	79	83
87	81	82
85	83	71
95	70	82
85	80	85

จากข้อมูลจงวิเคราะห์ข้อมูลและสรุปผลการทดลอง

7.10 เอกสารอ้างอิง

1. Rosner B. Fundamentals of biostatistics: Seventh edition. Boston : Brooks/Cole, Cengage Learning; 2011.
2. Schober P, Vetter TR. Two-Sample Unpaired t Tests in Medical Research. Anesthesia & Analgesia. 2019;129(4):911.
3. Ghasemi A, Zahediasl S. Normality tests for statistical analysis: a guide for non-statisticians. Int J Endocrinol Metab. 2012;10(2):486-9.
4. Kim HY. Statistical notes for clinical researchers: assessing normal distribution (1). Restor Dent Endod. 2012;37(4):245-8.
5. Mishra P, Pandey CM, Singh U, Gupta A, Sahu C, Keshri A. Descriptive statistics and normality tests for statistical data. Ann Card Anaesth. 2019;22(1):67-72.
6. Kim TK. T test as a parametric statistic. Korean J Anesthesiol. 2015;68(6):540-6.
7. Witt PL, McGrain P. Comparing two sample means t tests. Phys Ther. 1985;65(11):1730-3.
8. Schober P, Vetter TR. Two-Sample Unpaired t Tests in Medical Research. Anesthesia & Analgesia. 2019;129(4).
9. Hart A. Mann-Whitney test is not just a test of medians: differences in spread can be important. Bmj. 2001;323(7309):391-3.
10. Kim TK. T test as a parametric statistic. Korean J Anesthesiol. 2015;68(6):540-6.
11. Rosner B, Glynn RJ, Lee ML. The Wilcoxon signed rank test for paired comparisons of clustered data. Biometrics. 2006;62(1):185-92.
12. Altman DG. Practical statistics for medical research: CRC press; 1990.
13. Thompson HW, Mera R, Prasad C. The Analysis of Variance (ANOVA). Nutr Neurosci. 1999;2(1):43-55.
14. Kim TK. Understanding one-way ANOVA using conceptual figures. Korean J Anesthesiol. 2017;70(1):22-6.
15. Bewick V, Cheek L, Ball J. Statistics review 10: further nonparametric methods. Crit Care. 2004;8(3):196-9.
16. Peters G-J, Peters MG-J, XLConnect S. Package 'userfriendlyscience'. Vienna: R Foundation for Statistical Computing. 2018.

17. Games PA, Howell JF. Pairwise Multiple Comparison Procedures with Unequal N's and/or Variances: A Monte Carlo Study. Journal of Educational Statistics. 1976;1(2):113-25.
18. Brown AM. A new software for carrying out one-way ANOVA post hoc tests. Comput Methods Programs Biomed. 2005;79(1):89-95.
19. Theodorsson-Norheim E. Kruskal-Wallis test: BASIC computer program to perform nonparametric one-way analysis of variance and multiple comparisons on ranks of several independent samples. Comput Methods Programs Biomed. 1986;23(1):57-62.
20. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med. 2001;16(9):606-13.

บทที่ 8

การวิเคราะห์ข้อมูลจำแนกประเภท

ในบทที่ผ่านมาเป็นการวิเคราะห์ข้อมูลแบบต่อเนื่อง ในบทนี้จะได้ศึกษาการวิเคราะห์ข้อมูลที่มีผลลัพธ์การศึกษาเป็นแบบจำแนกประเภทหรือตัวแปรกลุ่ม (Categorical data) การวิเคราะห์ข้อมูลใช้สถิติ Chi-square test และ Fisher's exact test ดังต่อไปนี้

8.1 การทดสอบความสัมพันธ์ของตัวแปรจำแนกประเภท 2 ตัวแปร กรณีตารางการณ์จรแบบ 2x2

การทดสอบความสัมพันธ์ของตัวแปรกลุ่ม 2 ตัวแปรที่แต่ละตัวแปรแบ่งออกเป็น 2 ระดับ ดังในตารางที่ 8.1

ตารางที่ 8.1 ตารางการณ์จร (Contingency table) แบบ 2x2

ปัจจัยที่ต้องการศึกษา (Factor)	ผลลัพธ์ที่ต้องการศึกษา (Outcome)	
	เกิดเหตุการณ์ (Event)	ไม่เกิดเหตุการณ์ (Non-event)
สัมผัสปัจจัย (Exposure)	a	b
ไม่สัมผัสปัจจัย (Non-exposure)	c	d

จากตารางดังกล่าวสามารถใช้สถิติทดสอบความสัมพันธ์ระหว่างสองตัวแปรเรียกว่า Chi-square test ซึ่งมีข้อตกลงเบื้องต้นที่ของการใช้สถิติตัวนี้คือ[1-4]

1. ถ้าค่าคาดหวังในแต่ละช่องมีค่าน้อยกว่า 5 น้อยกว่า 20% ของจำนวนช่องทั้งหมด ใช้สถิติ Chi-square test
2. ถ้าค่าคาดหวังในแต่ละช่องมีค่าน้อยกว่า 5 มากกว่า 20% ของจำนวนช่องทั้งหมด ใช้สถิติ Fisher's exact test

จากตารางการณ์จรแบบ 2x2 สามารถคำนวณค่าความน่าจะเป็นนอกจากนั้นแล้วในทางวิทยาการระบาดยังสามารถคำนวณค่าความเสี่ยงสัมพัทธ์ (Odds ratio หรือ Relative risk) ด้วย เพราะเมื่อทดสอบความสัมพันธ์ของสองตัวแปรพบว่ามีความสัมพันธ์กัน ลำดับต่อมาจะทำการคำนวณระดับความสัมพันธ์ (Strength of association) ว่ามีมากน้อยแค่ไหน ทั้งนี้ขึ้นอยู่กับรูปแบบงานวิจัยที่ทำการศึกษา

รูปแบบงานวิจัยแบบ Case-control study, Cross-sectional study หรือ Cohort study (ในกรณีที่ศึกษาในโรคที่เกิดน้อย) นิยมรายงานค่าความเสี่ยงสัมพัทธ์เรียกว่า Odds Ratio (OR)[5-8]

ในกรณีที่รูปแบบงานวิจัยที่เป็น Cohort study (Prospective study) นิยมรายงานค่าความเสี่ยงสัมพัทธ์ที่เรียกว่า Relative Risk (RR) อย่างไรก็ตามถ้าเป็นโรคที่เกิดขึ้นน้อยค่า RR กับ OR จะมีค่าใกล้เคียงกัน ในกรณีผลลัพธ์ที่ต้องการศึกษาพบได้น้อย[9,10] (Rare disease) นิยมรายงานด้วยค่า OR

ในกรณีที่รูปแบบงานวิจัยที่เป็น Cohort study ที่มีเวลาเข้ามาเกี่ยวข้องกับการเกิดผลลัพธ์ (Time-to event) เช่น การศึกษาปัจจัยเสี่ยงของการเกิดโรคมะเร็ง (Time-to event) นิยมรายงานค่าความเสี่ยงสัมพัทธ์ที่เรียกว่า Hazard Ratio (HR)[11-14]

ทั้งนี้สามารถคำนวนค่า OR และช่วงเชื่อมั่นที่ 95% (95% Confidence Interval: 95% CI) ได้โดยใช้ข้อมูลใน ตารางที่ 8.1 ดังนี้

การคำนวณค่า OR และช่วงเชื่อมั่น 95% สามารถทำได้ดังนี้[5]

$$OR = \frac{\frac{a}{b}}{\frac{c}{d}} = \frac{a \times d}{b \times c}$$

$$95\%\ CI = OR \times e^{\left[\mp 1.96\sqrt{\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}}\right]}$$

การคำนวณค่า RR และช่วงเชื่อมั่น 95%[15-17]

$$RR = \frac{\frac{a}{(a+b)}}{\frac{c}{(c+d)}} = \frac{a \times (c+d)}{c \times (a+b)}$$

$$95\%\ CI = RR \times e^{\left[\mp 1.96\sqrt{\frac{1}{a\times(a+b)}+\frac{1}{c\times(c+d)}}\right]}$$

การแปลผลค่าความเสี่ยงสัมพัทธ์กลุ่มคนที่สัมผัสปัจจัยเสี่ยงมีโอกาสเสี่ยงต่อการเกิดโรคเป็นกี่เท่า (OR, RR หรือ HR) เมื่อเทียบกับกลุ่มคนที่ไม่ได้สัมผัสปัจจัยเสี่ยง

ตัวอย่างที่ 8.1 การศึกษาของ Laopaiboon M และคณะ (2014) เรื่อง การตั้งครรภ์ในสตรีอายุมากและผลลัพธ์ของการตั้งครรภ์ การประเมินแบบหลายประเทศ (Advanced maternal age and pregnancy outcomes: a multicountry assessment.)[18] ได้ผลการศึกษาดังนี้

ตารางที่ 8.2 ความถี่และร้อยละของอายุแม่ตั้งครรภ์และผลลัพธ์ของการตั้งครรภ์

	Maternal age (years)				
Outcomes	Total	20–34	35–39	40–44	≥45
Maternal (n)	276,291	238,504	29,245	7,015	1,527
Maternal near miss[b]	1,372 (5)	1,007 (4)	243 (9)	98 (14)	24 (16)
Maternal death[b]	270 (1)	207 (1)	43 (2)	15 (2)	5 (3)
Severe maternal outcome[b]	1,642 (6)	1,214 (5)	286 (10)	113 (17)	29 (20)
Neonatal; total births (n)	276,291	238,504	29,245	7,015	1,527
Live births (n)	270,744	234,057	28,445	6,780	1,462
Preterm birth (< 37 weeks)[a]	16,966 (61)	14,352 (60)	1,964 (67)	548 (78)	102 (67)
Stillbirths[a]	5,547 (20)	4,447 (19)	800 (27)	235 (34)	65 (43)
Early neonatal mortality[b]	2,297 (9)	1,962 (8)	249 (9)	71 (11)	15 (10)
Perinatal mortality[a]	7,776 (28)	6,357 (27)	1,036 (35)	303 (43)	80 (53)
Low birthweight (< 2500 g)[b]	28,543 (105)	24,985 (107)	2,649 (93)	774 (114)	135 (92)
Neonatal intensive care unit admission[b]	17,247 (64)	14,498 (62)	2,044 (72)	633 (93)	72 (49)
Apgar score < 7 at 5 minutes[b]	6,896 (26)	5,887 (25)	721 (25)	214 (32)	74 (51)

a: per 1000 total births., b: per 1000 live births.

ที่มา: ดัดแปลงมาจากLaopaiboon M, Lumbiganon P, **Intarut N**, Mori R, Ganchimeg T, Vogel JP, et al. Advanced maternal age and pregnancy outcomes: a multicountry assessment. BJOG. 2014;121 Suppl 1:49-56.Table 2. Prevalence of adverse pregnancy outcomes in relation to maternal age.[18]

จากตารางที่ 8.2 ถ้าสนใจปัจจัยอายุแม่ตั้งครรภ์ (อายุ 20-34 ปี, อายุมากกว่าหรือเท่ากับ 35 ปี) กับการเข้าพักฟื้นที่ห้องอภิบาลทารกแรกเกิด (Neonatal intensive care unit admission NICU) จากโจทย์ดังกล่าวสามารถเขียนเป็นสมมุติฐานทางสถิติและตาราง 2x2 ได้ดังนี้

ตัวแปรอายุแม่ตั้งครรภ์แบ่งออกเป็น อายุ 20-34 ปี และ อายุมากกว่าหรือเท่ากับ 35 ปี

ตัวแปรการเข้าพักฟื้นที่ห้องอภิบาลทารกแรกเกิด เข้าพักฟื้น NICU และ ไม่เข้าพักฟื้น NICU

สมมุติฐานงานวิจัย :อายุแม่ตั้งครรภ์มีความสัมพันธ์กับการเข้าพักฟื้นที่ห้องอภิบาลทารกแรกเกิด

สมมุติฐานทางสถิติ :

H_0: อายุแม่ตั้งครรภ์ ไม่มีความสัมพันธ์กับการเข้าพักฟื้นที่ห้องอภิบาลทารกแรกเกิด

H_a: อายุแม่ตั้งครรภ์มีความสัมพันธ์กับการเข้าพักฟื้นที่ห้องอภิบาลทารกแรกเกิด

จากนั้นทำการทดสอบสมมุติฐานทางสถิติ

ตารางที่ 8.3 ตารางแจกแจงความถี่ของข้อมูลอายุของแม่ตั้งครรภ์และการเข้าพักฟื้นที่ห้องอภิบาลทารกแรกเกิด

อายุแม่ตั้งครรภ์ (ปี)	การเข้าพักฟื้นที่ห้องอภิบาลทารกแรกเกิด		รวม
	ใช่	ไม่ใช่	
≥35	3,558	8,144	11,702
20-34	24,985	234,057	259,042
รวม	28,543	242,201	270,744

จากข้อมูลในตารางที่ 8.3 ทำการวิเคราะห์ข้อมูลโดยใช้สถิติ Chi-square test ในโปรแกรม R สามารถใช้คำสั่งได้สองรูปแบบดังนี้

รูปแบบที่ 1 ใช้คำสั่งจากชุดคำสั่งที่ให้มาในโปรแกรม R (chisq.test)

```
>df<- matrix(c(2749,33938,14498,219559), ncol=2)
> #test expected value
> pchi2<-chisq.test(df,correct = FALSE)
> pchi2$expected
          [,1]      [,2]
[1,]  2337.044  14909.96
[2,] 34349.956 219147.04
> pchi2

        Pearson's Chi-squared test

data:  df
X-squared = 89.713, df = 1, p-value < 2.2e-16
```

ค่าคาดหวัง

รูปที่ 8.1 ผลการวิเคราะห์ข้อมูลรูปแบบที่ 1 โดยใช้คำสั่ง chisq.test()

จากผลการวิเคราะห์ดังกล่าว ค่าความคาดหวังมีค่ามากกว่า 5 ทุกเซลล์ ดังนั้นจึงสามารถใช้สถิติ Chi-square test ได้ และจากผลการวิเคราะห์โดยใช้คำสั่ง chisq.test() ได้ค่าความน่าจะเป็นเท่ากับ < 2.2e-16 หรือเขียนเป็น < 0.001 สรุปได้ว่าอายุแม่ตั้งครรภ์มีความสัมพันธ์กับการเข้าพักฟื้นที่ห้องอภิบาลทารกแรกเกิดอย่างมีนัยสำคัญทางสถิติ (p-value < 0.001)

รูปแบบที่ 2 ใช้คำสั่ง cci จากชุดคำสั่ง epiDisplay (epiDisplay package)[19]

```
> library(epiDisplay)
Loading required package: foreign
Loading required package: survival
Loading required package: MASS
Loading required package: nnet
>cci(2749,33938,14498,219559)

          Exposure
Outcome    Non-exposed Exposed  Total
  Negative      219559   33938 253497
  Positive       14498    2749  17247
  Total         234057   36687 270744

OR =  1.23
95% CI =  1.18, 1.28
Chi-squared = 89.71, 1 d.f., P value = 0
Fisher's exact test (2-sided) P value = 0
```

รูปที่ 8.2 ผลการวิเคราะห์ข้อมูลรูปแบบที่ 2 โดยใช้คำสั่ง cci()

ผลการวิเคราะห์ข้อมูลทั้งสองแบบได้ค่าความน่าจะเป็นเท่ากันคือ <0.001 อย่างไรก็ตาม คำสั่งทางระบาดวิทยา cci() ที่อยู่ในชุดคำสั่ง epiDisplay จะได้ข้อมูลเกี่ยวกับความเสี่ยงสัมพัทธ์ (OR) ช่วงเชื่อมั่นที่ 95% และ สถิติ Fisher's exact test มาด้วย

การแปลผลการวิเคราะห์ข้อมูลอายุแม่ตั้งครรภ์มีความสัมพันธ์กับการเข้าพักฟื้นที่ห้องอภิบาลทารกแรกเกิดอย่างมีนัยสำคัญทางสถิติ (p-value < 0.001) โดยที่แม่ที่อายุมากกว่า 34 ปี มีโอกาสเสี่ยงต่อการเข้าพักฟื้นที่ห้องอภิบาลทารกแรกเกิด (NICU) เป็น 1.23 (95% CI: 1.18, 1.28) เท่าเมื่อเทียบกับแม่ที่อายุอยู่ระหว่าง 20-34 ปี

จากตัวอย่างที่ 8.1 เป็นการทดสอบความสัมพันธ์ของตัวแปรกลุ่มสองตัวแปรแบบตาราง 2x2 ในกรณีปัจจัยที่ต้องศึกษามีมากกว่า 2 กลุ่ม หรือเรียกตารางการณ์จรว่า ตาราง 2xR ดังแสดงในตารางที่ 8.4 เป็นตารางแบบ 2x3

ตารางที่ 8.4 ตารางการณ์จร (Contingency table) แบบ 2x3

ปัจจัยที่ต้องการศึกษา (Factor)	ผลลัพธ์ที่ต้องการศึกษา (Outcome)	
	เกิดเหตุการณ์ (Event)	ไม่เกิดเหตุการณ์ (Non-event)
Level 1	a	b
Level 2	c	d
Level 3	e	f

การวิเคราะห์ความสัมพันธ์ทางสถิติจะใช้ Chi-square test หรือ Fisher's exact test เหมือนตัวอย่างที่ผ่านมา อย่างไรก็ตามการคำนวณค่า OR ต้องทำการแจกแจงตารางให้เป็นแบบ 2x2 ก่อน หรือเรียกว่าตัวแปรหุ่น (Dummy variable)[20] จึงจะคำนวณค่า OR แต่ละคู่ได้

จากข้อมูลในตารางที่ 8.2 ถ้าเราสนใจศึกษา ปัจจัยอายุแม่ตั้งครรภ์ (อายุ 20-34 ปี, อายุ 35-39 ปี, อายุ 40-44 ปี, มากกว่าหรือเท่ากับ 45 ปี) กับการคลอดก่อนกำหนด (อายุครรภ์น้อยกว่า 37 สัปดาห์: Preterm) สามารถแจกแจงความถี่ลงในตาราง 2x4 ได้ตารางที่ 8.5

ตารางที่ 8.5 ตารางแจกแจงความถี่ของอายุของแม่ตั้งครรภ์กับการคลอดตาราง 2x4

อายุของแม่ตั้งครรภ์ (ปี)	คลอดก่อนกำหนด		รวม
	Preterm (< 37 สัปดาห์)	Term (≥ 37 สัปดาห์)	
≥45	102	1,360	1,462
40-44	548	6,232	6,780
35-39	1,964	26,481	28,445
20-34	14,352	219,705	234,057
รวม	16,966	253,778	270,744

การทดสอบทางสถิติ Chi-square test สามารถทำได้ดังนี้

```
>df<- matrix(c(102,1360,548,6232,1964,26481,14352,219705), ncol=2)
> #test expected value
> pchi2<-chisq.test(df,correct = FALSE)
> #expected value
> pchi2$expected
          [,1]      [,2]
[1,]  62.89326  2003.107
[2,] 847.53687 26993.463
[3,] 453.58641 14446.414
[4,] 6877.98346 219059.017
> pchi2

        Pearson's Chi-squared test

data:  df
X-squared = 427.51, df = 3, p-value < 2.2e-16
```

ค่าคาดหวัง

รูปที่ 8.3 ผลการวิเคราะห์ข้อมูลและค่าคาดหวัง (Expected value)

จากผลการวิเคราะห์ข้อมูลดังกล่าวพบว่า ค่าคาดหวังมีค่ามากกว่า 5 ทุกเซลล์ดังนั้น จึงใช้สถิติ Chi-square test ได้ พบว่าอายุแม่ตั้งครรภ์มีความสัมพันธ์กับการคลอดก่อนกำหนด (Preterm < 37 สัปดาห์) อย่างมีนัยสำคัญทางสถิติ (p-value < 0.001) ลำดับต่อมาทำการคำนวณค่าความเสี่ยงสัมพัทธ์รายคู่ดังนี้

การคำนวณค่า OR สามารถคำนวณได้ทั้งหมด 3 คู่ โดยมีกลุ่มอายุ 20-34 ปี เป็นกลุ่มอ้างอิง

การคำนวณค่า OR ของคู่ที่ 1

ตารางที่ 8.6 ตารางแจกแจงความถี่ของกลุ่มอายุ 20-34 ปี เปรียบเทียบกับ 35-39 ปี

อายุของแม่ตั้งครรภ์ (ปี)	คลอดก่อนกำหนด		OR (95% CI)
	Preterm	Term	
35-39	1,964	26,481	1.14 (1.08, 1.19)
20-34	14,352	219,705	

คำสั่งในโปรแกรม R: *cci(1964,26481,14352,219705)*

การคำนวณค่า OR ของคู่ที่ 2

ตารางที่ 8.7 ตารางแจกแจงความถี่ของกลุ่มอายุ 20-34 ปี เปรียบเทียบกับ 40-44 ปี

อายุของแม่ตั้งครรภ์ (ปี)	คลอดก่อนกำหนด		OR (95% CI)
	Preterm	Term	
40-44	548	6,232	1.35 (1.23, 1.47)
20-34	14,352	219,705	

คำสั่งในโปรแกรม R: *cci(548,6232,14352,219705)*

การคำนวณค่า OR ของคู่ที่ 3

ตารางที่ 8.8 ตารางแจกแจงความถี่ของกลุ่มอายุ 20-34 ปี เปรียบเทียบกับ ≥ 45 ปี

อายุของแม่ตั้งครรภ์ (ปี)	คลอดก่อนกำหนด		OR (95% CI)
	Preterm	Term	
≥ 45	102	1,360	1.15 (0.94, 1.41)
20-34	14,352	219,705	

คำสั่งในโปรแกรม R: *cci(102,1360,14352,219705)*

เมื่อคำนวณค่า OR ในแต่ละคู่แล้วสามารถนำมาเขียนรายงานรวมกัน โดยที่กลุ่มอ้างอิงจะมีค่าเท่ากับ 1 ได้ดังนี้

ตารางที่ 8.9 ค่าความเสี่ยงสัมพัทธ์ระหว่างอายุของแม่ตั้งครรภ์กับการคลอดก่อนกำหนด

อายุของแม่ตั้งครรภ์ (ปี)	คลอดก่อนกำหนด		OR (95% CI)
	Preterm	Term	
≥ 45	102	1,360	1.15 (0.94, 1.41)
40-44	548	6,232	1.35 (1.23, 1.47)
35-39	1,964	26,481	1.14 (1.08, 1.19)
20-34	14,352	219,705	1

อย่างไรก็ตามสามารถใช้คำสั่งของโปรแกรม R โดยใช้การวิเคราะห์แบบ Simple logistic regression[21] (จะได้ศึกษาในบทที่ 10) เพื่อให้สามารถคำนวณค่า OR และ 95% CI ได้ในครั้งเดียว

```
> outcome<-c(1,1,1,1,0,0,0,0)
>var<-c(3,2,1,0,3,2,1,0)
>freq<-c(102,548,1964,14352,1360,6232,26481,219705)
>df<-data.frame(outcome,var,freq)
>df$var<-as.factor(df$var)
>mylogit<- glm(outcome ~ var,
+                data = df, family = "binomial",weights =freq )
>logistic.display(mylogit)

Logistic regression predicting outcome

            OR(95%CI)         P(Wald's test) P(LR-test)
var: ref.=0                                  < 0.001
   1        1.14 (1.08,1.19)  < 0.001
   2        1.35 (1.23,1.47)  < 0.001
   3        1.15 (0.94,1.41)  0.18

Log-likelihood = -63386.7976
No. of observations = 8
AIC value = 126781.5951
```

รูปที่ 8.4 ผลการวิเคราะห์ข้อมูลเพื่อคำนวณค่า Odds ratio โดยใช้ Simple linear regression

ผลการวิเคราะห์แบบ Simple logistic regression ได้ค่า OR เหมือนกับการวิเคราะห์รายคู่ที่ผ่านมา

8.2 การวิเคราะห์ข้อมูลโดยใช้สถิติ Fisher's exact test

ในกรณีที่เราไม่สามารถวิเคราะห์ข้อมูลโดยใช้ Chi-square test ได้เนื่องจากละเมิดข้อตกลงเบื้องต้นของการใช้สถิติ จะใช้สถิติ Fisher's exact test[22-24] ในการทดสอบความสัมพันธ์ของสองตัวแปรดังตัวอย่างที่ 8.2

ตัวอย่างที่ 8.2 การศึกษาแบบ Case-control study เพื่อทดสอบความสัมพันธ์ระหว่างการได้รับควันบุหรี่มือสองกับการเป็นโรคมะเร็งปอด ได้ข้อมูลดังนี้

ตารางที่ 8.10 ความถี่ของข้อมูลการได้รับควันบุหรี่มือสองกับการเป็นโรคมะเร็งปอด

	โรคมะเร็งปอด	
ได้รับควันบุหรี่มือสอง	เป็นโรค	ไม่เป็นโรค
ได้รับ	52	2
ไม่ได้รับ	3	16

จากข้อมูลที่ได้สามารถทำการวิเคราะห์ได้ 2 วิธีดังนี้

รูปแบบที่ 1 ใช้คำสั่งจากชุดคำสั่งที่ให้มาในโปรแกรม R (fisher.test)

```
> dat2 <- matrix(c(52,2,3,16), ncol=2)
> pchi2<-chisq.test(dat2,correct = FALSE)
Warning message:
In chisq.test(dat2, correct = FALSE) :
Chi-squared approximation may be incorrect
> pchi2

        Pearson's Chi-squared test

data:  dat2
X-squared = 49.034, df = 1, p-value = 2.515e-12

>fisher.test(dat2)

        Fisher's Exact Test for Count Data

data:  dat2
p-value = 2.538e-11
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
   17.45888 1524.70497
sample estimates:
odds ratio
  115.7632
```

รูปที่ 8.5 ผลการวิเคราะห์ข้อมูลรูปแบบที่ 1 โดยใช้คำสั่ง fisher.test()

ผลการวิเคราะห์โดยวิธีที่ 1 เมื่อทำการวิเคราะห์โดยใช้ Chi-square test พบว่ามีข้อความเตือนมาให้ทราบว่าสถิติ Chi-square test อาจจะไม่เหมาะสม (Chi-squared approximation may be incorrect) ดังนั้นการรายงานค่าสถิติที่ได้จาก Chi-square test อาจจะไม่เหมาะสม จึงใช้ Fisher's exact test แทนซึ่งได้ค่า p-value = 2.538e-11 หรือ p-value < 0.001

รูปแบบที่ 2 ใช้คำสั่งจากชุดคำสั่ง epiDisplay (cci)[19]

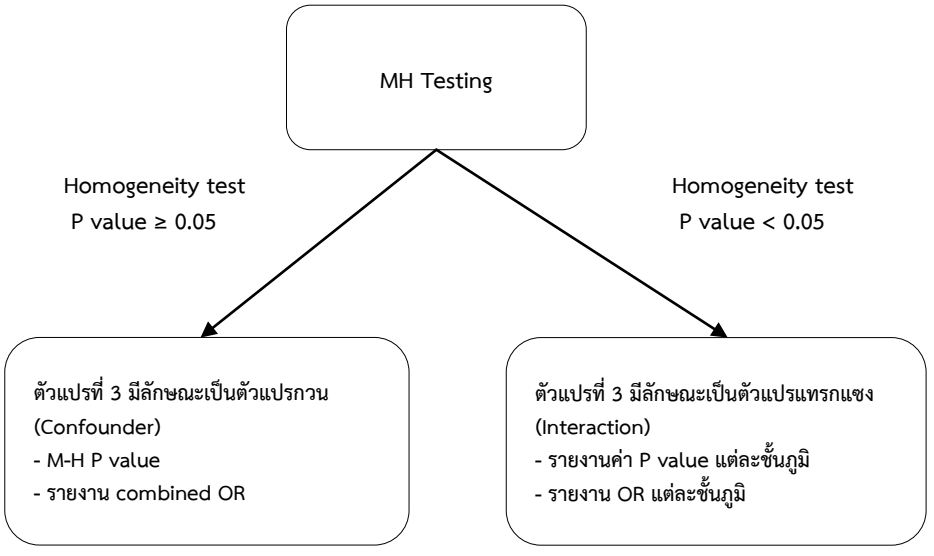
```
> library(epiDisplay)
>cci(52,2,3,16)

          Exposure
Outcome    Non-exposed Exposed Total
  Negative          16       2    18
  Positive           3      52    55
  Total             19      54    73

OR =  138.67
95% CI =  21.27, 904.11
Chi-squared = 49.03, 1 d.f., P value = 0
Fisher's exact test (2-sided) P value = 0

Warning message:
In chisq.test(table1, correct = FALSE) :
Chi-squared approximation may be incorrect
```

รูปที่ 8.6 ผลการวิเคราะห์ข้อมูลรูปแบบที่ 2 โดยใช้คำสั่ง cci()

การวิเคราะห์แบบที่ 2 นี้ (รูปที่ 8.6) ใช้คำสั่ง cci จากชุดคำสั่งของ epiDisplay ซึ่งพบว่ามีการแจ้งเตือนถึงการใช้สถิติ Chi-square test ว่าอาจจะไม่เหมาะสม (Chi-squared approximation may be incorrect) จึงใช้ค่า p-value จาก Fisher's exact test แทนการแปลผลการวิเคราะห์ข้อมูล การได้รับควันบุหรี่มือสองมีความสัมพันธ์การเป็นโรคมะเร็งปอดอย่างมีนัยสำคัญทางสถิติ (p-value < 0.001)ทั้งนี้พบว่าคนที่ได้รับควันบุหรี่มือสองมีโอกาสเสี่ยงต่อการเกิดโรคมะเร็งปอดเป็น 138.67 เท่าเมื่อเทียบกับคนที่ไม่ได้รับควันบุหรี่มือสอง (95% CI: 21.27, 904.11)

8.3 การวิเคราะห์ข้อมูลแบบ Mantel-Haenszelmethod

ในการทดสอบความสัมพันธ์ของตัวแปรจำแนกประเภทสองตัวแปร อาจมีปัจจัยที่ 3 เข้ามาเกี่ยวข้องมีอิทธิพลต่อความสัมพันธ์ของสองตัวแปรที่ทดสอบ การทดสอบปัจจัยที่ 3 ที่คาดว่ามีความเกี่ยวข้องกับความสัมพันธ์แบบสองตัวแปรนั้นใช้การทดสอบที่เรียกว่า Mantel-Haenszel testing หรือ Cochran–Mantel–Haenszel test (MH)[25-28]



รูปที่ 8.7 การทดสอบสถิติ Mantel-Haenszel

จากรูปที่ 8.7 การทดสอบ MH ถ้าค่าความน่าจะเป็น (p-value) มีค่ามากกว่า 0.05 หมายความว่า ตัวแปรที่ 3 นั้นมีลักษณะเป็น Confounding effect (ตัวแปรกวน) และการรายงานผลความเสี่ยงสัมพัทธ์ควรใช้ Combined OR ในขณะเดียวกันถ้าค่า p-value น้อยกว่า 0.05 ตัวแปรที่ 3 มีลักษณเป็น Interaction effect (ตัวแปรแทรกแซง) และการรายงานผลความเสี่ยงสัมพัทธ์ควรแยกรายงานผลเป็นชั้นภูมิตามลักษณะตัวแปรที่ 3 เช่นตัวแปรที่ 3 มี 4 กลุ่ม ควรรายงานค่าความเสี่ยงสัมพัทธ์ทั้ง 4 กลุ่ม

ตัวอย่างที่ 8.3 Chisha Y และคณะ (2017)[29] ได้ทำการศึกษาความชุกและปัจจัยที่มีความสัมพันธ์ของโรคเบาหวานเข้าจอประสาทตา (Diabetic Retinopathy) รูปแบบการศึกษาเป็นแบบภาคตัดขวาง ในตัวอย่างนี้ทำการทดสอบความสัมพันธ์ของประวัติการเป็นโรคเบาหวานในครอบครัวกับการเกิดโรคเบาหวานเข้าจอประสาทตา โดยทดสอบว่าระยะเวลาการเป็นโรคเบาหวาน มีผลต่อความสัมพันธ์ดังกล่าวหรือไม่ การแจกแจงข้อมูลในดังตารางที่ 8.11

ตารางที่ 8.11 ตารางแจกแจงความถี่ของประวัติการเป็นโรคเบาหวานในครอบครัวกับการเกิดเบาหวานเข้าจอประสาทตาแบ่งตามระยะเวลาการเป็นโรคเบาหวาน

ชั้นภูมิ (Stratum)		เบาหวานเข้าจอประสาทตา	
ระยะเวลาการเป็นโรคเบาหวาน	ประวัติการเป็นโรคเบาหวานในครอบครัว	เป็น	ไม่เป็น
0-6 ปี	มี	2	21
	ไม่มี	13	177
≥6 ปี	มี	11	4
	ไม่มี	9	32

จากข้อมูลในตารางที่ 8.11 สามารถทำการวิเคราะห์โดยใช้โปรแกรม R และใช้ชุดคำสั่งจาก epiDisplay (mhor)[19] ได้ดังนี้

```
> library(epiDisplay)
Loading required package: foreign
Loading required package: survival
Loading required package: MASS
Loading required package: nnet
>tbl<- array(c(2,13,21,177,11,9,4,32),
+              dim=c(2,2,2),
+              dimnames=list(histdm=c("Yes","No"),
+                            retiodm=c("Case","Control"),
+                            duradm=c("0-6yrs",">=6")))
>tbl
, ,duradm = 0-6yrs

retiodm
histdm Case Control
   Yes    2      21
   No    13     177

, ,duradm = >=6

retiodm
histdm Case Control
   Yes   11       4
   No     9      32

>mhor(mhtable=tbl,decomal=4,graph = TRUE, design = "case control")

Stratified analysis by  duradm
                OR lower lim. upper lim. P value
duradm 0-6yrs  1.29      0.133       6.36 0.66835
duradm>=6     9.29      2.132      50.31 0.00101
M-H combined   4.13      1.701      10.02 0.00234

M-H Chi2(1) = 9.26 , P value = 0.002
Homogeneity test, chi-squared 1 d.f. = 3.49 , P value = 0.062
```

รูปที่ 8.8 ผลการวิเคราะห์ประวัติการเป็นโรคเบาหวานในครอบครัวกับการเกิดเบาหวานเข้าจอประสาทตาโดยคิดผลกระทบจากตัวแปรระยะเวลาการเป็นโรคเบาหวาน

จากผลการวิเคราะห์ดังกล่าวสรุปได้ว่า การทดสอบความเท่ากันของค่าความเสี่ยงสัมพัทธ์ (Homogeneity testing) มีค่า p-value เท่ากับ 0.062 หมายความว่าตัวแปรระยะเวลาการเป็นโรคเบาหวานมีลักษณะเป็นตัวแปรกวนดังนั้นใช้สถิติ MH testing (M-H combined) ในการรายงานผล

การแปลผล หลังจากควบคุมตัวแปรด้านระยะเวลาการเป็นโรคเบาหวานแล้วพบว่า การมีประวัติการเป็นโรคเบาหวานในครอบครัวกับการเกิดเบาหวานเข้าจอประสาทตามีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติ (p-value = 0.002) โดยพบว่าผู้ที่มีประวัติการเป็นโรคเบาหวานในครอบครัวมีโอกาสเสี่ยงต่อการเกิดเบาหวานเข้าจอประสาทตาเป็น 4.13 เท่า (95% CI: 1.70, 10.02) เมื่อเทียบกับคนที่ไม่มีประวัติการเป็นโรคเบาหวานในครอบครัว

8.4 บทสรุป

แนวทางการวิเคราะห์ข้อมูลในกรณีที่ข้อมูลเป็นชนิดจำแนกประเภทจำนวนสองกลุ่มใช้สถิติ Chi-square test หรือ Fisher's exact testนอกจากนั้นแล้วถ้ามีปัจจัยตัวแปรที่ 3 ซึ่งเป็นตัวแปรกลุ่มมีผลต่อความสัมพันธ์ของปัจจัยที่ทดสอบ จะใช้สถิติ Mantel-Haenszel ทดสอบ ทั้งนี้การวิเคราะห์ข้อมูลและคำสั่งที่ใช้ในโปรแกรม R สรุปในตารางที่ 8.12

ตารางที่ 8.12 สรุปสถิติและคำสั่งในการวิเคราะห์ของมูลตัวแปรกลุ่ม

การทดสอบ	**สถิติ**	**คำสั่งในโปรแกรม R**
การทดสอบความสัมพันธ์ของตัวแปรจำแนกประเภท 2 ตัวแปร กรณีเป็นตารางการณ์จรแบบ 2x2 และ 2xR	Chi-square test	chisq.test(ชื่อตารางข้อมูล,correct = FALSE)
	Fisher's exact test	fisher.test(ชื่อตารางข้อมูล) *cci(จำนวน case exposed, จำนวน control exposed, จำนวน case unexposed, จำนวน control unexposed)* ในกรณีที่ข้อมูลถูกเก็บไว้ในชุดข้อมูล *cc(ตัวแปร outcome, ตัวแปร exposure)* ทั้งนี้สามารถ help เพิ่มดูตัวอย่างโดยพิมพ์ ?cci
การวิเคราะห์ข้อมูล Mantel-Haenszel method		mhor(mhtable==ชื่อตารางความถี่,decomal=4, graph= TRUE, design = "case control")

8.5 แบบฝึกหัด

1. การศึกษาความสัมพันธ์ของการใช้ยาเม็ดคุมกำเนิด (Oral Contraceptive: OC) กับการเกิดภาวะกล้ามเนื้อหัวใจตาย (Myocardial Infarction: MI) ในผู้หญิงอายุ 40 ถึง 44 ปีจากการติดตามผลผู้ใช้งาน OC จำนวน 5000 รายและไม่ใช้ OC จำนวน 10000 เป็นระยะเวลา 3 ปี พบว่าผู้ใช้ OC เกิด MI จำนวน 13 ราย และผู้ไม่ใช้ OC เกิด MI จำนวน 7 รายจงทดสอบความสัมพันธ์ดังกล่าว[30]

2. การศึกษาความสัมพันธ์ของการรับประทานอาหารรสเค็มกับการเสียชีวิตจากการเกิดโรคหัวใจและหลอดเลือดโดยการศึกษาประวัติการรักษาของคนไข้ของผู้ป่วยที่เสียชีวิตอายุ 50 ถึง 54 ปี มีรูปแบบการศึกษาเป็นแบบ Retrospective study จากการเก็บข้อมูลผู้ที่เสียชีวิตซึ่งมีสาเหตุมาจากโรคหัวใจและหลอดเลือดจำนวน 35 คน พบว่ารับประทานอาหารเค็มจำนวน 5 คน ในขณะที่ผู้ที่เสียชีวิตซึ่งมีสาเหตุมาจากโรคอื่นจำนวน 25 คน พบว่ารับประทานอาหารเค็มจำนวน 2 คน จงทดสอบความสัมพันธ์ดังกล่าว[30]

3. การศึกษาความสัมพันธ์ของผู้ที่ได้รับควันบุหรี่มือสองกับการเป็นโรคมะเร็งปอด โดยที่คาดว่าการสูบบุหรี่มีความสัมพันธ์กับการเป็นโรคมะเร็งดังกล่าวด้วย ผลการศึกษาดังกล่าวได้ทำการแจกแจงความถี่ในตารางที่ 8.13 จากข้อมูลดังกล่าวจงทดสอบความสัมพันธ์ของการได้รับควันบุหรี่มือสองกับการเป็นโรคมะเร็งปอดและทดสอบปัจจัยทางด้านการสูบบุหรี่ต่อความสัมพันธ์ดังกล่าว

ตารางที่ 8.13 ความถี่ของการได้รับควันบุหรี่มือสองกับการเป็นโรคมะเร็งแยกชั้นภูมิตามตัวแปรการสูบบุหรี่

การสูบบุหรี่	การเป็นโรคมะเร็ง	การได้รับควันบุหรี่มือสอง	
สูบ	เป็น	120	111
	ไม่เป็น	80	155
	รวม	200	266
ไม่สูบ	เป็น	161	117
	ไม่เป็น	130	124
	รวม	291	241

8.6 เอกสารอ้างอิง

1. Lydersen S, Fagerland MW, Laake P. Recommended tests for association in 2 x 2 tables. Stat Med. 2009;28(7):1159-75.
2. Richardson JT. The analysis of 2 × 2 contingency tables--yet again. Stat Med. 2011;30(8):890; author reply 1-2.
3. Shan G, Gerstenberger S. Fisher's exact approach for post hoc analysis of a chi-squared test. PLoS One. 2017;12(12):e0188709.
4. Hess AS, Hess JR. Understanding tests of the association of categorical variables: the Pearson chi-square test and Fisher's exact test. Transfusion. 2017;57(4):877-9.
5. Szumilas M. Explaining odds ratios. J Can Acad Child Adolesc Psychiatry. 2010;19(3):227-9.
6. Cummings P. The relative merits of risk ratios and odds ratios. Arch Pediatr Adolesc Med. 2009;163(5):438-45.
7. Hoppe FM, Hoppe DJ, Walter SD. Explaining odds ratios as conditional risk ratios. J Clin Epidemiol. 2018;97:123-4.
8. Persoskie A, Ferrer RA. A Most Odd Ratio:: Interpreting and Describing Odds Ratios. Am J Prev Med. 2017;52(2):224-8.
9. Moser BK, Halabi S. Estimation and testing of the relative risk of disease in case-control studies with a set of k matched controls per case with known prevalence of disease. Stat Med. 2012;31(1):29-44.
10. Greenland S, Thomas DC, Morgenstern H. The rare-disease assumption revisited. A critique of "estimators of relative risk for case-control studies". Am J Epidemiol. 1986;124(6):869-83.
11. Hernán MA. The hazards of hazard ratios. Epidemiology. 2010;21(1):13-5.
12. Lin JC, Lee WC. Hazard ratio bias in cohort studies. Epidemiology. 2013;24(5):777-8.
13. Spruance SL, Reid JE, Grace M, Samore M. Hazard ratio in clinical trials. Antimicrob Agents Chemother. 2004;48(8):2787-92.
14. Sedgwick P, Joekes K. Interpreting hazard ratios. Bmj. 2015;351:h4631.
15. Cummings P. The relative merits of risk ratios and odds ratios. Arch Pediatr Adolesc Med. 2009;163(5):438-45.

16. Mayer DM, Young A. When to use relative risk or odds ratios in describing study results. Ann Emerg Med. 2005;46(4):385-6; author reply 6.
17. Brumback B, Berg A. On effect-measure modification: Relationships among changes in the relative risk, odds ratio, and risk difference. Stat Med. 2008;27(18):3453-65.
18. Laopaiboon M, Lumbiganon P, Intarut N, Mori R, Ganchimeg T, Vogel JP, et al. Advanced maternal age and pregnancy outcomes: a multicountry assessment. BJOG. 2014;121 Suppl 1:49-56.
19. Chongsuvivatwong V. epiDisplay: Epidemiological Data Display Package. 2018.
20. LaVeist TA. Beyond dummy variables and sample selection: what health services researchers ought to know about race as a variable. Health Serv Res. 1994;29(1):1-16.
21. Sperandei S. Understanding logistic regression analysis. Biochem Med (Zagreb). 2014;24(1):12-8.
22. Kim HY. Statistical notes for clinical researchers: Chi-squared test and Fisher's exact test. Restor Dent Endod. 2017;42(2):152-5.
23. Lydersen S, Fagerland MW, Laake P. Recommended tests for association in 2 x 2 tables. Stat Med. 2009;28(7):1159-75.
24. Connelly LM. Fisher's Exact Test. Medsurg Nurs. 2016;25(1):58, 61.
25. Agresti A. An Introduction to Categorical Data Analysis, Second Edition. 2007.
26. Berger VW, Stefanescu C, Zhou YY. The analysis of stratified 2 x 2 contingency tables. Biom J. 2006;48(6):992-1007.
27. Lu K. Multiple imputation score tests and an application to Cochran-Mantel-Haenszel statistics. Stat Med. 2020;39(27):4025-36.
28. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. J Natl Cancer Inst. 1959;22(4):719-48.
29. Chisha Y, Terefe W, Assefa H, Lakew S. Prevalence and factors associated with diabetic retinopathy among diabetic patients at Arbaminch General Hospital, Ethiopia: Cross sectional study. PLoS One. 2017;12(3):e0171987.
30. Rosner B. Fundamentals of biostatistics: Seventh edition. Boston : Brooks/Cole, Cengage Learning; 2011.

บทที่ 9

การวิเคราะห์สหสัมพันธ์ของเพียร์สัน และการถดถอยเชิงเส้น

จากบทที่ 7 ได้ศึกษาถึงการวิเคราะห์ข้อมูลที่มีผลลัพธ์ทางการศึกษาเป็นแบบข้อมูลต่อเนื่องและในบทที่ 8 ได้ศึกษาการวิเคราะห์ข้อมูลตัวแปรจำแนกประเภทซึ่งเป็นการวิเคราะห์แบบสองตัวแปรและการวิเคราะห์วิธีการของ Mantel-Haenszel ในบทนี้จะได้ศึกษาการทดสอบความสัมพันธ์ของตัวแปรต่อเนื่องจำนวนสองตัวแปรด้วยสหสัมพันธ์ของเพียร์สัน (Person's correlation) และการถดถอยเชิงเส้น (Linear regression)

9.1 การทดสอบสหสัมพันธ์ของเพียร์สัน

ในการวิเคราะห์ข้อมูลสองตัวแปรที่มีการวัดเป็นแบบตัวแปรต่อเนื่องทั้งคู่ จะใช้สถิติที่เรียกว่า สหสัมพันธ์ของเพียร์สัน (Person's correlation)[1-5] โดยมีสูตรการคำนวณดังนี้

$$r = \frac{\sum_{i=1}^{n}(x - \bar{x})(y - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x - x)^2 \sum_{i=1}^{n}(y - \bar{y})^2}}$$

โดยที่

x_i หมายถึงข้อมูลของตัวแปรที่หนึ่ง, $\bar{x}$ หมายถึงค่าเฉลี่ยข้อมูลของตัวแปรที่หนึ่ง

y_i หมายถึงข้อมูลของตัวแปรที่สอง, $\bar{y}$ หมายถึงค่าเฉลี่ยข้อมูลของตัวแปรที่สอง

ข้อตกลงเบื้องต้นของการใช้สถิติ[6]

1. ตัวอย่างสุ่มมาอย่างอิสระ
2. ประชากรมีการแจงแจกปกติ

ค่าสหสัมพันธ์ของเพียร์สัน (r) จะบอกระดับความสัมพันธ์ของสองแปรโดยค่าที่คำนวณได้จะมีค่าตั้งแต่ -1 ถึง 1 ดังแสดงความหมายไว้ในตารางที่ 9.1

ตารางที่ 9.1 ความหมายของขนาดความสัมพันธ์ของค่าสัมพัทธ์

ขนาดของความสัมพันธ์	ความหมาย
0.91, 1.00 (–0.91, –1.00)	ความสัมพันธ์เชิงบวก (ลบ) สูงมาก
0.71, 0.90 (–0.71, –0.90)	ความสัมพันธ์เชิงบวก (ลบ) สูง
0.51, 0.70 (–0.51, –0.70)	ความสัมพันธ์เชิงบวก (ลบ) ปานกลาง
0.31, 0.50 (–0.31, –0.50)	ความสัมพันธ์เชิงบวก (ลบ) ต่ำ
0.00, 0.30 (0.00, –0.30)	ความสัมพันธ์ที่ไม่มีสำคัญ (negligible correlation)

ที่มา: ดัดแปลงมาจาก Mukaka MM. Statistics Corner: A guide to appropriate use of Correlation coefficient in medical research. Malawi Med J. 2012;24(3):69-71.Table 1.Rule of Thumb for Interpreting the Size of a Correlation Coefficient[7]

ตัวอย่างที่ 9.1 การทดสอบความสัมพันธ์ระหว่างคะแนนกิจกรรมทางกาย (Physical Activity) กับดัชนีมวลกาย (Body Mass Index: BMI) โดยมีข้อมูลดังตารางที่ 9.2

ตารางที่ 9.2 ตารางแสดงข้อมูลคะแนนกิจกรรมทางกายและดัชนีมวลกาย

id	PA (X)	BMI (Y)
1	65	30
2	74	22
3	56	34
4	93	21
5	51	30
6	87	24
7	86	26
8	73	24
9	95	22
10	81	21
11	71	30
12	84	21
13	84	24
14	68	25
15	59	29

วิธีทำ

ขั้นตอนที่ 1 ทดสอบการแจกแจงของข้อมูล

```
> id<-c(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15)
> pa<-c(65,74,56,93,51,87,86,73,95,81,71,84,84,68,59)
> bmi<-c(30,22,34,21,30,24,26,24,22,21,30,21,24,25,29)
> dt<-data.frame(id,pa,bmi)
> #test normality
> shapiro.test(dt$pa)

        Shapiro-Wilk normality test

data:  dt$pa
W = 0.96023, p-value = 0.6965

> shapiro.test(dt$bmi)

        Shapiro-Wilk normality test

data:  dt$bmi
W = 0.89478, p-value = 0.07924
```

รูปที่ 9.1 ผลการทดสอบการแจกแจงข้อมูล

จากการทดสอบการแจกแจงของข้อมูลพบว่าคะแนนกิจกรรมทางกาย (p-value = 0.697) กับดัชนีมวลกาย (p-value = 0.079) มีการแจกแจงแบบปกติดังนั้นจึงทดสอบความสัมพันธ์โดยใช้ Person's correlation

ขั้นตอนที่ 2 การทดสอบทางสถิติ

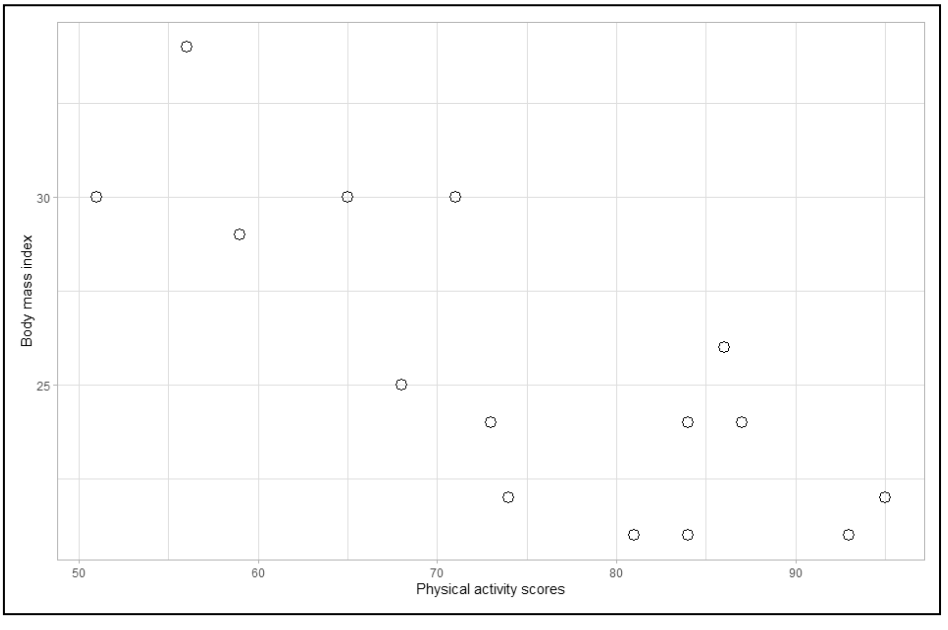
```
> cor.test(pa, bmi, method=c("pearson"))

        Pearson's product-moment correlation

data:  pa and bmi
t = -4.8121, df = 13, p-value = 0.0003394
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.9309121 -0.4881269
sample estimates:
       cor
-0.8002802
```

รูปที่ 9.2 ผลการวิเคราะห์ความสัมพันธ์ของคะแนนกิจกรรมทางกายกับดัชนีมวลกาย

ผลการวิเคราะห์พบว่าคะแนนกิจกรรมทางกายกับดัชนีมวลกายมีความสัมพันธ์ระดับสูงแปรผกผันกัน (r = -0.8, 95% CI: -0.9, -0.5) อย่างมีนัยสำคัญทางสถิติ นอกจากนั้นแล้วยังสามารถแสดงความสัมพันธ์ดังกล่าวด้วยกราฟโดยใช้ชุดคำสั่งของ ggplot2[8] ดังรูปที่ 9.3

```
library(ggplot2)
g0<-ggplot(dt, aes(x=pa, y=bmi)) +
  geom_point(shape = 21,size=4)
g1 <- g0+labs(x = "Physical activity scores", y = "Body mass index")
g1 + theme_light()
```

รูปที่ 9.3 คำสั่งวาดกราฟจากชุดคำสั่ง ggplot2



รูปที่ 9.4 กราฟความสัมพันธ์ระหว่างคะแนนกิจกรรมทางกายและดัชนีมวลกาย

จากรูปที่ 9.4 จะเห็นว่าคะแนนกิจกรรมทางกายกับดัชนีมวลกายมีความสัมพันธ์ต่อกันทั้งนี้เป็นความสัมพันธ์ในเชิงลบ

9.2 การใช้สถิติสหสัมพันธ์ของสเปียร์แมน

จากการทดสอบโดยใช้สหสัมพันธ์ของเพียร์สันข้อมูลจะต้องมีการแจกแจงแบบปกติทั้งสองตัวแปร อย่างไรก็ตามอาจจะมีบางตัวแปรหรือทั้งสองตัวแปร มีการแจกแจงไม่ปกติ การทดสอบแบบสหสัมพันธ์ของเพียร์สันจะไม่เหมาะสม กรณีดังกล่าวจะใช้สถิตินอนพาราเมตริกที่ชื่อว่า Spearman's rank correlation[9-12]

ตัวอย่างที่ 9.2 การทดสอบความสัมพันธ์ระหว่างคะแนน Apgar score ในเด็กแรกเกิดจำนวน 13 คน โดยการวัด Apgar score หลังจากคลอด 1 นาทีและ 5 นาที ได้ข้อมูลดังนี้[13]

ตารางที่ 9.3 ข้อมูล Apgar score ที่ 1 นาทีและ 5 นาที

apgar1min	apgar 5 min
10	10
3	6
8	9
9	10
8	9
9	10
8	9
8	9
8	9
8	9
7	9
8	9
7	9

วิธีทำ

ขั้นตอนที่ 1 ทดสอบการแจกแจงของข้อมูล

```
> apgar1min<-c(10,3,8,9,8,9,8,8,8,8,7,8,7)
> apgar5min<-c(10,6,9,10,9,10,9,9,9,9,9,9,9)
> dt<-data.frame(apgar1min,apgar5min)
> #test normality
> shapiro.test(dt$apgar1min)

        Shapiro-Wilk normality test

data:  dt$apgar1min
W = 0.74244, p-value = 0.001527

> shapiro.test(dt$apgar5min)

        Shapiro-Wilk normality test

data:  dt$apgar5min
W = 0.61504, p-value = 8.857e-05
```

รูปที่ 9.5 ผลการทดสอบการแจกแจงข้อมูลของ Apgar score

การทดสอบการแจกแจงของคะแนน Apgar Score ที่ 1 นาที และ 5 นาที พบว่ามีการแจกแจงไม่ปกติทั้งสองตัวแปร ดังนั้นในการทดสอบความสัมพันธ์จะใช้สถิติ Spearman's rank correlation

ขั้นตอนที่ 2 ทดสอบทางสถิติ

```
> cor.test(dt$apgar1min,dt$apgar5min, method=c("spearman"))

        Spearman's rank correlation rho

data:  dt$apgar1min and dt$apgar5min
S = 41.636, p-value = 5.641e-05
alternative hypothesis: true rho is not equal to 0
sample estimates:
      rho
0.8856149
```

รูปที่ 9.6 ผลการวิเคราะห์ข้อมูลของ Apgar score ที่ 1 นาทีและ 5 นาที

การทดสอบทางสถิติพบว่าคะแนน Apgar ที่ 1 นาที หลังคลอดมีความสัมพันธ์กับคะแนน Apgar ที่ 5 นาที อย่างมีนัยสำคัญทางสถิติ (p-value < 0.001) โดยที่มีค่าระดับของความสัมพันธ์เท่ากับ 0.89 อยู่ในระดับสูง

9.3 สมการถดถอยเชิงเส้นอย่างง่าย

การวิเคราะห์ความสัมพันธ์ทำให้ทราบว่าสองตัวแปรที่ทดสอบมีความสัมพันธ์กันหรือไม่ และทราบระดับความสัมพันธ์ (ค่า r) แต่อย่างไรก็ตามนักวิจัยอยากทราบต่อว่าถ้าค่าของตัวแปรต้นเปลี่ยนแปลงไปจำนวนหนึ่งแล้ว ค่าของตัวแปรตามจะเปลี่ยนแปลงไปเท่าไหร่ การหาคำตอบดังกล่าวจะใช้การวิเคราะห์สมการถดถอยเชิงเส้น การวิเคราะห์โดยใช้สมการถดถอยเชิงเส้นที่มีตัวแปรต้น 1 ตัวแปรและตัวแปรตาม 1 ตัวแปร จะเรียกว่า การวิเคราะห์การถดถอยอย่างง่าย (Simple linear regression)[14-17] ถ้าตัวแปรต้นมากกว่า 1 ตัวแปรและตัวแปรตาม 1 ตัวแปร เรียกว่า การวิเคราะห์สมการถดถอยเชิงเส้นพหุคูณ (Multiple linear regression) ทั้งนี้สามารถเขียนเป็นสมการเชิงเส้นอย่างง่ายได้ดังนี้

$$y = \alpha + \beta \times x + \varepsilon$$

โดยที่

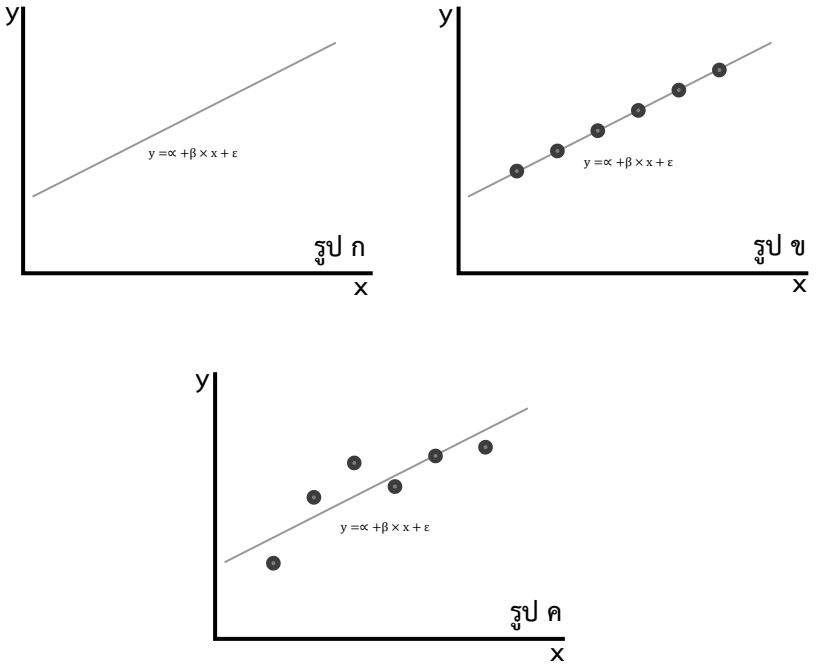
y หมายถึง ตัวแปรตามหรือตัวแปรไม่อิสระหรือ Dependent variable หรือ Outcome

x หมายถึง ตัวแปรต้นหรือตัวแปรอิสระหรือ Independent variable หรือ Factors

α หมายถึง จุดตัดแกน y

β หมายถึง ค่าสัมประสิทธิ์ของสมการความสัมพันธ์เชิงเส้นตรง หรือ ความชันของกราฟเส้นตรง

จากสมการเชิงเส้นสามารถวาดเป็นกราฟได้ดังนี้



รูปที่ 9.7 กราฟสมการเส้นตรงอย่างง่าย
ที่มา: ดัดแปลงมาจาก Rosner B. Fundamentals of biostatistics: Seventh edition. Boston: Brooks/Cole, Cengage Learning, 2011. page 428.

รูป ก สมการเชิงเส้นระหว่างความสัมพันธ์ของตัวแปร x กับตัวแปร y

รูป ข สมการเชิงเส้นที่ลากผ่านข้อมูลอย่างสมบูรณ์แบบผ่านจุดทุกจุกของข้อมูล ในกรณีนี้ค่าความแปรปรวนของข้อมูล (Regression line) มีค่าเท่ากับศูนย์

รูป ค สมการเชิงเส้นที่ลากผ่านข้อมูล ที่คาดว่าเข้าใกล้ข้อมูลทุกจุดอย่างมากที่สุด ในกรณีนี้ค่าความแปรปรวนของข้อมูล (Regression line) มีค่ามากกว่าศูนย์ทั้งนี้สมการที่ลากผ่านกลุ่มข้อมูลที่ใกล้เคียงที่สุดใช้วิธีการที่เรียกว่า Least square mean

ตัวอย่างที่ 9.3 การทดสอบความสัมพันธ์คะแนนกิจกรรมทางกายกับดัชนีมวลกายโดยใช้ข้อมูลในตารางที่ 9.2 โดยใช้วิธี Simple linear regression

```
> id<-c(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15)
> pa<-c(65,74,56,93,51,87,86,73,95,81,71,84,84,68,59)
> bmi<-c(30,22,34,21,30,24,26,24,22,21,30,21,24,25,29)
> dt<-data.frame(id,pa,bmi)
> m0<-lm(bmi~pa,data=dt)
> summary(m0)

Call:
lm(formula = bmi ~ pa, data = dt)

Residuals:
    Min      1Q  Median      3Q     Max
-3.8103 -2.1656 -0.1673  1.6785  3.7911

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 43.89332    3.87241  11.335 4.13e-08 ***
pa          -0.24437    0.05078  -4.812 0.000339 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.564 on 13 degrees of freedom
Multiple R-squared:  0.6404,  Adjusted R-squared:  0.6128
F-statistic: 23.16 on 1 and 13 DF,  p-value: 0.0003394
```

รูปที่ 9.8 ผลการวิเคราะห์ข้อมูลโดยใช้ Simple linear regression

ผลการวิเคราะห์ข้อมูลในรูปที่ 9.8 พบว่าเมื่อคะแนนกิจกรรมทางกายเพิ่มขึ้น 1 หน่วย ดัชนีมวลกายจะลดลง -0.24 (p-value < 0.001) โดยที่มีค่า R-squared เท่ากับ 0.64 หมายถึง สมการเชิงเส้น (Regression line) เข้าได้ดีกับข้อมูล 64% ของข้อมูลชุดนี้

9.4 การวิเคราะห์การถดถอยพหุคูณ

ในการวิจัยเพื่อทดสอบความสัมพันธ์ของตัวแปรอิสระ(ตัวแปรต้น) ซึ่งมีหลายตัวแปร ซึ่งคาดว่าน่ามีความสัมพันธ์ต่อตัวแปรไม่อิสระ (ตัวแปรตาม) ในกรณีนี้จะใช้การวิเคราะห์แบบ Multiple linear regression)[18, 19] โดยที่สามารถเขียนสมการถดถอยพหุคูณได้ดังนี้

$$y = \beta_0 + \beta_1 \times x_1 + \beta_i \times x_i + \varepsilon_i$$

โดยที่

y = ผลลัพธ์ที่ตัวการศึกษา หรือ ตัวแปรตาม หรือ Outcome มีความหมายเหมือนกัน

x = ตัวแปรอิสระ หรือ Factors หรือ ตัวแปรต้น มีความหมายเหมือนกัน

β_0 = จุดตัดแกน Y

β = ค่า Coefficients

ε_i = Residual errors

ขั้นตอนการฟิตโมเดล (Fitting model)[20-22] จะประกอบไปด้วย 2 ขั้นตอนหลักๆ คือ

1. การฟิตโมเดล (Fitting model) เทคนิคการฟิตโมเดลประกอบไปด้วย 3 เทคนิคหลักๆ คือ

1.1 Backward elimination

เทคนิคนี้เริ่มต้นด้วยตัวแปรที่ได้จากการวิเคราะห์แบบสองตัวแปร (Univariable analysis) ที่มีค่า p-value น้อยกว่า 0.05 (หรือ p-value = 0.15-0.20 ขึ้นอยู่กับขนาดตัวอย่างของงานวิจัย[23, 24]) และตัวแปรกวน (Known confounder) เข้าสู่โมเดลเริ่มต้นจากนั้นทำการฟิตโมเดลโดยการเอาตัวแปรที่ไม่มีนัยสำคัญทางสถิติออก จนได้โมเดลสุดท้าย (ทุกตัวแปรในโมเดล มีค่าความน่าจะเป็น น้อยกว่า 0.05) ในตำราเล่มนี้ใช้วิธีการฟิตโมเดลแบบ Backward elimination

1.2 Forward Elimination

การฟิตโมเดลทำตรงกันข้ามกับวิธี Backward elimination โดยทำการเอาตัวแปรเข้าในโมเดลทีละตัวจนได้โมเดลสุดท้าย

1.3 Stepwise

การฟิตโมเดลโดยวิธีนี้เป็นการนำตัวแปรทั้งหมดเข้าโมเดลแล้วทำการกำหนดเงื่อนไขในการนำตัวแปรเข้าโมเดล แล้วกำหนดเกณฑ์ในการดึงออกของตัวแปรที่ไม่มีความสัมพันธ์กับโมเดลนี้ วิธีนี้แตกต่างกับสองวิธีด้านบนคือ การคัดเลือกตัวแปรเข้าโมเดล วิธี Backward และ Forward elimination มีการทดสอบสองตัวแปรเพื่อคัดเลือกตัวแปรอิสระเข้าสู่โมเดล

2. การทดสอบข้อตกลงเบื้องต้นของโมเดล

ข้อตกลงเบื้องต้นของสมการถดถอยเชิงเส้นดังนี้

2.1 Linearity

ความสัมพันธ์ของตัวแปรต้น (Independent factors) และตัวแปรตาม (Dependent factors) ต้องมีลักษณะเป็นเชิงเส้น

2.2 Residual errors มีการแจกแจงปกติ

2.3 Homogeneity of residuals variance คือตัวแปรอิสระมีความแปรปรวนที่เท่ากัน

การตรวจสอบข้อตกลงเบื้องต้นดังกล่าวสามารถใช้กราฟตรวจสอบได้ดังนี้

1. Residuals vs Fitted	ใช้ตรวจสอบความเป็นเชิงเส้นของตัวแปรต้นกับตัวแปรตาม (ควรเป็นเส้นตรงและเส้นตามขวาง)
2. Normal Q-Q	ใช้ตรวจสอบการแจกแจงของค่า residual (ข้อมูลควรอยู่ใกล้เส้นทะแยงมุมมากที่สุด)
3. Scale-Location	ใช้ตรวจสอบความเท่ากันของความแปรปรวน (จุดข้อมูลควรกระจายเส้นตามขวางควรเป็นเส้นตรง)
4. Residuals vs Leverage	ใช้ตรวจสอบค่าที่ผิดปกติไปจากกลุ่มของข้อมูลนั้นหรือเรียกว่า Outlier

ตัวอย่างที่ 9.4การวิเคราะห์ Multiple linear regression

การศึกษาเพื่อทดสอบความสัมพันธ์ของระดับน้ำตาลในเลือด (Fasting plasma glucose) และระดับกรดยูริก (Serum uric acid) โดยมีรูปแบบศึกษาเป็นแบบภาคตัดขวาง (Cross-sectional study) ข้อมูลเป็นการจำลองดังแสดงในตารางที่ 9.4

ตารางที่ 9.4 ตารางข้อมูลระดับน้ำตาลในเลือดและระดับกรดยูริกและตัวแปรกวน

id	uric	fpg	age	bmi	whr	gender	smk
1	286	5.3	56	25.42	0.92	1	1
2	289	5.2	43	26.2	0.97	1	1
3	288	5.3	58	25.72	0.94	0	1
4	276	5.6	57	24.65	0.98	0	1
5	276	6.5	57	22.74	0.99	1	0
6	259	6.3	44	23.28	0.81	0	0
7	290	5.1	57	26.55	0.99	0	1
8	271	6.6	54	21.92	0.88	1	0
9	288	5.2	50	26.05	0.97	0	1
10	275	6.6	55	22.2	0.96	1	1
11	260	6.6	45	22.46	0.82	1	0
12	290	5	57	26.62	1	1	0
13	272	5.7	54	24.55	0.92	0	1
14	289	5.2	46	25.92	0.94	0	1
15	285	5.3	57	25.33	0.89	1	1
16	279	5.6	58	24.83	0.86	0	1
17	262	6	49	23.8	0.84	1	0
18	282	5.4	48	25.04	0.87	0	1
19	275	5.6	56	24.63	0.96	0	1
20	267	5.9	49	24.08	0.84	0	1
21	269	5.9	49	24.13	0.85	0	0
22	272	5.6	55	24.58	0.95	0	1
23	257	6.4	41	23.24	0.81	0	0
24	261	6.1	46	23.68	0.82	1	1
25	269	6.5	50	22.59	0.85	0	0
26	284	5.4	47	25.21	0.89	0	1
27	277	6.5	46	22.96	1	0	0
28	280	5.6	50	24.97	0.87	0	1
29	270	5.7	51	24.47	0.86	0	0
30	255	6.4	40	23.1	0.8	1	0

ตารางที่ 9.5 ดัชนีข้อมูลของการทดสอบความสัมพันธ์ของระดับน้ำตาลในเลือดและระดับกรดยูริก

ชื่อตัวแปร	**ความหมาย**	**ค่า**
id	ระดับที่ของข้อมูล	ตัวแปรต่อเนื่อง
uric	ระดับกรดยูริก (μmol/L)	ตัวแปรต่อเนื่อง
fpg	ระดับน้ำตาลในเลือด (mmol/L)	ตัวแปรต่อเนื่อง
age	อายุ (ปี)	ตัวแปรต่อเนื่อง
bmi	ดัชนีมวลกาย (กิโลกรัมต่อตารางเมตร)	ตัวแปรต่อเนื่อง
whr	สัดส่วนรอบเอวต่อรอบสะโพก (เซนติเมตร/เซ็นติเมตร)	ตัวแปรต่อเนื่อง
gender	เพศ	0 = ผู้หญิง, 1 =ผู้ชาย
smk	การสูบบุหรี่	0 = ไม่สูบ, 1 =สูบ

การวิเคราะห์ข้อมูล

ขั้นตอนที่ 1 สำรวจข้อมูลและจัดการข้อมูล

```
>setwd("D:\\epibios")
> library(epiDisplay)
> dt<-read.csv("data_linear.csv")
> #exploratory
>summ(dt$uric)
 obs. mean   median  s.d.   min.   max.
 30   275.1  275.5   10.736 255    290
> shapiro.test(dt$uric)

        Shapiro-Wilk normality test

data:  dt$uric
W = 0.9443, p-value = 0.1188

>summ(dt$fpg)
 obs. mean   median  s.d.   min.   max.
 30   5.803  5.65    0.526  5      6.6
>shapiro.test(dt$fpg)

        Shapiro-Wilk normality test

data:  dt$fpg
W = 0.90889, p-value = 0.01395

>summ(dt$bmi)
 obs. mean   median  s.d.   min.   max.
 30   24.364 24.565  1.331  21.92  26.62
> shapiro.test(dt$bmi)

        Shapiro-Wilk normality test

data:  dt$bmi
W = 0.96984, p-value = 0.5349

>summ(dt$whr)
 obs. mean   median  s.d.   min.   max.
 30   0.902  0.89    0.065  0.8    1
> shapiro.test(dt$whr)

        Shapiro-Wilk normality test

data:  dt$whr
W = 0.92758, p-value = 0.04235
```

รูปที่ 9.9 ผลการสำรวจข้อมูลและจัดการข้อมูล

ผลการสำรวจข้อมูลในรูปที่ 9.9 สรุปได้ว่าระดับกรดยูริก (p-value = 0.119) และดัชนีมวลกาย (p-value = 0.535) มีการแจกแจงปกติ ส่วนระดับน้ำตาลในเลือด (p-value = 0.014) และสัดส่วนรอบเอวต่อรอบสะโพกมีการแจกแจงไม่ปกติ (p-value = 0.042)

การสำรวจและจัดการข้อมูลตัวแปรเพศ (gender)

```
> tab1(dt$gender)
dt$gender :
        Frequency Percent Cum. percent
0              19    63.3         63.3
1              11    36.7        100.0
  Total        30   100.0        100.0
>dt$gender<-factor(dt$gender,levels = c("0","1"),labels =
c("female","male"))
> tab1(dt$gender)
dt$gender :
        Frequency Percent Cum. percent
female         19    63.3         63.3
male           11    36.7        100.0
  Total        30   100.0        100.0
>shapiro.test(dt$uric[dt$gender=="female"])

        Shapiro-Wilk normality test

data:  dt$uric[dt$gender == "female"]
W = 0.9577, p-value = 0.528

>shapiro.test(dt$uric[dt$gender=="male"])

        Shapiro-Wilk normality test

data:  dt$uric[dt$gender == "male"]
W = 0.9106, p-value = 0.2479
```

รูปที่ 9.10 ผลการสำรวจและจัดการข้อมูลตัวแปรเพศ

การทดสอบการแจกแจงข้อมูลของระดับกรดยูริกแยกตามเพศพบว่ามีการแจกแจงปกติทั้งกลุ่มเพศชาย (p-value = 0.248) และเพศหญิง (p-value = 0.528)

การสำรวจและจัดการข้อมูลตัวแปรการสูบบุหรี่ (smk)

```
>tab1(dt$smk)
dt$smk :
        Frequency Percent Cum. percent
0              12      40           40
1              18      60          100
  Total        30     100          100
> dt$smk<-factor(dt$smk,levels = c("0","1"),labels = c("no","yes"))
> tab1(dt$smk)
dt$smk :
        Frequency Percent Cum. percent
no             12      40           40
yes            18      60          100
  Total        30     100          100
>shapiro.test(dt$uric[dt$smk=="no"])

        Shapiro-Wilk normality test

data:  dt$uric[dt$smk == "no"]
W = 0.93493, p-value = 0.4353

>shapiro.test(dt$uric[dt$smk=="yes"])

        Shapiro-Wilk normality test

data:  dt$uric[dt$smk == "yes"]
W = 0.92861, p-value = 0.1836
```

รูปที่ 9.11 ผลการสำรวจและจัดการข้อมูลตัวแปรการสูบบุหรี่

การทดสอบการแจกแจงข้อมูลของระดับกรดยูริกในแยกตามการสูบบุหรี่พบว่ามีการแจกแจงปกติทั้งกลุ่มที่สูบบุหรี่ (p-value = 0.184) และไม่สูบบุหรี่ (p-value = 0.435)

การสำรวจและจัดการข้อมูลตัวแปรอายุ (age)

```
>summ(dt$age)
 obs. mean   median  s.d.   min.   max.
 30   50.833 50      5.453  40     58
>dt$age_gr<-cut(dt$age, breaks=c(min(dt$age,na.rm =
TRUE),45,55,max(dt$age,na.rm = TRUE)),
+                labels = c("<46","46-55",">55yrs"),
+                include.lowest = TRUE)
>tab1(dt$age_gr)
dt$age_gr :
        Frequency Percent Cum. percent
<46             5    16.7         16.7
46-55          16    53.3         70.0
>55yrs          9    30.0        100.0
  Total        30   100.0        100.0
>shapiro.test(dt$uric[dt$age_gr=="<46"])

        Shapiro-Wilk normality test

data:  dt$uric[dt$age_gr == "<46"]
W = 0.6833, p-value = 0.006353

>shapiro.test(dt$uric[dt$age_gr=="46-55"])

        Shapiro-Wilk normality test

data:  dt$uric[dt$age_gr == "46-55"]
W = 0.95571, p-value = 0.5851

>shapiro.test(dt$uric[dt$age_gr==">55yrs"])

        Shapiro-Wilk normality test

data:  dt$uric[dt$age_gr == ">55yrs"]
W = 0.8592, p-value = 0.09399
```

รูปที่ 9.12 ผลการสำรวจและจัดการข้อมูลตัวแปรอายุ

การทดสอบการแจกแจงข้อมูลของตัวแปรอายุ ซึ่งทำการแบ่งเป็นกลุ่มอายุเป็น 3 กลุ่ม ดังนี้คือ อายุ น้อยกว่าหรือเท่ากับ 46 ปี, อายุ 47 ถึง 55 ปี และมากกว่าหรือเท่ากับ 56 ปี โดยการใช้คำสั่ง *cut()* จากนั้นทำการทดสอบการแจกแจงของระดับกรดยูริกแยกตามกลุ่มอายุโดยที่พบว่ามีการแจกแจงของข้อมูลปกติในกลุ่มอายุ ระหว่าง 47 ถึง 55 ปี (p-value = 0.585) และ อายุมากกว่าหรือเท่ากับ 56 ปี (p-value = 0.094) ส่วนกลุ่มอายุน้อยกว่าหรือเท่ากับ 46 ปี มีการแจกแจงไม่ปกติ (p-value = 0.006)

ขั้นตอนที่ 2 การวิเคราะห์สองตัวแปร

```
>cor.test(dt$uric,dt$bmi)

        Pearson's product-moment correlation

data:  dt$uric and dt$bmi
t = 6.3658, df = 28, p-value = 6.878e-07
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.5653859 0.8842890
sample estimates:
      cor
0.7690138

>cor.test(dt$uric,dt$fpg,method = c("spearman"))

        Spearman's rank correlation rho

data:  dt$uric and dt$fpg
S = 8145.1, p-value = 5.117e-08
alternative hypothesis: true rho is not equal to 0
sample estimates:
       rho
-0.8120373

Warning message:
In cor.test.default(dt$uric, dt$fpg, method = c("spearman")) :
  Cannot compute exact p-value with ties
>cor.test(dt$uric,dt$whr,method = c("spearman"))

        Spearman's rank correlation rho

data:  dt$uric and dt$whr
S = 1037.3, p-value = 6.798e-07
alternative hypothesis: true rho is not equal to 0
sample estimates:
      rho
0.7692309

Warning message:
In cor.test.default(dt$uric, dt$whr, method = c("spearman")) :
  Cannot compute exact p-value with ties
>
```

รูปที่ 9.13 ผลการวิเคราะห์ข้อมูลแบบสองตัวแปร

จากผลการทดสอบในรูปที่ 9.13 พบว่าระดับกรดยูริกกับปัจจัยด้านดัชนีมวลกาย (p-value = 6.878e-07 หรือ p-value < 0.001) ระดับน้ำตาลในเลือด (p-value = 5.117e-08 หรือ p-value < 0.001) และ สัดส่วนรอบเอวต่อรอบสะโพก (p-value = 6.798e-07 หรือ p-value < 0.001) ซึ่งพบว่ามีความสัมพันธ์กับระดับกรดยูริกทั้ง 3 ตัวแปรจึงจะคัดตัวแปรทั้ง 3 เข้าไปวิเคราะห์ในขั้นตอนการวิเคราะห์หลายตัวแปร (Multiple linear regression)

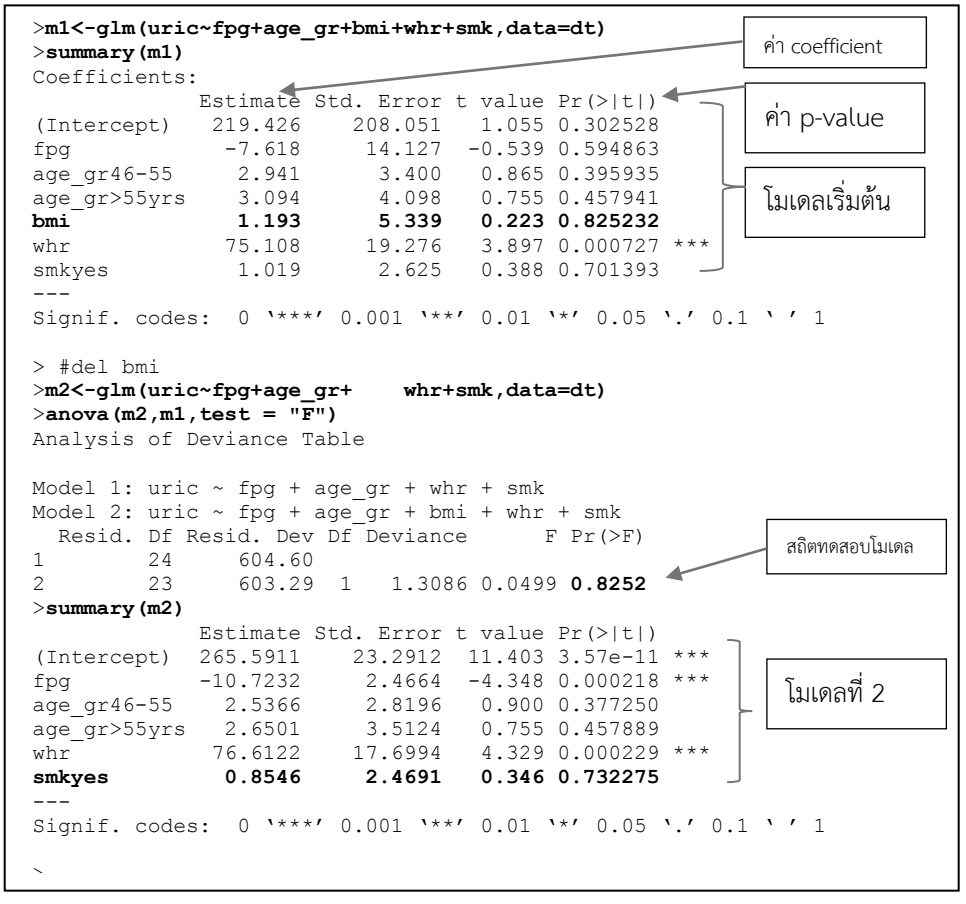
```
>kruskal.test(uric ~ age_gr, data = dt)

        Kruskal-Wallis rank sum test

data:  uric by age_gr
Kruskal-Wallis chi-squared = 9.326, df = 2, p-value = 0.009438

>t.test(dt$uric~dt$gender)

        Welch Two Sample t-test

data:  dt$uric by dt$gender
t = 0.52057, df = 16.703, p-value = 0.6095
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -7.068024 11.690034
sample estimates:
mean in group female   mean in group male
            275.9474             273.6364

>t.test(dt$uric~dt$smk)

        Welch Two Sample t-test

data:  dt$uric by dt$smk
t = -3.4062, df = 20.67, p-value = 0.002704
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -19.288844  -4.655601
sample estimates:
 mean in group no mean in group yes
         267.9167          279.8889
```

รูปที่ 9.14 ผลการวิเคราะห์ความแตกต่างของค่าเฉลี่ยระหว่างสองตัวแปร

จากการทดสอบความแตกต่างกันของระดับกรดยูริกระหว่างเพศชายและเพศหญิงไม่มีความแตกต่างกัน (p-value = 0.609) ระหว่างผู้สูบและไม่สูบบุหรี่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ (p-value = 0.003) และ ระหว่างกลุ่มอายุมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ (p-value = 0.009) ดังนั้นจะคัดเลือกตัวแปร กลุ่มอายุและการสูบบุหรี่ เข้าไปสู่การวิเคราะห์แบบหลายตัวแปร

ขั้นตอนที่ 3 การวิเคราะห์หลายตัวแปร (Multiple linear regression)

จากขั้นตอนการวิเคราะห์แบบสองตัวแปรพบว่าตัวแปรที่มีค่า p-value น้อยกว่า 0.05 ประกอบไปด้วย 5 ตัวแปรคือ ดัชนีมวลกาย ระดับน้ำตาลในเลือด สัดส่วนรอบเอวต่อรอบสะโพก การสูบบุหรี่ และกลุ่มอายุ ดังนั้นในขั้นการวิเคราะห์แบบหลายตัวแปรจะนำ 5 ตัวแปรนี้มาวิเคราะห์



รูปที่ 9.15 ผลการวิเคราะห์ข้อมูล Multiple linear regression โมเดลเริ่มต้นและโมเดลที่ 2

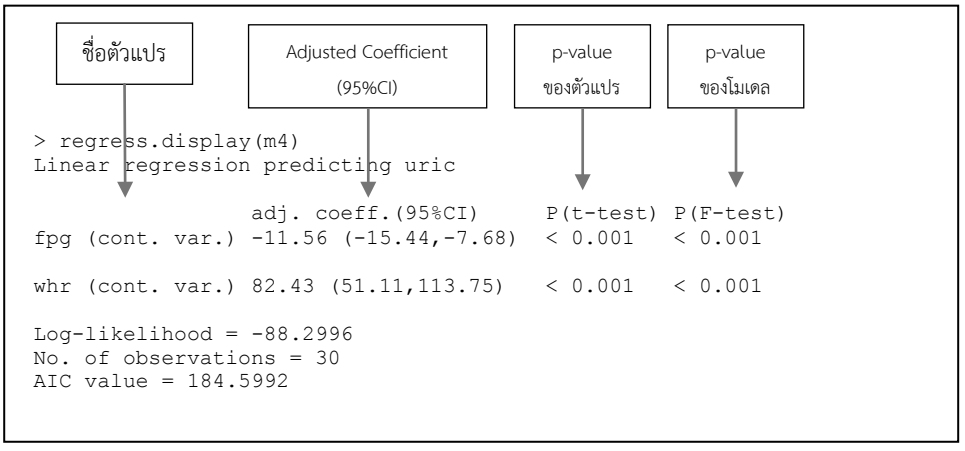
การวิเคราะห์ Multiple linear regression จะใช้คำสั่ง glm() จากการวิเคราะห์โมเดลเริ่มต้นพบว่า ค่า p-value ของตัวแปร bmi (ดัชนีมวลกาย) มีค่ามากที่สุด คือ 0.825 ดังนั้นจึงฟิตโมเดลที่ 2 โดยไม่มีตัวแปร bmi และทำการทดสอบ ANOVA โดยใช้คำสั่ง *anova(m2,m1, test = "F")* เพื่อทดสอบว่าตัวแปรที่ดึงออกมีความสำคัญกับโมเดลหรือไม่ (ถ้าค่า p-value < 0.05 ให้ดึงตัวแปรนั้นกลับเข้าไปดังเดิม) จากผลลัพธ์ดังกล่าวพบว่าเมื่อดึงตัวแปร bmi ออกแล้วการทดสอบ ANOVA ได้ค่า p-value = 0.825 จึงพิจารณาโมเดลที่ 2 (m2) ต่อไปและพบว่าตัวแปร smk มีค่า p-value สูงที่สุดคือ 0.732 จึงพิจารณาดึงตัวแปรนี้ออกและทดสอบ ANOVA

```
>m3<-glm(uric~fpg+age_gr+    whr     ,data=dt)
>anova(m3,m2,test = "F")
Analysis of Deviance Table

Model 1: uric ~ fpg + age_gr + whr
Model 2: uric ~ fpg + age_gr + whr + smk
  Resid. Df Resid. Dev Df Deviance      F Pr(>F)
1        25     607.62
2        24     604.60  1   3.0178 0.1198 0.7323
>summary(m3)
             Estimate Std. Error t value Pr(>|t|)
(Intercept)   268.435     21.406  12.540 2.78e-12 ***
fpg           -11.194      2.021  -5.538 9.32e-06 ***
age_gr46-55     2.739      2.710   1.011 0.321806
age_gr>55yrs    2.784      3.429   0.812 0.424553
whr            76.892     17.367   4.427 0.000164 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> m4<-glm(uric~fpg+           whr     ,data=dt)
> anova(m4,m3,test = "F")
Analysis of Deviance Table

Model 1: uric ~ fpg + whr
Model 2: uric ~ fpg + age_gr + whr
  Resid. Df Resid. Dev Df Deviance      F Pr(>F)
1        27     632.68
2        25     607.62  2   25.057 0.5155 0.6034
> summary(m4)

            Estimate Std. Error t value Pr(>|t|)
(Intercept)   267.88      20.95  12.784 5.75e-13 ***
fpg           -11.56       1.89  -6.117 1.56e-06 ***
whr            82.43      15.27   5.400 1.04e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

p-value จากการทดสอบ ANOVA

รูปที่ 9.16 ผลการวิเคราะห์ข้อมูล Multiple linear regression โมเดลที่ 3 และ 4

จากการวิเคราะห์โมเดลที่ 3 (m3) พบว่าเมื่อดึงตัวแปร smk ออกและโมเดลที่ 4 (m4) ได้ดึงตัวแปรกลุ่มอายุออกจะได้โมเดลสุดท้ายคือ m4 ซึ่งประกอบไปด้วยปัจจัย ระดับน้ำตาลในเลือดและสัดส่วนรอบเอวต่อรอบสะโพก จากนั้นจึงรายงานค่าสัมประสิทธิ์ของโมเดล (Coefficient) ดังรูปที่ 9.17

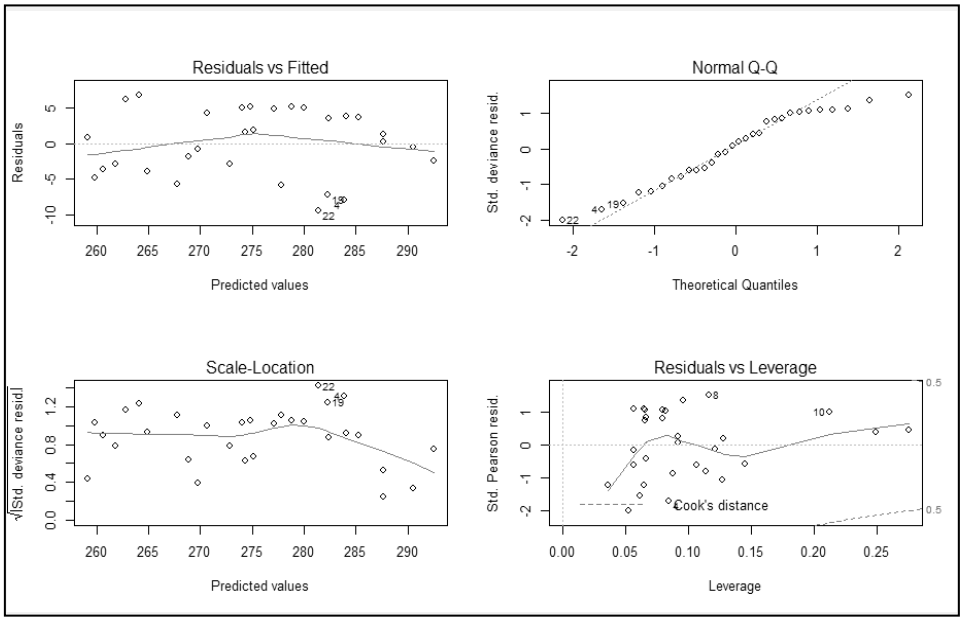


รูปที่ 9.17 ผลการวิเคราะห์ Multiple linear regression โมเดลสุดท้าย

แปลผลการวิเคราะห์ข้อมูล

หลังจากควบคุมปัจจัยด้านสัดส่วนเส้นรอบเอวต่อเส้นรอบสะโพกแล้วพบว่าเมื่อระดับน้ำตาลในเลือดมีค่าเพิ่มขึ้น 1 หน่วยจะทำให้ระดับกรดยูริกลดลง 11.56 หน่วย (95% CI: 7.68, 15.44) ในขณะที่ถ้าควบคุมปัจจัยด้าน ระดับน้ำตาลในเลือด แล้วพบว่าถ้าสัดส่วนเส้นรอบเอวต่อเส้นรอบสะโพกเพิ่มขึ้น 1 หน่วย ระดับกรดยูริกจะมีค่าเพิ่มขึ้น 82.43 หน่วย (95% CI: 51.11, 113.75)

การตรวจสอบโมเดลใช้คำสั่ง *plot(m4)* ได้ผลลัพธ์ดังรูปที่ 9.18



รูปที่ 9.18 กราฟการทดสอบข้อตกลงเบื้องต้นของโมเดล

จากกราฟรูปที่ 9.18สามารถแปลผลได้ดังนี้

1. รูป Residuals vs Fitted ลักษณะของข้อมูลมีความเป็นเชิงเส้นเพราะข้อมูลมีการกระจายโดยรอบเส้น Residual = 0

2. รูป Normal Q-Q ลักษณะของข้อมูลค่อนข้างมีการแจกแจงปกติเพราะข้อมูลมีการเกาะกลุ่มอยู่ในเส้นทแยง

3. รูป Scale-Location เป็นการทดสอบความเท่ากันของความแปรปรวน โดยดูจากเส้นที่ลากผ่านข้อมูลควรเป็นเส้นตรง ทั้งนี้สามารถทดสอบทางสถิติโดยใช้วิธีการของ T. S. Breusch and A. R. Pagan[25] ในโปรแกรม R สามารถเรียกใช้ได้จากชุดคำสั่ง lmtest[26] ผลการทดสอบความเท่ากันของความแปรปรวนดังรูปที่ 9.19 พบว่ามีความแปรปรวนเท่ากัน (p-value = 0.936)

```
> model <- lm(uric~fpg+whr,data=dt)
> bptest(model)
        studentized Breusch-Pagan test
data:  model
BP = 0.13252, df = 2, p-value = 0.9359
```

รูปที่ 9.19 ผลการทดสอบความเท่ากันของความแปรปรวน (Homoscedasticity)

4. รูป Residual vs Leverageเพื่อทดสอบโมเดลมีข้อมูลที่แยกออกจากกลุ่มหรือไม่ (Outliers)ในรูปนี้ดูได้จากเส้นปะ (สีแดง) ถ้ามีข้อมูลที่เลยออกจากเส้นปะนี้หมายถึงเมื่อลบข้อมูลนี้ออกจะมีอิทธิพลต่อผลลัพธ์ของโมเดลเป็นอย่างมาก

9.5 บทสรุป

การทดสอบความสัมพันธ์ของตัวแปรต่อเนื่องสองตัวแปรจะใช้สถิติ Pearson's correlation ถ้าข้อมูลมีการแจกแจงไม่ปกติทั้งสองตัวแปรหรือตัวแปรใดตัวแปรหนึ่ง จะใช้สถิติ Spearman's rank correlation การทดสอบดังกล่าวสามารถให้ข้อมูลได้ว่าสองตัวแปรที่ทดสอบมีความสัมพันธ์กันและระดับของความสัมพันธ์ อย่างไรก็ตามนักวิจัยอยากจะทราบว่า ถ้าตัวแปรต้นเปลี่ยนค่าไปหนึ่งหน่วยแล้วตัวแปรตามจะเปลี่ยนไปเท่าใด ให้ใช้สถิติ linear regression ทั้งนี้คำสั่งการวิเคราะห์ข้อมูลได้สรุปในตารางที่ 9.6

ตารางที่ 9.6 คำสั่งในโปรแกรม R ที่ใช้ในการวิเคราะห์ Correlation และ Multiple linear regression

การทดสอบ	คำสั่ง
Person's correlation	กรณีที่ข้อมูลมีการแจกแจงปกติ *cor.test(ชื่อตัวแปรที่ 1, ชื่อตัวแปรที่ 2, method=c("pearson"))*
	กรณีที่ข้อมูลมีการแจกแจงไม่ปกติ *cor.test(ชื่อตัวแปรที่ 1, ชื่อตัวแปรที่ 2, method=c("spearman"))*
Linear regression	*lm(ชื่อตัวแปรต้น~ชื่อตัวแปรอิสระ ถ้ามีหลายตัว คั่นด้วยเครื่องหมาย (+),data=ชื่อชุดข้อมูล)*
	ทดสอบตัวแปรระหว่างโมเดล *anova(ชื่อโมเดลที่ 1,ชื่อโมเดลที่ 2,test = "F")*
	รายงานค่า Coefficient และ ช่วงเชื่อมั่นที่ 95% *regress.display(ชื่อโมเดลสุดท้าย)*

9.6 แบบฝึกหัด

1. นักวิจัยต้องการทราบว่าค่าระดับน้ำตาลในเลือดเพิ่มขึ้น 1 หน่วยจะส่งผลให้ค่าความดันโลหิตซนิดซิสโตลิกเปลี่ยนไปกี่หน่วย จะใช้การวิเคราะห์แบบใด
2. จงบอกข้อดีของการวิเคราะห์ข้อมูลแบบ Multiple linear regression
3. การทดสอบผลของตัวแปรต่อโมเดลใช้สถิติอะไร
4. สหสัมพันธ์ของเพียร์สัน (r) มีค่าอยู่ระหว่างเท่าใด
5. การฟิตโมเดล (Fitting model) มีกี่วิธี อะไรบ้าง

9.7 เอกสารอ้างอิง

1. Arndt S, Turvey C, Andreasen NC. Correlating and predicting psychiatric symptom ratings: Spearman's r versus Kendall's tau correlation. J Psychiatr Res. 1999;33(2):97-104.
2. Bewick V, Cheek L, Ball J. Statistics review 7: Correlation and regression. Crit Care. 2003;7(6):451-9.
3. Miot HA. Correlation analysis in clinical and experimental studies. J Vasc Bras. 2018;17(4):275-9.
4. Schober P, Boer C, Schwarte LA. Correlation Coefficients: Appropriate Use and Interpretation. Anesth Analg. 2018;126(5):1763-8.
5. Williams S. Pearson's correlation coefficient. N Z Med J. 1996;109(1015):38.
6. Altman DG. Practical statistics for medical research: CRC press; 1990.
7. Mukaka MM. Statistics Corner: A guide to appropriate use of Correlation coefficient in medical research. Malawi Medical Journal. 2012;24(3):69-71.
8. Wickham H. ggplot2: Elegant Graphics for Data Analysis: Springer-Verlag New York; 2016.
9. Schober P, Boer C, Schwarte LA. Correlation Coefficients: Appropriate Use and Interpretation. Anesth Analg. 2018;126(5):1763-8.
10. de Winter JC, Gosling SD, Potter J. Comparing the Pearson and Spearman correlation coefficients across distributions and sample sizes: A tutorial using simulations and empirical data. Psychol Methods. 2016;21(3):273-90.

11. Rosner B, Glynn RJ. Interval estimation for rank correlation coefficients based on the probit transformation with extension to measurement error correction of correlated ranked data. Stat Med. 2007;26(3):633-46.
12. Cureton EE. The average Spearman rank correlation when ties are present: a correction. Psychometrika. 1965;30(3):377.
13. Rosner B. Fundamentals of biostatistics: Seventh edition. Boston : Brooks/Cole, Cengage Learning; 2011.
14. Schneider A, Hommel G, Blettner M. Linear regression analysis: part 14 of a series on evaluation of scientific publications. Dtsch Arztebl Int. 2010;107(44):776-82.
15. Kim HY. Statistical notes for clinical researchers: simple linear regression 1 - basic concepts. Restor Dent Endod. 2018;43(2):e21.
16. Zou KH, Tuncali K, Silverman SG. Correlation and simple linear regression. Radiology. 2003;227(3):617-22.
17. Aggarwal R, Ranganathan P. Common pitfalls in statistical analysis: Linear regression analysis. Perspect Clin Res. 2017;8(2):100-2.
18. Kleinbaum DG, Kupper LL, Nizam A, Rosenberg ES. Applied regression analysis and other multivariable methods: Nelson Education; 2013.
19. Aiken LS, West SG, Pitts SC, Baraldi AN, Wurpts IC. Multiple linear regression. Handbook of Psychology, Second Edition. 2012;2.
20. Zhang Z. Variable selection with stepwise and best subset approaches. Ann Transl Med. 2016;4(7):136.
21. Morozova O, Levina O, Uusküla A, Heimer R. Comparison of subset selection methods in linear regression in the context of health-related quality of life and substance abuse in Russia. BMC Med Res Methodol. 2015;15:71.
22. Vrieze SI. Model selection and psychological theory: a discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). Psychol Methods. 2012;17(2):228-43.
23. Hosmer Jr DW, Lemeshow S, Sturdivant RX. Applied logistic regression: John Wiley & Sons; 2013.

24. Chowdhury MZI, Turin TC. Variable selection strategies and its importance in clinical prediction modelling. Fam Med Community Health. 2020;8(1):e000262.
25. Breusch TS, Pagan AR. A Simple Test for Heteroscedasticity and Random Coefficient Variation. Econometrica. 1979;47(5):1287-94.
26. Zeileis A, Hothorn T. Diagnostic Checking in Regression Relationships. R News. 2002;2:3.

บทที่ 10

การวิเคราะห์การถดถอยลอจิสติก

ในบทที่ 9 ได้ศึกษาเกี่ยวกับการวิเคราะห์หลายตัวแปรในกรณีที่ผลลัพธ์ของการศึกษา (Outcome) เป็นตัวแปรต่อเนื่อง ในบทนี้จะได้ศึกษาการวิเคราะห์ข้อมูลในกรณีที่ Outcome มีลักษณะเป็นตัวแปรจำแนกประเภท (สองกลุ่ม) และตัวแปรต้นมี 1 ตัวแปรหรือมากกว่า[1, 2] การวิเคราะห์ข้อมูลจะใช้โมเดลทางสถิติชื่อ การถดถอยลอจิสติกอย่างง่าย (Simple logistic regression) ใช้ในกรณีที่ตัวแปรต้นมี 1 ตัวแปร[3,4] และการถดถอยพหุลอจิสติก (Multiple logistic regression)[5] ใช้ในกรณีที่ตัวแปรต้นมากกว่า 1 ตัวแปร

10.1 ตัวแบบการถดถอยพหุลอจิสติก (Multiple logistic regression model)

การวิเคราะห์ข้อมูลโดยใช้การถดถอยพหุลอจิสติก[5] นั้นเหมาะสมกับกรณีที่ผลลัพธ์ของการศึกษาแบ่งออกเป็น 2 กลุ่ม โดยมีสมการตัวแบบทางคณิตศาสตร์ดังนี้

$$\text{logit}(P) = \beta_0 + \beta_1 + x_1 + \beta_i + x_i + \varepsilon_i$$

โดยที่

P = ความน่าจะเป็นของการเกิดเหตุการณ์

X = ตัวแปรอิสระหรือตัวแปรต้น

β_0 = จุดตัดแกน Y เมื่อ X_i เท่ากับ 0

β_1 = ค่าความชันของกราฟเส้น (Coefficients)

ε_i = Residual errors (ความคลาดเคลื่อนของการพยากรณ์)

ทั้งนี้ตัวแปรตาม (ตัวแปร Y หรือ outcome หรือ dependent factor) มีตัวแปรเดียว และมีลักษณะเป็น dichotomous หรือ binary outcome (เช่น ป่วยหรือ ไม่ป่วย, เกิดโรคหรือไม่เกิดโรค, หายหรือไม่หาย) และตัวแปรอิสระ (X: risk factors, independent factors) มีได้หลายตัวแปรซึ่งอาจเป็นตัวแปรต่อเนื่องหรือตัวแปรจำแนกประเภท (ตัวแปรกลุ่ม) ก็ได้ ถ้าเป็นตัวแปร

กลุ่มเวลาทดสอบโมเดลต้องมีลักษณะเป็นตัวแปรหุ่น (Dummy variable)[6-8] ทั้งนี้การประมาณค่าพารามิเตอร์ของโมเดลใช้วิธีที่ชื่อว่า Maximum likelihood (ความน่าจะเป็นที่เป็นไปได้มากที่สุด) และการเปรียบเทียบโมเดลจะใช้ Likelihood ratio test (LR test)[9]

ข้อตกลงเบื้องต้นสาหรับสถิติได้[10]

1. ค่าคลาดเคลื่อน (Error) เป็นอิสระต่อกัน
2. ไม่มี Multicolinearity ระหว่างตัวแปรอิสระ

ตัวอย่างที่ 10.1 การศึกษาของ Wondie AG และคณะ (2019) ได้ทำการศึกษาปัจจัยที่มีผลต่อการผ่าตัดคลอดในประเทศเอธิโอเปีย[11] โดยมีตัวอย่างชุดข้อมูลดังตารางที่ 10.1

ตารางที่ 10.1 ตัวอย่างชุดข้อมูลการศึกษาปัจจัยที่มีผลต่อการผ่าตัดคลอดในประเทศเอธิโอเปีย

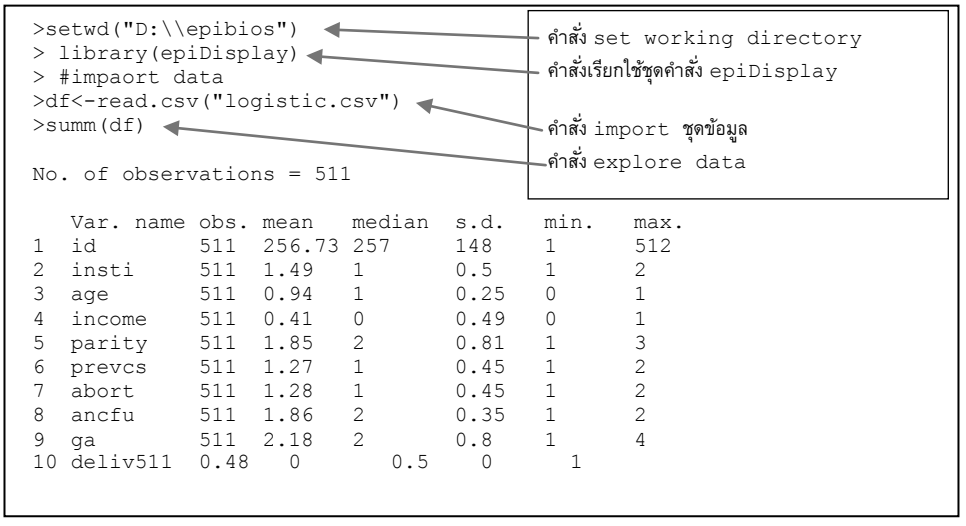
id	insti	age	income	parity	prevcs	abort	ancfu	ga	deliv
1	1	1	1	1	1	1	2	2	1
2	1	1	0	1	1	2	2	3	1
3	1	1	0	1	1	2	2	3	1
4	1	0	1	3	2	2	2	2	1
5	1	1	1	1	1	2	2	2	1
6	1	1	1	1	1	1	2	3	1
7	1	1	1	3	1	2	2	2	1
8	1	1	1	1	1	1	2	3	1
9	1	1	1	1	1	1	2	3	1
.	.	.	.	.	.	.	.	.	.

ตารางที่ 10.2 ดัชนีข้อมูลการศึกษาปัจจัยที่มีผลต่อการผ่าตัดคลอดในประเทศเอธิโอเปีย

ตัวแปร	ความหมาย	ค่า	ความหมาย
insti	โรงพยาบาล (Institution type)	1	Private Hospital
		2	Public Hospital
age	อายุ ปี (Age)	0	<35yrs
		1	≥35
income	รายได้ต่อเดือน (Monthly income)	0	≤2000
		1	>2000
parity	จำนวนครั้งของการคลอดบุตร (Parity)	1	First
		2	Second
		3	Third and above
prevcs	ประวัติการผ่าคลอด (Previous cesarean section)	1	No
		2	Yes
abort	ประวัติการแท้งบุตร (Abortion history)	1	No
		2	Yes
ancfu	การฝากครรภ์ (ANC followup)	1	No
		2	Yes
ga	อายุครรภ์ (Gestational age)	1	≤36
		2	37 - 39
		3	40 - 42
		4	≥43
deliv	วิธีการคลอด (Mode of delivery)	0	Vaginal delivery
		1	Cesarean section

วิธีทำ

ขั้นตอนที่ 1 นำเข้าข้อมูลและจัดการข้อมูล



รูปที่ 10.1 การนำเข้าข้อมูลและสำรวจข้อมูลเบื้องต้น

ในตัวอย่างนี้ได้ทำการนำเข้าข้อมูลจากข้อมูลที่อยู่ในรูปแบบ csv (comma delimited) โดยใช้คำสั่ง read.csv(“logistic.csv”) ชื่อของข้อมูลตัวอย่างคือ logistic.csv จากการสำรวจข้อมูลเบื้องต้นพบว่ามีตัวแปรกลุ่มที่มากกว่า 2 กลุ่มนั้น คือ ตัวแปร parity และตัวแปร ga ตัวแปรที่มีลักษณะเป็นตัวแปรจำแนกประเภทและมีระดับมากกว่า 2 ระดับ จะทำเป็นตัวแปรหุ่น (Dummy variable) ก่อนนำตัวแปรทั้งหมดเหล่านั้นเข้าไปในตัวแบบทางสถิติ การทำตัวแปร parity ให้เป็นตัวแปรหุ่นดังนี้

ตารางที่ 10.3 การแบ่งจำนวนครั้งของการคลอดบุตร (Parity)(ตัวแปรชื่อ parity)

1	ครั้งแรก (first)
2	ครั้งที่สอง (second)
3	ครั้งที่ 3 ขึ้นไป (third or more)

จากเกณฑ์ดังกล่าวเราจะให้กลุ่มคลอดครั้งแรก (first) เป็นกลุ่มอ้างอิงการสร้างตัวแปรหุ่นสามารถทำได้ 2 วิธี คือ

แบบที่ 1 สร้างตัวแปรและให้โค้ด

ตารางที่ 10.4 การสร้างตารางหุ่น

จำนวนครั้งของการคลอดบุตร (Parity)	first	second	third or more
f.parity	1	0	0
s.parity	0	1	0
tm.parity	0	0	1

ตารางที่ 10.4 กลุ่มอ้างอิงจะให้รหัสเป็น 0 ทั้งหมด ที่เหลือจะเป็นรหัส 1 ตามตัวแปรที่เรากำหนดเช่นกลุ่ม first parity (f.parity) จะให้โค้ดเป็น 1 ถ้า จำนวนครั้งของการคลอดบุตรมีค่าเท่ากับ 1 ค่าอื่นๆ จะให้รหัสเป็นเลข 0 ดังนั้นจะสามารถสร้างตัวแปรหุ่นได้ 3 ตัวแปรสามารถใช้โค้ดในโปรแกรม R ได้ดังนี้

```
f.parity<-ifelse(parity ==1, 1, 0)
s.parity<-ifelse(parity ==2, 1, 0)
tm.parity<-ifelse(parity ==3, 1, 0)
```

เมื่อนำตัวแปรเหล่านี้เข้าไปทดสอบในโมเดลจะต้องเอาเข้าทั้งชุด (3 ตัวแปร) นั้นคือ f.paritys.parityและ tm.parity เช่นในสมการถดถอยพหุลอจิสติก

```
deliv ~ age + income + prevcs+f.parity+s.parity+tm.parity
```

แบบที่ 2 สร้างตัวแปรกลุ่มขึ้นมาแล้วทำเป็นตัวแปรอินดิเคเตอร์ (Indicator) ในขั้นตอนการวิเคราะห์หลายตัวแปรเช่น

```
data$paritygr<-order(data$parity,
            levels=c(1,2,3),
                 lables=c("first", "second","third or more"))
```

ตารางที่ 10.5 ตัวแปรหุ่นโดยใช้วิธีกำหนดอินดิเคเตอร์ (Indicator) ในโมเดล

จำนวนครั้งของการคลอดบุตร (parity)	first	second	third or more
s.parity	0	1	0
tm.parity	0	0	1

เมื่อแทนค่าในสมการถดถอยพหุลอจิสติกได้ดังนี้

```
deliv ~ age + income + prevcs + paritygr
```

ทั้งนี้ตัวแปร parity ต้องมีคุณสมบัติเป็นแบบ factor เมื่อเปรียบเทียบค่าสัมประสิทธิ์ (Coefficient) ของทั้งสองวิธีจะมีค่าเท่ากัน การสร้างตัวแปรหุ่นทั้ง 2 วิธีที่ผ่านมาพบว่าเมื่อใช้วิธีตัวแปรหุ่นแบบที่ 1 สมมุติว่ากลุ่ม first เทียบกับกลุ่มอื่นๆ (second and third or more) แต่ถ้าเราใช้การทำตัวแปรหุ่นแบบที่ 2 มีกลุ่ม first เป็นอันดับแรกและเป็นกลุ่มอ้างอิงแปลผลได้ว่า กลุ่มที่ second เทียบกับ first และ กลุ่ม third or more เทียบกับ first ซึ่งในการวิจัยและการวิเคราะห์ข้อมูลนิยมสร้างตัวแปรหุ่นแบบที่ 2

```
> tab1(df$insti,graph = FALSE)
df$insti :
        Frequency Percent Cum. percent
1             259    50.7         50.7
2             252    49.3        100.0
  Total       511   100.0        100.0
>df$insti<-
factor(df$insti,levels=c("1","2"),labels=c("private","public"))
> tab1(df$insti,graph = FALSE)
df$insti :
        Frequency Percent Cum. percent
private       259    50.7         50.7
public        252    49.3        100.0
  Total       511   100.0        100.0
> tab1(df$age,graph = FALSE)
df$age :
        Frequency Percent Cum. percent
0              33     6.5          6.5
1             478    93.5        100.0
  Total       511   100.0        100.0
>df$age<-factor(df$age,levels=c("0","1"),labels=c("<35yrs",">= 35"))
> tab1(df$age,graph = FALSE)
df$age :
        Frequency Percent Cum. percent
<35yrs         33     6.5          6.5
>= 35         478    93.5        100.0
  Total       511   100.0        100.0
```

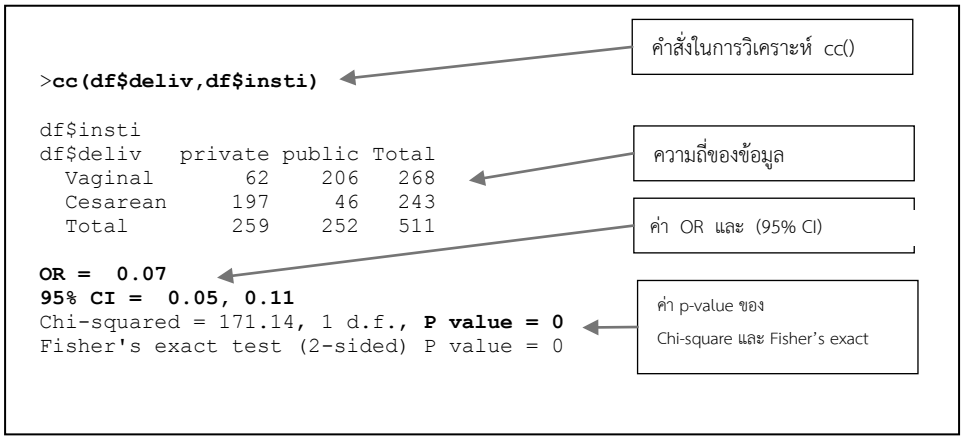
รูปที่ 10.2 การจัดการข้อมูลและการสำรวจข้อมูลตัวแปรสถาบันและกลุ่มอายุ

รูปที่ 10.2 การจัดการข้อมูลตัวแปร instและ age จากดรรชนีข้อมูล ตัวแปรทั้งสองตัวแปรกลุ่ม ดังนั้นต้องทำให้มีลักษณะเป็นแบบ factor (ใช้ในกรณีที่ข้อมูลมีสองกลุ่ม) หรือ order (ใช้ในกรณีที่ข้อมูลมีมากกว่าสองกลุ่ม) จากตัวอย่างจะใช้คำสั่งในโปรแกรม R ดังนี้

```
df$insti<-factor(df$insti,
                 levels=c("1","2"),
                 labels=c("private","public"))
```

ขั้นตอนที่ 2 การวิเคราะห์ข้อมูลสองตัวแปร

การวิเคราะห์ความสัมพันธ์ระหว่างประเภทของโรงพยาบาลกับวิธีการคลอด



รูปที่ 10.3 ผลการวิเคราะห์สัมพันธ์ระหว่างประเภทของโรงพยาบาลกับวิธีการคลอด

ผลการวิเคราะห์ข้อมูลพบว่าประเภทของโรงพยาบาลกับวิธีการคลอดมีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติ (p-value < 0.001) โดยโรงพยาบาลในสังกัดของรัฐ (Public hospital) มีโอกาสในการผ่าคลอดเท่ากับ 0.07 เท่าเมื่อเทียบกับโรงพยาบาลในสังกัดเอกชน (Private hospital) (95% CI: 0.05, 0.11) หรือ โรงพยาบาลในสังกัดของเอกชนมีโอกาสในการผ่าคลอดเท่ากับ 14.3 (1/0.07) เท่าเมื่อเทียบกับโรงพยาบาลในสังกัดของรัฐ (95% CI: 9.1 (1/0.11), 20.0 (1/0.05))

การวิเคราะห์ความสัมพันธ์ระหว่างอายุของแม่ตั้งครรภ์กับวิธีการคลอด

```
> cc(df$deliv,df$age)

df$age
df$deliv<35yrs >= 35 Total
  Vaginal      16   252   268
  Cesarean     17   226   243
  Total        33   478   511

OR =  0.84
95% CI =  0.42, 1.71
Chi-squared = 0.22, 1 d.f., P value = 0.638
Fisher's exact test (2-sided) P value = 0.72
```

รูปที่ 10.4 ผลการวิเคราะห์สัมพันธ์ระหว่างอายุแม่ตั้งครรภ์กับวิธีการคลอด

ผลการวิเคราะห์ข้อมูลพบว่าอายุแม่ตั้งครรภ์กับวิธีการคลอดไม่มีความสัมพันธ์กัน (p-value = 0.638)

การวิเคราะห์ความสัมพันธ์ระหว่างรายได้ต่อเดือนกับวิธีการคลอด

```
> cc(df$deliv,df$income)

df$income
df$deliv<= 2000 >2000 Total
  Vaginal      168   100   268
  Cesarean     136   107   243
  Total        304   207   511

OR =  1.32
95% CI =  0.93, 1.88
Chi-squared = 2.39, 1 d.f., P value = 0.122
Fisher's exact test (2-sided) P value = 0.126
```

รูปที่ 10.5 ผลการวิเคราะห์สัมพันธ์ระหว่างรายได้ต่อเดือนกับวิธีการคลอด

ผลการวิเคราะห์ข้อมูลพบว่ารายได้ต่อเดือนกับวิธีการคลอดไม่มีความสัมพันธ์กัน (p-value = 0.122)

การวิเคราะห์ความสัมพันธ์ระหว่างจำนวนครั้งของการคลอดบุตรกับวิธีการคลอด

```
> cc(df$deliv,df$parity)

df$parity
df$deliv         First Second >=3
  Vaginal        104   88     76
  Cesarean       108   75     60

  Odds ratio     1     0.82   0.76
  lower 95% CI         0.53   0.48
  upper 95% CI         1.26   1.2

Chi-squared = 1.776 , 2 d.f., P value = 0.412
Fisher's exact test (2-sided) P value = 0.408
```

รูปที่ 10.6 ผลการวิเคราะห์สัมพันธ์ระหว่างจำนวนครั้งของการคลอดบุตรกับวิธีการคลอด

ผลการวิเคราะห์ข้อมูลพบว่าจำนวนครั้งของการคลอดบุตรกับวิธีการคลอดไม่มีความสัมพันธ์กัน (p-value = 0.412) จากผลการวิเคราะห์พบว่าตัวแปร parity มี 3 กลุ่ม โดยที่กลุ่มแรก (first) เป็นกลุ่มอ้างอิง จะมีค่า OR เท่ากับ 1 มีความหมายคือคนที่คลอดครั้งที่ 2 มีโอกาสเสี่ยงทำให้มีการผ่าคลอดเท่ากับ 0.82 เท่าเมื่อเทียบกับคนที่คลอดครั้งแรก และคนที่คลอดครั้งที่ 3 ขึ้นไป มีโอกาสเสี่ยงทำให้มีการผ่าคลอดเท่ากับ 0.76 เท่าเมื่อเทียบกับคนที่คลอดครั้งแรก

การวิเคราะห์ความสัมพันธ์ระหว่างประวัติการผ่าคลอดกับวิธีการคลอด

```
> cc(df$deliv,df$prevcs)

df$prevcs
df$deliv    no yes Total
Vaginal  229  39   268
  Cesarean 143 100   243
  Total    372 139   511

OR =  4.11
95% CI =  2.68, 6.28
Chi-squared = 45.54, 1 d.f., P value = 0
Fisher's exact test (2-sided) P value = 0
```

รูปที่ 10.7 ผลการวิเคราะห์สัมพันธ์ประวัติการผ่าคลอดกับระหว่างวิธีการคลอด

ผลการวิเคราะห์ข้อมูลพบว่าประวัติการผ่าคลอดกับวิธีการคลอดมีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติ (p-value < 0.001) โดยพบว่าคนที่มีประวิติผ่าคลอดมีโอกาสเสี่ยงทำให้ผ่าคลอดเท่ากับ 4.11 เท่าเมื่อเทียบกับคนที่ไม่มีประวัติ (95% CI: 2.68, 6.28)

การวิเคราะห์ความสัมพันธ์ระหว่างประวัติการแท้งบุตรกับวิธีการคลอด

```
> cc(df$deliv,df$abort)

df$abort
df$deliv    no yes Total
Vaginal  195  73   268
  Cesarean 175  68   243
  Total    370 141   511

OR =  1.04
95% CI =  0.7, 1.53
Chi-squared = 0.04, 1 d.f., P value = 0.851
Fisher's exact test (2-sided) P value = 0.921
```

รูปที่ 10.8 ผลการวิเคราะห์สัมพันธ์วิธีการคลอดกับระหว่างวิธีการคลอด

ผลการวิเคราะห์ข้อมูลแบบสองตัวแปรพบว่าการมีประวัติการทำแท้งกับวิธีการคลอดไม่มีความสัมพันธ์กัน (p-value = 0.851)

การวิเคราะห์ความสัมพันธ์ระหว่างอายุครรภ์กับวิธีการคลอด

```
> cc(df$deliv,df$ga)

df$ga
df$deliv<=36 37-39 40-42 >=43
  Vaginal        54   155   43    16
  Cesarean       36   118   69    20

  Odds ratio     1    1.14  2.4   1.87
  lower 95% CI        0.68  1.31  0.8
  upper 95% CI        1.92  4.42  4.42

Chi-squared = 13.905 , 3 d.f., P value = 0.003
Fisher's exact test (2-sided) P value = 0.003
```

รูปที่ 10.9 ผลการวิเคราะห์สัมพันธ์ระหว่างอายุครรภ์กับวิธีการคลอด

ผลการวิเคราะห์ข้อมูลพบว่าวิธีการคลอดกับอายุครรภ์มีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติ (p-value = 0.003) โดยพบว่าคนที่มีอายุครรภ์ 37-39 สัปดาห์ มีโอกาสเสี่ยงในการผ่าคลอดเท่ากับ 1.14 (95% CI: 0.68, 1.92) เท่าเมื่อเทียบกับคนที่มีอายุครรภ์น้อยกว่าหรือเท่ากับ 36 สัปดาห์ และคนที่มีอายุครรภ์ 40-42 สัปดาห์ มีโอกาสเสี่ยงในการผ่าคลอดเท่ากับ 2.4 (95% CI: 1.31, 4.42) เท่าเมื่อเทียบกับคนที่มีอายุครรภ์น้อยกว่าหรือเท่ากับ 36 สัปดาห์นอกจากนั้นแล้วคนที่มีอายุครรภ์มากกว่าหรือเท่ากับ 43 สัปดาห์ มีโอกาสเสี่ยงในการผ่าคลอดเท่ากับ 1.87 (95% CI: 0.80, 4.42) เท่าเมื่อเทียบกับคนที่มีอายุครรภ์น้อยกว่าหรือเท่ากับ 36 สัปดาห์

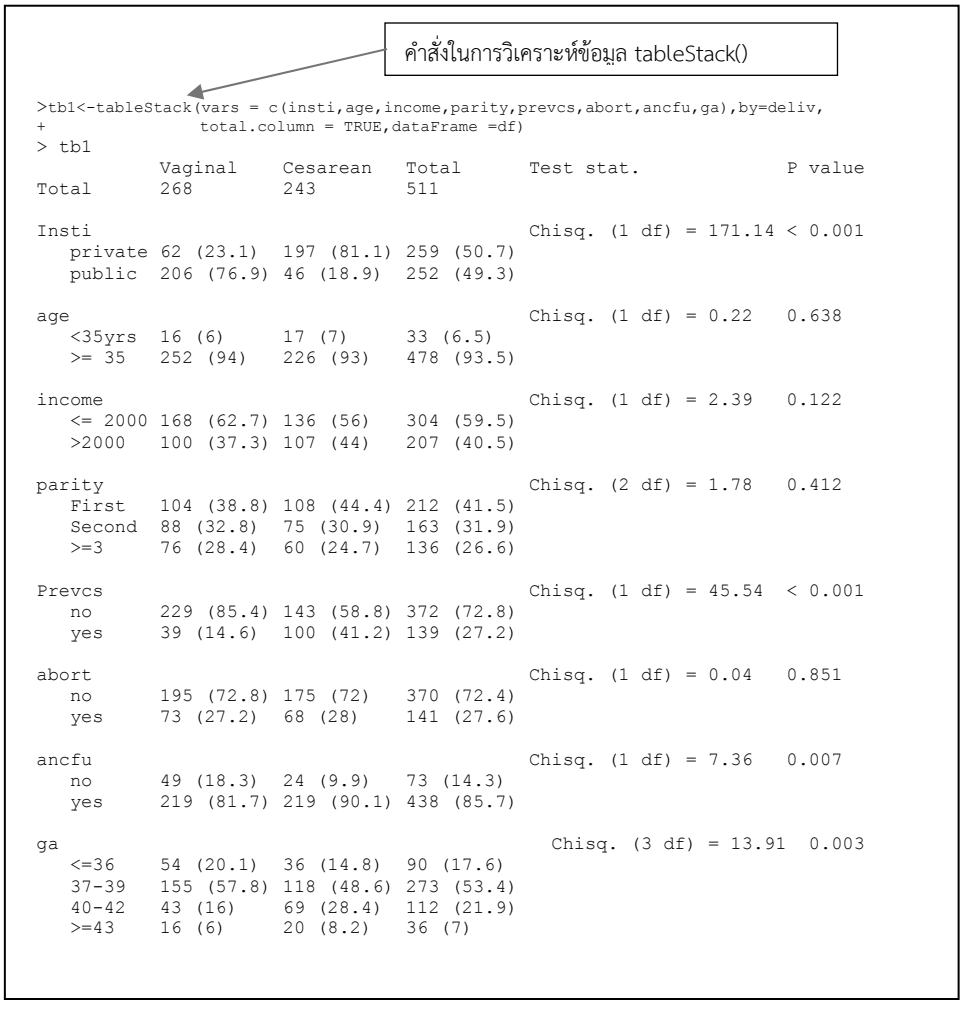
การวิเคราะห์ความสัมพันธ์ระหว่างการฝากครรภ์กับวิธีการคลอด

```
> cc(df$deliv,df$ancfu)

df$ancfu
df$deliv    no yes Total
  Vaginal   49 219   268
Cesarean  24 219   243
  Total     73 438   511

OR =  2.04
95% CI =  1.21, 3.44
Chi-squared = 7.36, 1 d.f., P value = 0.007
Fisher's exact test (2-sided) P value = 0.008
```

รูปที่ 10.10 ผลการวิเคราะห์สัมพันธ์ระหว่างวิธีการคลอดกับวิธีการคลอด

ผลการวิเคราะห์ข้อมูลพบว่าการฝากครรภ์กับวิธีการคลอดมีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติ (p-value = 0.007) โดยพบว่าคนที่มีการฝากครรภ์มีโอกาสเสี่ยงในการผ่าคลอดเท่ากับ 2.04 (95% CI: 1.21, 3.44) เท่าเมื่อเทียบกับคนที่ไม่มีการฝากครรภ์

จากผลการวิเคราะห์ทั้ง 8 ตัวแปรพบตัวแปรที่มีนัยสำคัญทางสถิติประกอบไปด้วย ชนิดของโรงพยาบาล (insti), ประวัติการผ่าคลอด (prevcs), อายุครรภ์ (ga), และ การฝากครรภ์ (ancfu) ทั้งนี้จะนำตัวแปรเหล่านี้เข้าไปวิเคราะห์ในขั้นตอน Multiple logistic regression อย่างไรก็ตามในชุดคำสั่งของ *epiDisplay* สามารถใช้คำสั่ง *tableStack* ในการวิเคราะห์ข้อมูลในครั้งเดียวได้ดังนี้




```
>tb1<-tableStack(vars = c(insti,age,income,parity,prevcs,abort,ancfu,ga),by=deliv,
+               total.column = TRUE,dataFrame =df)
> tb1
           Vaginal    Cesarean   Total      Test stat.              P value
Total      268        243        511

Insti                                       Chisq. (1 df) = 171.14 < 0.001
   private 62 (23.1)  197 (81.1) 259 (50.7)
   public  206 (76.9) 46 (18.9)  252 (49.3)

age                                         Chisq. (1 df) = 0.22   0.638
   <35yrs  16 (6)     17 (7)     33 (6.5)
   >= 35   252 (94)   226 (93)   478 (93.5)

income                                      Chisq. (1 df) = 2.39   0.122
   <= 2000 168 (62.7) 136 (56)   304 (59.5)
   >2000   100 (37.3) 107 (44)   207 (40.5)

parity                                      Chisq. (2 df) = 1.78   0.412
   First   104 (38.8) 108 (44.4) 212 (41.5)
   Second  88 (32.8)  75 (30.9)  163 (31.9)
   >=3     76 (28.4)  60 (24.7)  136 (26.6)

Prevcs                                      Chisq. (1 df) = 45.54  < 0.001
   no      229 (85.4) 143 (58.8) 372 (72.8)
   yes     39 (14.6)  100 (41.2) 139 (27.2)

abort                                       Chisq. (1 df) = 0.04   0.851
   no      195 (72.8) 175 (72)   370 (72.4)
   yes     73 (27.2)  68 (28)    141 (27.6)

ancfu                                       Chisq. (1 df) = 7.36   0.007
   no      49 (18.3)  24 (9.9)   73 (14.3)
   yes     219 (81.7) 219 (90.1) 438 (85.7)

ga                                            Chisq. (3 df) = 13.91  0.003
   <=36    54 (20.1)  36 (14.8)  90 (17.6)
   37-39   155 (57.8) 118 (48.6) 273 (53.4)
   40-42   43 (16)    69 (28.4)  112 (21.9)
   >=43    16 (6)     20 (8.2)   36 (7)
```

รูปที่ 10.11 ผลการวิเคราะห์แบบสองตัวแปร

จากการใช้คำสั่ง *tableStack()* ดังรูปที่ 10.11 สามารถวิเคราะห์ข้อมูลและค้นหาตัวแปรที่จะนำเข้าไปวิเคราะห์แบบหลายตัวแปรได้รวดเร็วขึ้น และยังสามารถส่งออกผลการวิเคราะห์ให้อยู่ในรูปแบบ csv (comma delimited) คล้ายๆรูปแบบ excel โดยใช้คำสั่ง *write.csv()* เพื่อให้ง่ายต่อการรายงานผลการวิเคราะห์ข้อมูล

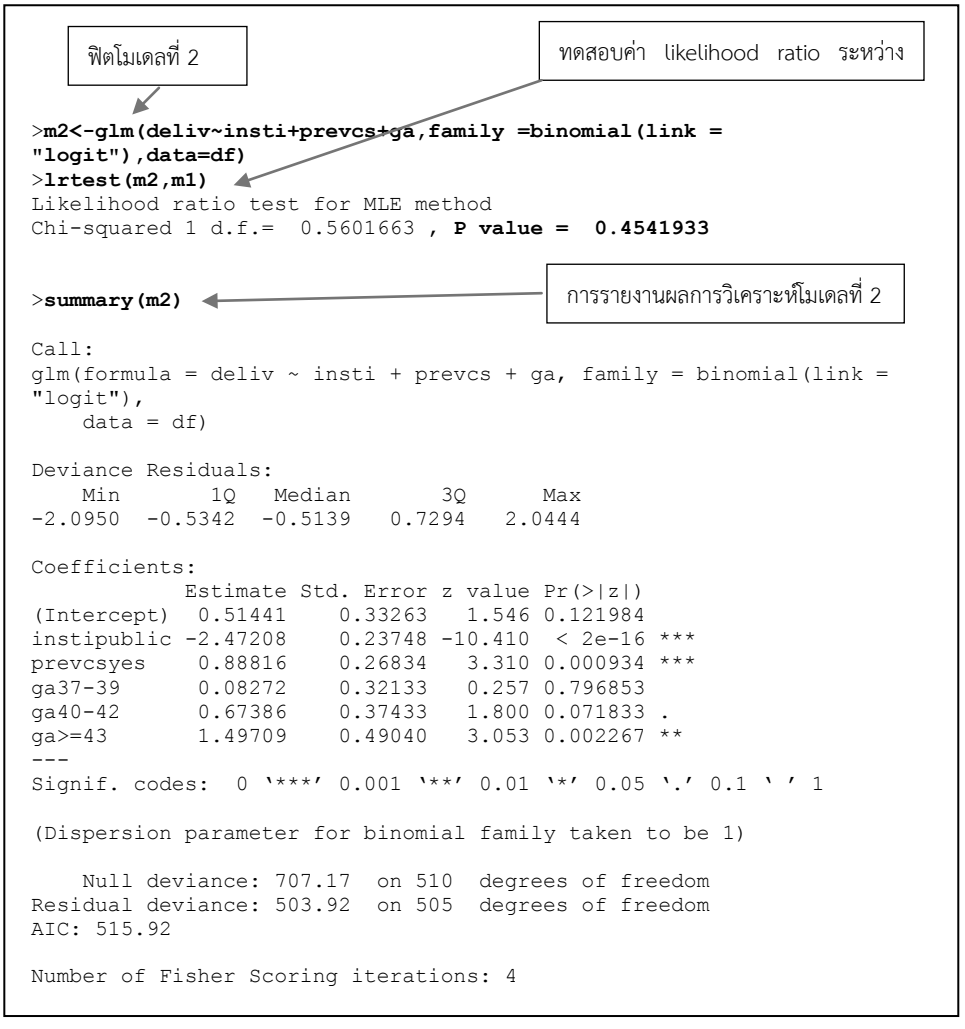
ขั้นตอนที่ 3 การวิเคราะห์หลายตัวแปรโดยใช้ Multiple logistic regression

การฟิต โมเดลที่ 1

```
> m1<-glm(deliv~insti+prevcs+ancfu+ga,family =binomial(link =
"logit"),data=df)
> summary(m1)

Call:
glm(formula = deliv ~ insti + prevcs + ancfu + ga, family =
binomial(link = "logit"),
    data = df)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.1041  -0.5471  -0.4864   0.7277   2.1376

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)   0.2651     0.4712   0.563 0.573720
instipublic  -2.4425     0.2401 -10.171  < 2e-16 ***
prevcsyes     0.9043     0.2693   3.358 0.000785 ***
ancfuyes      0.2515     0.3375   0.745 0.456142
ga37-39       0.1024     0.3221   0.318 0.750417
ga40-42       0.6770     0.3744   1.809 0.070521 .
ga>=43        1.5319     0.4936   3.104 0.001912 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 707.17  on 510  degrees of freedom
Residual deviance: 503.36  on 504  degrees of freedom
AIC: 517.36

Number of Fisher Scoring iterations: 4
```

รูปที่ 10.12 ผลการวิเคราะห์ข้อมูลด้วย Multiple logistic regression โมเดลที่ 1

ผลการวิเคราะห์ในรูปที่ 10.12 พบว่าโมเดลชื่อ m1 ตัวแปร ancfuyes (ancfu) ไม่มีนัยสำคัญทางสถิติ (p-value = 0.456) ดังนั้นจึงทำการขจัดตัวแปร ancfu ออกจากโมเดลและทำการฟิตโมเดลที่ 2

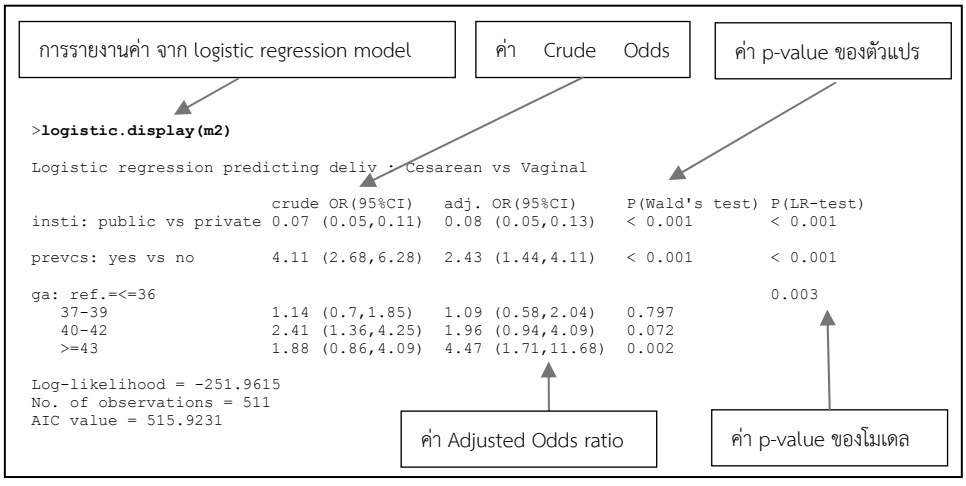



```
>m2<-glm(deliv~insti+prevcs+ga,family =binomial(link =
"logit"),data=df)
>lrtest(m2,m1)
Likelihood ratio test for MLE method
Chi-squared 1 d.f.=  0.5601663 , P value =  0.4541933

>summary(m2)

Call:
glm(formula = deliv ~ insti + prevcs + ga, family = binomial(link =
"logit"),
    data = df)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.0950  -0.5342  -0.5139   0.7294   2.0444

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  0.51441    0.33263   1.546 0.121984
instipublic -2.47208    0.23748 -10.410  < 2e-16 ***
prevcsyes    0.88816    0.26834   3.310 0.000934 ***
ga37-39      0.08272    0.32133   0.257 0.796853
ga40-42      0.67386    0.37433   1.800 0.071833 .
ga>=43       1.49709    0.49040   3.053 0.002267 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 707.17  on 510  degrees of freedom
Residual deviance: 503.92  on 505  degrees of freedom
AIC: 515.92

Number of Fisher Scoring iterations: 4
```

รูปที่ 10.13 ผลการวิเคราะห์ข้อมูลด้วย Multiple logistic regression โมเดลที่ 2

จากรูปที่ 10.13พบว่าเมื่อทดสอบ *lrtest()* ระหว่างโมเดลm1 และ m2 ใช้คำสั่ง *lrtest (m2,m1)* ได้ค่า p-value = 0.454 หมายความว่าตัวแปรชื่อ ancfu ไม่มีความสำคัญกับโมเดลนี้ (สามารถขจัดตัวแปรนี้ออกได้) จากนั้นจึงพิจารณาโมเดลที่ 2 (m2) พบว่ามีค่า p-value น้อยกว่า 0.05 ทุกตัวแปร (ตัวแปร ga แบ่งออกเป็น 4 กลุ่ม ในกลุ่ม ga ≥ 43 มีค่า p-value = 0.002 จึงต้องเก็บตัวแปร ga นี้ไว้ในโมเดลทั้งที่กลุ่มที่เหลือมีค่า p-value มากกว่า 0.05)

โมเดลสุดท้ายจึงประกอบไปด้วยตัวแปร ประเภทของโรงพยาบาล (insti), ประวัติการผ่าคลอด (prevcs), และอายุครรภ์ (ga) การรายงานค่า OR โดยใช้คำสั่งในชุดคำสั่ง *epiDisplay* ชื่อว่า *logistic.display()* ได้ผลดังนี้



รูปที่ 10.14 ผลการวิเคราะห์ข้อมูลด้วย Multiple logistic regression โมเดลสุดท้าย

จากการวิเคราะห์ข้อมูลรูปที่ 10.14 พบว่าหลังจากควบคุมปัจจัยด้าน ประเภทของโรงพยาบาลและอายุครรภ์แล้วผู้ที่มีประวัติการผ่าคลอดมีโอกาสเสี่ยงที่จะมีการผ่าคลอดเป็น 2.43 เท่า (95% CI:1.44,4.11) เมื่อเทียบกับผู้ที่ไม่มีประวัติการผ่าคลอด

จากนั้นทำการทดสอบความเหมาะสมของโมเดลหรือความเข้ากันได้ของชุดข้อมูลที่ได้ด้วย Goodness of fit[12-15] ในชุดคำสั่งชื่อว่า *ResourceSelection*[16] และใช้คำสั่ง *hoslem.test()* ได้ผลการทดสอบดังนี้

```
> library(ResourceSelection)
>df$deliv<-as.numeric(df$deliv)
>df$deliv<-ifelse(df$deliv==2,1,0)
> tab1(df$deliv)
df$deliv :
        Frequency Percent Cum. percent
0             268    52.4         52.4
1             243    47.6        100.0
  Total       511   100.0        100.0
>hoslem.test(df$deliv, fitted(m2))

        Hosmer and Lemeshow goodness of fit (GOF) test

data:  df$deliv, fitted(m2)
X-squared = 5.6434, df = 8, p-value = 0.6871
```

รูปที่ 10.15 ผลการวิเคราะห์ข้อมูลการทดสอบ Goodness of fit ของโมเดลสุดท้าย

จากรูปที่ 10.15 ค่า p-value ของ Hosmer and Lemeshow goodness of fit (GOF) test[17-20] เท่ากับ 0.687 สรุปได้ว่าข้อมูลที่มีสามารถเข้าได้ดีกับโมเดลนี้

10.2 บทสรุป

จากการวิเคราะห์ข้อมูลโดยใช้โมเดลทางสถิติมีข้อดีคือสามารถควบคุมตัวแปรหรือปัจจัยที่มีลักษณะตัวแปรกวนได้หลายตัวแปรและสามารถใช้คำสั่งในชุดคำสั่งของ epiDisplay เพื่อให้สามารถวิเคราะห์ข้อมูล ส่งออกตารางผลการวิเคราะห์ข้อมูลให้อยู่ในรูปแบบ csv เพื่อช่วยลดความผิดพลาดในการพิมพ์ตัวเลขสำหรับการรายงานผลการวิเคราะห์และยังทำให้มีความรวดเร็วในการรายงาน

10.3 แบบฝึกหัด

1. การวิเคราะห์ด้วย Multiple logistic regression สามารถวิเคราะห์ข้อมูลที่มีตัวแปรอิสระได้สูงสุดกี่ตัวแปร
2. การทดสอบ Goodness of fit ถ้าทดสอบโมเดลสุดท้ายแล้วมีค่าความน่าจะเป็นน้อยกว่า 0.05 ซึ่งหมายความว่าโมเดลที่ได้เข้าได้ไม่ดีกับข้อมูล จะมีวิธีการปรับแก้ไขโมเดลอย่างไร
3. ผลลัพธ์การศึกษาเป็นตัวแปรต่อเนื่องเช่น ดัชนีมวลกาย (Body mass index: BMI) อย่างไรก็ตามจะทำการจัดกลุ่มใหม่ให้เป็น 2 กลุ่ม คือ BMI < 30 kg/m^2 และ BMI ≥ 30kg/m^2 จากนั้นทำการทดสอบหลายตัวแปรปัจจัยเสี่ยงหลายตัวแปร กรณีนี้จะใช้โมเดลทางสถิติชื่อว่าอะไร
4. ในการวิเคราะห์ข้อมูลแบบสองตัวแปรเราสามารถใช้คำสั่งอะไรในชุดคำสั่ง epiDisplay เพื่อให้ได้ผลการวิเคราะห์แบบครั้งเดียวได้หลายตัวแปร
5. การฟิตโมเดล กรณีที่ตัวแปรที่ขจัดออกจากโมเดลมีค่า p-value < 0.05 เราจะทำอย่างไรกับตัวแปรนั้น

10.4 เอกสารอ้างอิง

1. Sperandei S. Understanding logistic regression analysis. Biochem Med (Zagreb). 2014;24(1):12-8.
2. Stoltzfus JC. Logistic regression: a brief primer. Acad Emerg Med. 2011;18(10):1099-104.
3. Lee J. An insight on the use of multiple logistic regression analysis to estimate association between risk factor and disease occurrence. Int J Epidemiol. 1986;15(1):22-9.
4. Kasza J, Wolfe R. Interpretation of commonly used statistical regression models. Respirology. 2014;19(1):14-21.
5. Kleinbaum DG, Klein M. Logistic Regression: A Self-Learning Text: Springer New York; 2010.
6. LaVeist TA. Beyond dummy variables and sample selection: what health services researchers ought to know about race as a variable. Health Serv Res. 1994;29(1):1-16.
7. Crown WH. There's a reason they call them dummy variables: a note on the use of structural equation techniques in comparative effectiveness research. Pharmacoeconomics. 2010;28(10):947-55.
8. Polissar L, Diehr P. Regression analysis in health services research: the use of dummy variables. Med Care. 1982;20(9):959-66.
9. Kleinbaum DG, Klein M, Pryor ER. Logistic regression : a self-learning text. 3rd ed. New York: Springer; 2010.
10. Hosmer DW, Lemeshow S. Applied Logistic Regression: Wiley; 2004.
11. Wondie AG, Zeleke AA, Yenus H, Tessema GA. Cesarean delivery among women who gave birth in Dessie town hospitals, Northeast Ethiopia. Plos One. 2019;14(5).
12. Agustin MZ, Peña EA. A basis approach to goodness-of-fit testing in recurrent event models. J Stat Plan Inference. 2005;133(2):285-303.
13. Gong G, Quante AS, Terry MB, Whittemore AS. Assessing the goodness of fit of personal risk models. Stat Med. 2014;33(18):3179-90.

14. Fagerland MW, Hosmer DW. A goodness-of-fit test for the proportional odds regression model. Stat Med. 2013;32(13):2235-49.
15. Mandel M, Betensky RA. Testing goodness of fit of a uniform truncation model. Biometrics. 2007;63(2):405-12.
16. Lele SR, Keim JL, Solymos P. ResourceSelection: Resource Selection (Probability) Functions for Use-Availability Data 2019 [Available from: https://CRAN.R-project.org/package=ResourceSelection.
17. Hosmer DW, Hosmer T, Le Cessie S, Lemeshow S. A comparison of goodness-of-fit tests for the logistic regression model. Stat Med. 1997;16(9):965-80.
18. Lemeshow S, Hosmer DW, Jr. A review of goodness of fit statistics for use in the development of logistic regression models. Am J Epidemiol. 1982;115(1):92-106.
19. Paul P, Pennell ML, Lemeshow S. Standardizing the power of the Hosmer-Lemeshow goodness of fit test in large data sets. Stat Med. 2013;32(1):67-80.
20. Fagerland MW, Hosmer DW. A goodness-of-fit test for the proportional odds regression model. Stat Med. 2013;32(13):2235-49.

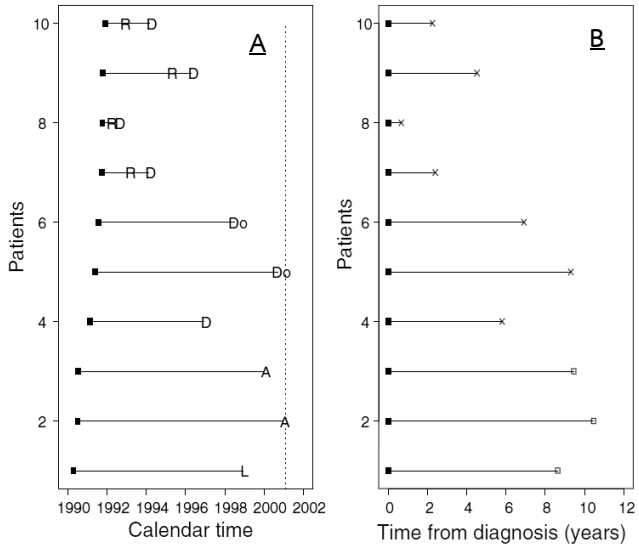
บทที่ 11

การวิเคราะห์อัตราการรอดชีพ

ในบทที่ 10 ได้ศึกษาถึงการวิเคราะห์ข้อมูลในกรณีที่ผลลัพธ์ของการศึกษาเป็นแบบจำแนกประเภทมีจำนวน 2 กลุ่ม เช่น เป็นโรคและไม่เป็นโรค โดยที่มีตัวแปรอิสระหลายตัว ใช้สถิติ Multiple logistic regression ในขณะเดียวกันถ้าการเกิดเหตุการณ์ที่เราสนใจนั้นขึ้นอยู่กับเวลาหรือสนใจเวลาตั้งแต่ติดตามอาสาสมัครงานวิจัยไปจนถึงเกิดเหตุการณ์หรือสิ้นสุดการศึกษา ลักษณะการเกิดเหตุการณ์ดังกล่าวเรียกว่า time-to event ลักษณะของข้อมูลดังกล่าวใช้การวิเคราะห์อัตราการรอดชีพหรือ Survival analysis[1] ในบทนี้จะได้เรียนรู้หลักการเบื้องต้นของการวิเคราะห์ดังกล่าว

11.1 หลักการการวิเคราะห์อัตราการรอดชีพ

ในการวิเคราะห์ข้อมูลที่มีผลลัพธ์ทางการศึกษาเป็นแบบ Time-to event นั้น ประกอบไปด้วยคำศัพท์ที่ใช้ในการวิเคราะห์ข้อมูลดังรูปที่ 11.1 เป็นตัวอย่างของการเป็นโรคมะเร็งรังไข่ที่มีการตรวจวินิจฉัยในระหว่างปี 1990 ถึงปี 1999 และติดตามผลถึงปี 2000[2]



รูปที่ 11.1 รูปตัวอย่างการเกิดเหตุการณ์โรคมะเร็งรังไข่
ที่มา: Clark TG, Bradburn MJ, Love SB, Altman DG. Survival analysis part I: basic concepts and first analyses. Br J Cancer. 2003;89(2):232-8. Figure 1 Converting calendar time in the ovarian cancer study to a survival analysis format.[3]

จากรูปที่ 11.1 (A) เส้น Dashed หมายถึงวันที่มีการติดตามผลการรักษาครั้งสุดท้าย (Last follow up หรือ Last visit), R หมายถึงการกลับมาเป็นโรคซ้ำ (Relapse), D หมายถึงการเสียชีวิตจากการเป็นโรคมะเร็งรังไข่ (Ovarian cancer), DO หมายถึงการเสียชีวิตจากโรคอื่นที่ไม่ใช่โรคมะเร็งรังไข่, A หมายถึงการมารับการติดตามผลครั้งสุดท้ายและยังมีชีวิตอยู่, L หมายถึงผู้ป่วยขาดการติดต่อ (Loss to follow-up) จากตัวอย่างมีการติดตามผลตั้งแต่เริ่มต้นที่ผู้ป่วยโรคมะเร็งรังไข่ได้รับการตรวจวินิจฉัย จนถึงสิ้นสุดการติดตามผลในปี 2000 จากข้อมูลพบว่ามีการกลับมาเป็นซ้ำ (R) จำนวน 4 ราย มีการเสียชีวิตจากการเป็นโรคมะเร็งรังไข่ (D) จำนวน 5 ราย, มีการเสียชีวิตจากการเป็นโรคอื่นที่ไม่ใช่เสียชีวิตจากโรคมะเร็งรังไข่ (DO) จำนวน 2 ราย, ยังคงมีชีวิตอยู่ ณ เวลาสุดท้ายที่ติดตามผลได้ (A) จำนวน 2 ราย และมีการขาดการติดต่อ (L) จำนวน 1 ราย

เมื่อทำการจัดรูปแบบข้อมูลให้อยู่ในรูปที่ 11.1(B)โดยที่เครื่องหมาย x หมายถึงมีการเสียชีวิต มีจำนวนทั้งหมด 7 ราย และ หมายถึง มีการขาดการติดต่อหรือไม่เกิดเหตุการณ์ที่สนใจมีจำนวนทั้งหมด 3 ราย ทั้งนี้การเกิดเหตุการณ์ที่สนใจ (Event) เช่น การเสียชีวิต (Death) การเป็นโรค (Disease) การกลับมาเป็นโรคซ้ำ (Relapse) หรือการหายจากโรค (Recovery) เรียกว่า failure และการไม่เกิดเหตุการณ์เรียกว่า Censoring จากรูปที่ 11.1(A) จะเห็นได้ว่าตอนที่เข้าสู่การศึกษา ณ จุดเวลาใดๆการเริ่มติดตามผล ณ เวลาที่เท่ากัน เหตุการณ์นี้เรียกว่า Left censored[4] ในขณะเดียวกัน ณ เวลาสิ้นสุดการศึกษาหรือการมาติดตามผลการรักษา ณ เวลาสุดท้าย ของแต่ละคนไม่เท่ากันและแต่ละคนอาจจะเกิดเหตุการณ์ก่อนสิ้นสุดแล้วติดตามผลไม่ได้ หรืออาจจะเกิดเหตุการณ์หลังสิ้นสุดเวลาการศึกษาเหตุการณ์นี้เรียกว่า Right censored[1] จากเหตุการณ์ทั้งสองเราจึงต้องวิเคราะห์ข้อมูลโดยใช้การวิเคราะห์แบบ Survival analysis ซึ่งในตำราเล่มนี้จะยกตัวอย่างการวิเคราะห์ข้อมูลประกอบไปด้วย การวิเคราะห์สองตัวแปรใช้สถิติชื่อ The logrank test[5], Kaplan-Meier curves[6,7] และการวิเคราะห์หลายตัวแปรใช้ตัวแบบ Cox proportional hazards[8-10]

ตัวอย่างที่ 11.1 การศึกษาของ Cho H และ คณะ (2016) ได้ทำการศึกษาปัจจัยที่มีผลต่อการต่อต้านไม่ยอมรับเซลล์ของผู้บริจาค (Graft failure) และการเสียชีวิตของผู้ป่วยสูงอายุที่ได้รับการปลูกถ่ายไต[11] เมื่อทำการจัดการข้อมูลแล้วจะได้ดัชนีข้อมูลดังตารางที่ 11.1

ตารางที่ 11.1 ดัชนีข้อมูลปัจจัยที่มีผลต่อการต่อต้านไม่ยอมรับเซลล์ผู้บริจาค

ชื่อตัวแปร	ความหมาย	ค่าของข้อมูล
id	identification	
rage	อายุของผู้รับ (Recipient age)	ตัวแปรต่อเนื่อง
rsex	เพศของผู้รับ (Recipient sex)	1=หญิง, 2=ชาย
hla_mis	Human leukocyte antigen mismatches	ตัวแปรต่อเนื่อง
aboi_no	การมีหมู่เลือดที่ไม่เข้ากัน (ABO incompatibility)	0=no, 1=yes
d_age	อายุของผู้ให้	ตัวแปรต่อเนื่อง
cmv_yn	Cytomegalovirus (Adverse event)	0=no, 1=yes
bkvcmv_yn	BK virus infection (Adverse event)	0=no, 1=yes
hcv_r	Hepatitis C ของผู้รับ	0=no, 1=yes
tdialymnt	Time on dialysis prior to transplantation (เดือน)	ตัวแปรต่อเนื่อง
fumont	Follow up time to death (เดือน)	ตัวแปรต่อเนื่อง
outcome	Death and censored	0=no, 1=yes

จากตารางที่ 11.1 เหตุการณ์ที่สนใจคืออัตราการเสียชีวิต (outcome) เวลาที่ติดตามผลตั้งแต่การได้รับการผ่าตัดไตจนถึงสิ้นสุดการศึกษาคือตัวแปร fumont โดยที่ failure คือ มีการเสียชีวิตและ Censored คือไม่มีการเสียชีวิต การวิเคราะห์ข้อมูลสามารถทำได้ดังนี้

ขั้นตอนที่ 1 การจัดการข้อมูล

```
setwd("D:\\epibios")
library(epiDisplay)
library(ggfortify)

#import data
>dtasurv<-read.csv("survival.csv")
>summ(dtasurv)

No. of observations = 225

   Var. name obs. mean   median  s.d.   min.   max.
1  id        225  113    113     65.1   1      225
2  rage      225  63.2   63      2.68   60     73
3  rsex      225  1.36   1       0.48   1      2
4  hla_mis   222  3.63   3       1.35   0      6
5  aboi_no   225  0.14   0       0.35   0      1
6  d_age     225  43.13  41      12.69  1      80
7  cmv_yn    225  0.2    0       0.4    0      1
8  bkvcmv_yn 225  0.09   0       0.29   0      1
9  hcv_r     225  0      0       0.07   0      1
10 tdialymnt225  30.6   9.1     44.97  0      275
11 fumont225  57.92  39.93   49.58  0.83   231.7
12 outcome   225  0.18   0       0.39   0      1
>dta<-na.omit(dtasurv)
>summ(dta)

No. of observations = 222

   Var. name obs. mean   median  s.d.   min.   max.
1  id        222  113.38 113.5   64.75  1      225
2  rage      222  63.2   63      2.66   60     73
3  rsex      222  1.37   1       0.48   1      2
4  hla_mis   222  3.63   3       1.35   0      6
5  aboi_no   222  0.14   0       0.34   0      1
6  d_age     222  43.17  41      12.73  1      80
7  cmv_yn    222  0.21   0       0.41   0      1
8  bkvcmv_yn 222  0.09   0       0.29   0      1
9  hcv_r     222  0      0       0.07   0      1
10 tdialymnt222  30.13  8.55    44.42  0      275
11 fumont222  58.02  39.87   49.8   0.83   231.7
12 outcome   222  0.18   0       0.39   0      1
```

รูปที่ 11.2 การนำเข้าข้อมูลและตรวจสอบข้อมูล

จากการสำรวจข้อมูลเบื้องต้นพบว่ามีจำนวนข้อมูลทั้งหมด 225 ราย อย่างไรก็ตามตัวแปร hla_mis มีทั้งหมด 222 ราย วิธีการจัดการข้อมูลต้องกลับไปดูว่าข้อมูลที่หายไปของตัวแปร hla_mis นั้นเป็นค่าที่เกิดการสูญหายจริงหรือไม่ ในเบื้องจะทำการลบข้อมูลออกไปเพื่อให้ได้ข้อมูลที่สมบูรณ์ โดยใช้คำสั่ง na.omit (ชื่อชุดข้อมูล)

ขั้นตอนที่ 2 การวิเคราะห์สองตัวแปร

การทดสอบความสัมพันธ์ระหว่างอายุของผู้รับบริจาค (Recipient age) กับการเสียชีวิตของผู้ป่วยที่ได้รับการผ่าตัดเปลี่ยนถ่ายไต (Renal transplant)

```
>summ(dta$rage)
 obs. mean   median  s.d.   min.   max.
222  63.203 63      2.664  60     73
> dta$rage65 <-ifelse(dta$rage>65,1,0)
> tabl(dta$rage65)
dta$rage65 :
        Frequency Percent Cum. percent
0             183    82.4         82.4
1              39    17.6        100.0
  Total       222   100.0        100.0
>survdiff(Surv(fumont, outcome) ~ rage65, data = dta)
Call:
survdiff(formula = Surv(fumont, outcome) ~ rage65, data = dta)

           N Observed Expected (O-E)^2/E (O-E)^2/V
rage65=0 183       35    34.17    0.0203     0.141
rage65=1  39        5     5.83    0.1188     0.141

Chisq= 0.1  on 1 degrees of freedom, p= 0.7
```

การทดสอบ logrank test

ค่า p-value

รูปที่ 11.3 ผลการวิเคราะห์ความสัมพันธ์ระหว่างอายุของผู้รับบริจาคกับการเสียชีวิตของผู้ป่วยที่ได้รับการผ่าตัดเปลี่ยนถ่ายไต

จากผลการวิเคราะห์ข้อมูลพบว่าอายุของผู้รับบริจาคไตไม่มีความสัมพันธ์กับการเสียชีวิตของผู้รับการผ่าตัดเปลี่ยนไต (p-value = 0.7)

การทดสอบความสัมพันธ์ระหว่างอายุของบริจาคไต (Recipient age) กับการเสียชีวิตของผู้ป่วยที่ได้รับการผ่าตัดเปลี่ยนถ่ายไต

```
>summ(dta$d_age)
 obs. mean   median  s.d.   min.   max.
222  43.167 41      12.726 1      80
> dta$d_age60 <-ifelse(dta$d_age>50,1,0)
> tab1(dta$d_age60)
dta$d_age60 :
        Frequency Percent Cum. percent
0             152    68.5         68.5
1              70    31.5        100.0
  Total       222   100.0        100.0
>survdiff(Surv(fumont, outcome) ~ d_age60, data = dta)
Call:
survdiff(formula = Surv(fumont, outcome) ~ d_age60, data = dta)

            N Observed Expected (O-E)^2/E (O-E)^2/V
d_age60=0 152       30    30.38   0.00479     0.021
d_age60=1  70       10     9.62   0.01512     0.021

Chisq= 0  on 1 degrees of freedom, p= 0.9
```

รูปที่ 11.4 ผลการวิเคราะห์ความสัมพันธ์ระหว่างอายุของบริจาคไตกับการเสียชีวิตของผู้ป่วยที่ได้รับการผ่าตัดเปลี่ยนถ่ายไต

จากผลการวิเคราะห์ข้อมูลพบว่าอายุของผู้บริจาคไตไม่มีความสัมพันธ์กับการเสียชีวิตของผู้รับการผ่าตัดเปลี่ยนไต (p-value = 0.9)

การทดสอบความสัมพันธ์ระหว่างเพศของผู้รับบริจาคไต (Recipient age) กับการเสียชีวิตของผู้ป่วยที่ได้รับการผ่าตัดเปลี่ยนถ่ายไต

```
>survdiff(Surv(fumont, outcome) ~ rsex, data = dta)
Call:
survdiff(formula = Surv(fumont, outcome) ~ rsex, data = dta)

         N Observed Expected (O-E)^2/E (O-E)^2/V
rsex=1 140       26     25.4    0.0151    0.0415
rsex=2  82       14     14.6    0.0263    0.0415

Chisq= 0  on 1 degrees of freedom, p= 0.8
```

รูปที่ 11.5 ผลการวิเคราะห์ความสัมพันธ์ระหว่างเพศของผู้รับบริจาคไตกับการเสียชีวิตของผู้ป่วยที่ได้รับการผ่าตัดเปลี่ยนถ่ายไต

จากผลการวิเคราะห์ข้อมูลพบว่าเพศของผู้รับบริจาคไตไม่มีความสัมพันธ์กับการเสียชีวิตของผู้รับการผ่าตัดเปลี่ยนไต (p-value = 0.8)

การทดสอบความสัมพันธ์ระหว่างการมีหมู่เลือดที่ไม่เข้ากัน (ABO incompatibility) กับการเสียชีวิตของผู้ป่วยที่ได้รับการผ่าตัดเปลี่ยนถ่ายไต

```
>survdiff(Surv(fumont, outcome) ~ aboi_no, data = dta)
Call:
survdiff(formula = Surv(fumont, outcome) ~ aboi_no, data = dta)

            N Observed Expected (O-E)^2/E (O-E)^2/V
aboi_no=0 192       34    37.55     0.336      5.79
aboi_no=1  30        6     2.45     5.149      5.79

Chisq= 5.8  on 1 degrees of freedom, p= 0.02
```

รูปที่ 11.6 ผลการวิเคราะห์ความสัมพันธ์ระหว่างการมีหมู่เลือดที่ไม่เข้ากันกับการเสียชีวิตของผู้ป่วยที่ได้รับการผ่าตัดเปลี่ยนถ่ายไต

จากผลการวิเคราะห์ข้อมูลพบว่าการมีหมู่เลือดที่ไม่เข้ากันมีความสัมพันธ์กับการเสียชีวิตของผู้รับการผ่าตัดเปลี่ยนไต (p-value = 0.02)

การทดสอบความสัมพันธ์ระหว่างการเกิด Cytomegalovirus กับการเสียชีวิตของผู้ป่วยที่ได้รับการผ่าตัดเปลี่ยนถ่ายไต

```
>survdiff(Surv(fumont, outcome) ~ cmv_yn, data = dta)
Call:
survdiff(formula = Surv(fumont, outcome) ~ cmv_yn, data = dta)

           N Observed Expected (O-E)^2/E (O-E)^2/V
cmv_yn=0 176       27     33.7      1.33      8.57
cmv_yn=1  46       13      6.3      7.12      8.57

Chisq= 8.6  on 1 degrees of freedom, p= 0.003
```

รูปที่ 11.7 ผลการวิเคราะห์ความสัมพันธ์ระหว่างการเกิด Cytomegalovirus กับการเสียชีวิตของผู้ป่วยที่ได้รับการผ่าตัดเปลี่ยนถ่ายไต

จากผลการวิเคราะห์ข้อมูลพบว่าการเกิด Cytomegalovirus มีความสัมพันธ์กับการเสียชีวิตของผู้รับการผ่าตัดเปลี่ยนไต (p-value = 0.003)

การทดสอบความสัมพันธ์ระหว่างการติดเชื้อไวรัส BK (BK virus infection) กับการเสียชีวิตของผู้ป่วยที่ได้รับการผ่าตัดเปลี่ยนถ่ายไต

```
>survdiff(Surv(fumont, outcome) ~ bkvcmv_yn, data = dta)
Call:
survdiff(formula = Surv(fumont, outcome) ~ bkvcmv_yn, data = dta)

               N Observed Expected (O-E)^2/E (O-E)^2/V
bkvcmv_yn=0 201       39    35.77     0.292      2.77
bkvcmv_yn=1  21        1     4.23     2.470      2.77

Chisq= 2.8  on 1 degrees of freedom, p= 0.1
```

รูปที่ 11.8 ผลการวิเคราะห์ความสัมพันธ์ระหว่างการติดเชื้อไวรัส BK กับการเสียชีวิตของผู้ป่วยที่ได้รับการผ่าตัดเปลี่ยนถ่ายไต

จากผลการวิเคราะห์ข้อมูลพบว่าการติดเชื้ออไวรัส BK ไม่มีความสัมพันธ์กับการเสียชีวิตของผู้รับการผ่าตัดเปลี่ยนไต (p-value = 0.1)

การทดสอบความสัมพันธ์ระหว่างการติดเชื้อ Hepatitis C ของผู้รับกับการเสียชีวิตของผู้ป่วยที่ได้รับการผ่าตัดเปลี่ยนถ่ายไต

```
>survdiff(Surv(fumont, outcome) ~ hcv_r, data = dta)
Call:
survdiff(formula = Surv(fumont, outcome) ~ hcv_r, data = dta)

           N Observed Expected (O-E)^2/E (O-E)^2/V
hcv_r=0 221       39   39.812    0.0166      3.54
hcv_r=1   1        1    0.188    3.5092      3.54

Chisq= 3.5  on 1 degrees of freedom, p= 0.06
```

รูปที่ 11.9 ผลการวิเคราะห์ความสัมพันธ์ระหว่างการติดเชื้อ Hepatitis C ของผู้รับกับการเสียชีวิตของผู้ป่วยที่ได้รับการผ่าตัดเปลี่ยนถ่ายไต

จากผลการวิเคราะห์ข้อมูลพบว่าการติดเชื้อ Hepatitis C ของผู้รับไม่มีความสัมพันธ์กับการเสียชีวิตของผู้รับการผ่าตัดเปลี่ยนไต (p-value = 0.06)

การทดสอบความสัมพันธ์ระหว่างจำนวน Human leukocyte antigen mismatches กับการเสียชีวิตของผู้ป่วยที่ได้รับการผ่าตัดเปลี่ยนถ่ายไต

```
>hla_miscox<- coxph(Surv(fumont,outcome) ~hla_mis,data = dta)
> summary(hla_miscox)
Call:
coxph(formula = Surv(fumont, outcome) ~ hla_mis, data = dta)

  n= 222, number of events= 40

coefexp(coef) se(coef)     z Pr(>|z|)
hla_mis 0.3150    1.3702   0.1271 2.479   0.0132 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

exp(coef) exp(-coef) lower .95 upper .95
hla_mis      1.37     0.7298     1.068     1.758

Concordance= 0.643  (se = 0.055 )
Likelihood ratio test= 6.28  on 1 df,   p=0.01
Wald test            = 6.15  on 1 df,   p=0.01
Score (logrank) test = 6.19  on 1 df,   p=0.01
```

ค่า p-value

รูปที่ 11.10 ผลการวิเคราะห์ความสัมพันธ์ระหว่างความสัมพันธ์ระหว่างจำนวน Human leukocyte antigen mismatches กับการเสียชีวิตของผู้ป่วยที่ได้รับการผ่าตัดเปลี่ยนถ่ายไตกับ

จากผลการวิเคราะห์ข้อมูลพบว่า จำนวน Human leukocyte antigen mismatches มีความสัมพันธ์กับการเสียชีวิตของผู้รับการผ่าตัดเปลี่ยนไต (p-value = 0.013) ในกรณีที่ตัวแปรอิสระเป็นตัวแปรต่อเนื่องการวิเคราะห์สองตัวแปรจะใช้วิธีของ Simple cox proportional hazard model[12] ใช้คำสั่งในโปรแกรม R คือ *coxph()*

การทดสอบความสัมพันธ์ระหว่างเวลาตั้งแต่ฟอกไตจนถึงการผ่าตัดเปลี่ยนถ่ายไต (Time on dialysis prior to transplantation) กับการเสียชีวิตของผู้ป่วยที่ได้รับการผ่าตัดเปลี่ยนถ่ายไต

```
>tdialymntcox<- coxph(Surv(fumont,outcome) ~tdialymnt,data = dta)
> summary(tdialymntcox)
Call:
coxph(formula = Surv(fumont, outcome) ~ tdialymnt, data = dta)

  n= 222, number of events= 40

coefexp(coef) se(coef)     z Pr(>|z|)
tdialymnt0.005824  1.005841 0.002981 1.954   0.0507 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

exp(coef) exp(-coef) lower .95 upper .95
tdialymnt     1.006     0.9942         1     1.012

Concordance= 0.671  (se = 0.042 )
Likelihood ratio test= 3.21  on 1 df,   p=0.07
Wald test            = 3.82  on 1 df,   p=0.05
Score (logrank) test = 3.9  on 1 df,   p=0.05
```

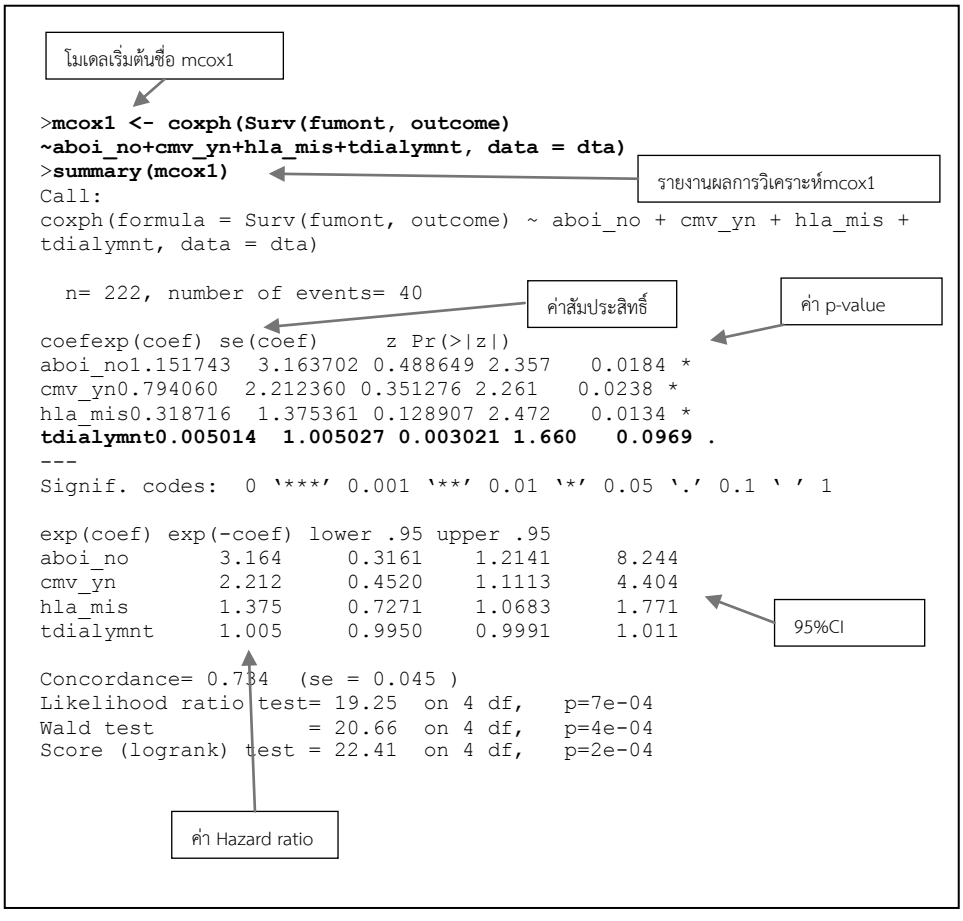
รูปที่ 11.11 ผลการวิเคราะห์ความสัมพันธ์ระหว่างเวลาตั้งแต่ฟอกไตจนถึงการผ่าตัดเปลี่ยนถ่ายไต กับการเสียชีวิตของผู้ป่วยที่ได้รับการผ่าตัดเปลี่ยนถ่ายไต

ผลการวิเคราะห์ข้อมูลพบว่าเวลาตั้งแต่ฟอกไตจนถึงการผ่าตัดเปลี่ยนถ่ายไต (Time on dialysis prior to transplantation) ไม่มีความสัมพันธ์กับการเสียชีวิตของผู้รับการผ่าตัดเปลี่ยนไต (p-value = 0.051)

จากการวิเคราะห์แบบสองตัวแปรพบว่าตัวแปรที่มีค่า p-value น้อยกว่า 0.05 มีจำนวน 3 ตัวแปร คือ aboi_no (ABO incompatibility), cmv_yn (Cytomegalovirus), และ hla_mis (Human leukocyte antigen mismatches) ส่วนตัวแปร tdialymnt (Time on dialysis prior to transplantation) มีค่า p-value ใกล้ 0.05 จึงจะพิจารณาเอาเข้าไปวิเคราะห์ในขั้นตอนหลายตัวแปรด้วย[13-16]

ขั้นตอนที่ 3 การวิเคราะห์หลายตัวแปร

การวิเคราะห์หลายตัวแปรจะใช้สถิติที่ชื่อว่า Cox proportionalhazards model[1] โดยใช้คำสั่งในโปรแกรม R คือ *coxph(model)* ในชุดคำสั่ง *survival* การวิเคราะห์ข้อมูลดังนี้



รูปที่ 11.12 ผลการวิเคราะห์ Cox proportional hazards model โมเดลที่ 1

ผลการวิเคราะห์ข้อมูลพบว่าตัวแปร tdialymnt มีค่า p-value เท่ากับ 0.097 ดังนั้นจะพิจารณาขจัดตัวแปรนี้ออกจากโมเดล จากนั้นทำการทดสอบ likelihood ratio[17,18] ระหว่างโมเดลโดยใช้คำสั่ง *lrtest(model2, model1)*

```
>mcox2 <- coxph(Surv(fumont, outcome) ~aboi_no+cmv_yn+hla_mis, data =
dta)
>lrtest(mcox2,mcox1)

Likelihood ratio test for Cox regression & conditional logistic
regression
Chi-squared 1 d.f.=  2.373838 , P value =  0.1233834
```

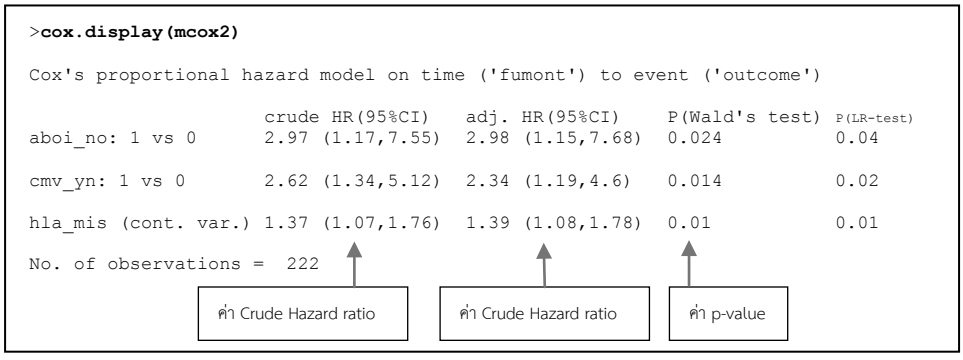
รูปที่ 11.13 ผลการวิเคราะห์ Cox proportional hazards model โมเดลที่ 2 และการทดสอบ likelihood ratio test

จากการทดสอบค่า likelihood ratio พบว่าได้ค่า p-value เท่ากับ 0.123 จึงสรุปได้ว่าตัวแปรที่ขจัดออก (tdialymnt) ไม่มีความสำคัญกับโมเดลจึงใช้ผลการวิเคราะห์ของโมเดลที่ 2 (mcox2)

```
>summary(mcox2)
Call:
coxph(formula = Surv(fumont, outcome) ~ aboi_no + cmv_yn + hla_mis,
    data = dta)

  n= 222, number of events= 40

coefexp(coef) se(coef)     z Pr(>|z|)
aboi_no 1.0910    2.9772   0.4836 2.256   0.0241 *
cmv_yn  0.8494    2.3381   0.3449 2.462   0.0138 *
hla_mis 0.3265    1.3861   0.1268 2.575   0.0100 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

exp(coef) exp(-coef) lower .95 upper .95
aboi_no     2.977     0.3359     1.154     7.682
cmv_yn      2.338     0.4277     1.189     4.597
hla_mis     1.386     0.7215     1.081     1.777

Concordance= 0.729  (se = 0.046 )
Likelihood ratio test= 16.88  on 3 df,   p=7e-04
Wald test            = 18.15  on 3 df,   p=4e-04
Score (logrank) test = 19.47  on 3 df,   p=2e-04
```

รูปที่ 11.14 ผลการวิเคราะห์ Cox proportional hazards model โมเดลที่ 2

ผลการวิเคราะห์ในรูปที่ 11.14 พบว่าตัวแปรที่อยู่ในโมเดลมีค่า p-value น้อยกว่า 0.05 ทั้งหมดจึงได้โมเดลสุดท้ายที่จะนำไปรายงานผลการศึกษา ในการรายงานผลจะใช้คำสั่ง *cox.display(model)* ซึ่งอยู่ในชุดคำสั่ง epiDisplay



รูปที่ 11.15 การรายงานค่า Hazard ration จาก Cox proportional hazards model โมเดลที่ 2

จากรูปที่ 11.15 ผลการวิเคราะห์ข้อมูลพบว่าหลังจากที่ควบคุมปัจจัยด้าน Human leukocyte antigen mismatches และการติดเชื้อ Cytomegalovirus พบว่า ผู้ป่วยที่มีความไม่เข้ากันของหมู่เลือด (ABO incompatibility) มีความเสี่ยงต่อการเสียชีวิตเป็น 2.98 (95% CI: 1.15,7.68) เท่าเมื่อเทียบกับผู้ที่มีหมู่เลือดเข้ากันได้ในขณะเดียวกันถ้าเราควบคุมปัจจัยด้าน Human leukocyte antigen mismatches และความไม่เข้ากันของหมู่เลือด (ABO incompatibility) พบว่า ผู้ป่วยที่มีการติดเชื้อ Cytomegalovirusมีความเสี่ยงต่อการเสียชีวิตเป็น 2.34 (95% CI: 1.19, 4.6) เท่าเมื่อเทียบกับผู้ที่ไม่มีการติดเชื้อ นอกจากนั้นแล้วถ้าเราควบคุมปัจจัยด้านการติดเชื้อ Cytomegalovirus และ ความไม่เข้ากันของหมู่เลือด (ABO incompatibility) พบว่า ทุกๆ ค่าที่เพิ่มขึ้น 1 หน่วยของ Human leukocyte antigen mismatches มีความเสี่ยงต่อการเสียชีวิตเป็น 1.39 (95% CI: 1.08, 1.78)

11.2 บทสรุป

การวิเคราะห์ข้อมูลอัตราการรอดชีพผลลัพธ์ของการศึกษามีลักษณะเป็นแบบ time to event ซึ่งผลลัพธ์ของการศึกษาแบ่งออกเป็นสองกลุ่มและมีระยะเวลาของการติดตามการเกิดผลลัพธ์ทางการศึกษาด้วย (Time-to event) ความเสี่ยงสัมพัทธ์ของการวิเคราะห์รายงานเป็น Hazard ratio (HR)

11.3 แบบฝึกหัด

1. ข้อแตกต่างระหว่าง Multiple logistic regression และ Cox proportional hazards model คืออะไร
2. ข้อแตกต่างของ Hazard ratio, Odds ratio และ Relative risk คืออะไร
3. กรณีที่อาสาสมัครในงานวิจัยไม่สามารถติดต่อได้ เรียกว่าอะไร
4. การวิเคราะห์แบบ Survival analysis เหมาะกับการศึกษาแบบใด
5. การวิเคราะห์ Survival analysis ในขั้นตอนวิเคราะห์สองตัวแปร ถ้าตัวแปรต้นหรือปัจจัยที่ต้องการศึกษาเป็นตัวแปรต่อเนื่อง จะใช้สถิติใด

11.4 เอกสารอ้างอิง

1. Kleinbaum DG, Klein M. Survival Analysis: A Self-Learning Text: Springer; 2005.
2. Clark TG, Stewart ME, Altman DG, Gabra H, Smyth JF. A prognostic model for ovarian cancer. Br J Cancer. 2001;85(7):944-52.
3. Clark TG, Bradburn MJ, Love SB, Altman DG. Survival analysis part I: basic concepts and first analyses. Br J Cancer. 2003;89(2):232-8.
4. Su YR, Wang JL. MODELING LEFT-TRUNCATED AND RIGHT-CENSORED SURVIVAL DATA WITH LONGITUDINAL COVARIATES. Ann Stat. 2012;40(3):1465-88.
5. Bland JM, Altman DG. The logrank test. Bmj. 2004;328(7447):1073.
6. Rich JT, Neely JG, Paniello RC, Voelker CC, Nussenbaum B, Wang EW. A practical guide to understanding Kaplan-Meier curves. Otolaryngol Head Neck Surg. 2010;143(3):331-6.
7. Ranstam J, Cook JA. Kaplan-Meier curve. Br J Surg. 2017;104(4):442.
8. Crichton N. Cox proportional hazards model. J Clin Nurs. 2002;11(6):723.
9. Tolosie K, Sharma MK. Application of cox proportional hazards model in case of tuberculosis patients in selected addis ababa health centres, ethiopia. Tuberc Res Treat. 2014;2014:536976.
10. Singh R, Mukhopadhyay K. Survival analysis in clinical trials: Basics and must know areas. Perspect Clin Res. 2011;2(4):145-8.
11. Cho H, Yu H, Shin E, Kim YH, Park SK, Jo MW. Risk Factors for Graft Failure and Death following Geriatric Renal Transplantation. PLoS One. 2016;11(4):e0153410.

12. George B, Seals S, Aban I. Survival analysis and regression models. J Nucl Cardiol. 2014;21(4):686-94.
13. Chowdhury MZI, Turin TC. Variable selection strategies and its importance in clinical prediction modelling. Fam Med Community Health. 2020;8(1):e000262.
14. Zhang Z. Variable selection with stepwise and best subset approaches. Ann Transl Med. 2016;4(7):136.
15. Heinze G, Wallisch C, Dunkler D. Variable selection - A review and recommendations for the practicing statistician. Biom J. 2018;60(3):431-49.
16. Sanchez-Pinto LN, Venable LR, Fahrenbach J, Churpek MM. Comparison of variable selection methods for clinical predictive modeling. Int J Med Inform. 2018;116:10-7.
17. Zhu H. Likelihood approaches for proportional likelihood ratio model with right-censored data. Stat Med. 2014;33(14):2467-79.
18. Deeks JJ, Altman DG. Diagnostic tests 4: likelihood ratios. Bmj. 2004;329(7458):168-9.

Index

ก

การกระจายของการเกิดโรค 1
การกลับมาเป็นโรคซ้ำ 146
การจัดสรรแบบสุ่ม 18
การจำแนกกลุ่มผิดประเภท 29
การทดลองเชิงสุ่มโดยมีกลุ่มเปรียบเทียบ 15
การทดลองแบบสองระยะไขว้กัน 15
การทดลองแบบสี่กลุ่มโซโลมอน 15, 21
การนำเข้าข้อมูล 51, 53
การฟิตโมเดล 108, 124, 142
การวัดแบบอัตราส่วน 40
การวัดผลก่อนการทดลอง 21
การวิเคราะห์อัตราการรอดชีพ 145
การศึกษาเชิงทดลอง 1, 15
การศึกษาเชิงสังเกตการณ์ 5
การศึกษาเชิงสุ่มแบบมีกลุ่มควบคุม 1
การศึกษาทางด้านระบาดวิทยาเชิงวิเคราะห์ 1
การศึกษาแบบตัดขวาง 5
การศึกษาแบบย้อนหลัง 5, 8
การศึกษาแบบสองระยะไขว้กัน 18
การศึกษาไปข้างหน้า 1, 10, 11
การสูญหายหรือติดต่ออาสาสมัครงานวิจัยไม่ได้ 29

ค

ความคลาดเคลื่อนเชิงระบบ 27
ความคลาดเคลื่อนแบบสุ่ม 27
ความตรงภายใน 18
ความถี่ของการเกิดโรค 1
ความเที่ยงตรงภายนอก 18
ความแปรปรวนของข้อมูล 106

ง

งานวิจัยทางคลินิก 18

ช

ชีวสถิติ 1, 2

ต

ตัวแปรกลุ่ม 40, 76, 81, 127
ตัวแปรกวน 20, 33, 92
ตัวแปรต่อเนื่อง 40, 41, 59, 76, 111, 147
ตัวแปรหุ่น 86, 130

ป

โปรแกรม R 57
โปรแกรม RStudio 47, 57

ม

มาตราวัดแบบนามบัญญัติ 39
มาตราวัดแบบมาตราอันดับ 40
มาตราอันตรภาค 40

ร

ระดับความสัมพันธ์ 81

ว

วิเคราะห์หลายตัวแปร 116

ส

สถิติเชิงอนุมาน 2, 39
สถิติพรรณนา 2, 39
สมการถดถอยเชิงเส้น 99
สมมติฐานทางสถิติ 41
สหสัมพันธ์ของเพียร์สัน 99, 124

อ

อคติ
    เกี่ยวกับการเลือกตัวอย่าง 27
    การเกิดอุบัติการณ์-ความชุก 27, 28
    การไม่ตอบรับ 27
    ตัวแปรกวน 27, 33
    อาสาสมัคร 27, 29
อุบัติการณ์ 7, 11

A

Adjusted odds ratio 34

attach 55

B

Backward elimination 108

Bias 13, 25, 27, 28, 35, 36

Box and Whisker Plots 73

C

Carry-over effect 20

Case-control study 1, 8, 10, 12, 31, 82, 90

Categorical data 40, 55, 81

Causation 10, 11

Chi-square test 81, 84, 85, 86, 87, 90, 91, 95

Clinical trial 18

Closed-end question 33

Cohort study 1, 5, 10, 12, 82

Confounding 27, 33, 35, 37, 92

Console 48, 49, 55

Continuous data 40, 55, 59

Cox proportional hazards regression 145

Cross-over design 15, 18, 23

Cross-sectional study 1, 5, 12, 82, 109

Crude odds ratio 34

D

data Frame 51

Descriptive statistics 2, 4, 39, 78

Determinants 1

Differential misclassification 29, 31

Distribution 1

Dummy variable 86, 128, 130

E

Environment 48, 49, 55, 58

epiDisplay 47, 57, 58, 86, 93, 98, 140, 142, 156

Experimental study design 5, 15

External validity 18

F

Factor 81, 86

failure 146, 147

Fisher's exact test 81, 86, 90, 91, 95, 97, 98

Forward Elimination 108

Friend controls 9

G

Games-Howell 72

ggplot2 47, 58, 102, 124

H

Homogeneity of residuals variance 109

I

Incidence-Prevalence Bias 27

Independent t-test 60, 62

Inferential statistics 2, 4, 39

Information bias 27, 29, 35, 37

Intention to-treat 29

Internal validity 18

Interval scale 40
Interviewer bias 31, 33

K
Kaplan-Meier curves 158
Kruskal Wallis Test 59, 73

L
Linearity 109
Long format 70
Loss to follow-up 7, 11, 18, 25, 27, 29, 36, 146

M
Mann-Whitney's U test 59, 63, 64, 76
Masking 33
Match case-control study 34
Matched case-control study 10
Maximum likelihood 128
Measurement bias 29
Misclassification 29, 31, 33, 37
multicolinearity 128
Multiple comparison 72
Multiple control per case 10
Multiple exposures 10
Multiple linear regression 34, 105, 107, 109, 116, 117, 118, 119, 120, 123, 124, 125
Multiple logistic regression 34, 127, 137, 139, 140, 141, 145, 158
Multiple outcomes 10

N
Neighborhoods controls 9
Nested case-control study 12
Neyman's Bias 13, 25
Nominal scale 39
Non-differential misclassification 29, 30, 31
Non-response Bias 25, 27
Normal Q-Q 109, 121

O
Observation bias 29
Odds Ratio 82
One-tailed 42
One-Way ANOVA 59
Ordinal scale 40
Overestimation 32, 33

P
Paired t-test 59, 64, 65, 66, 76
Person's correlation 99, 101, 123
Population controls 9
Power of analysis 7
Pretesting 21, 22, 23
Prevalence – incidence bias 7
Prospective cohort study 12

R
Random error 27
Randomization 18, 23
Randomized controlled trial 1, 15, 21, 23

Rare disease 10, 82
Ratio scale 40
Recall bias 10, 31
Relapse 146
Relative Risk 82
reshape 70
Residuals vs Fitted 109, 121
Residuals vs Leverage 109

S

Scale-Location 109, 121
Secondary outcome 18
Selection bias 24, 27, 35
Sequence 18
Simple logistic regression 89, 127
Solomon four group design 15, 21
Source 48, 55
Spearman Rank Correlation 103, 104
Statistical modeling 34
Stepwise 108
Stratified random sampling 34
Strength of association 81
Survival analysis 145, 146, 158, 159
Systematic error 27

T

The logrank test 158
Time-to event 145, 157
Two-tailed 42

U

Underestimation 32, 33
Unpaired t-test 59, 60, 61, 62, 63, 76

V

Values 46, 49, 51
Volunteer bias 27, 29, 36

W

Wide format 70
Wilcoxon signed rank test 59, 66, 67, 68, 78

เฉลย

เฉลยบทที่ 1

1. การศึกษาที่เกี่ยวข้องกับคนเพื่ออธิบายการกระจายของโรค (Distribution) และปัจจัยที่เกี่ยวข้องกับการเกิดโรคหรือปัจจัยที่เกี่ยวข้องกับการกระจายของการเกิดโรค (Determinants)
2. การศึกษาภาคตัดขวาง (Cross-sectional study) การศึกษาไปข้างหน้า (Cohort study) และการศึกษาย้อนหลัง (Case-control study)
3. ความถี่ของการเกิดโรค (Frequency) เกิดขึ้นเวลาไหน (Time) สถานที่ใด (Place) และเกิดขึ้นกับคนกลุ่มใด (Person)
4. ออกแบบงานวิจัย การคำนวณขนาดตัวอย่าง การสุ่มตัวอย่าง การเก็บข้อมูล การวิเคราะห์ข้อมูล การรายงานผลการวิเคราะห์ข้อมูลและการสรุปผล
5. สถิติเชิงอนุมาน

เฉลยบทที่ 2

1. Case-control study
2. Cross-sectional study และ Case-control study
3. Matched case-control study
4. Cross-sectional study
5. Cohort study

เฉลยบทที่ 3

1. เพื่อควบคุมปัจจัยกวนที่มีผลต่อการศึกษา
2. ขึ้นอยู่กับโรคที่ศึกษาและการเปลี่ยนแปลงของผลลัพธ์ทางการศึกษา
3. โรคเรื้อรัง
4. Half-life ที่นำมาทดลอง
5. เพื่อทดสอบ Pretesting effect

เฉลยบทที่ 4

1. ไม่อาจจำกำจัดให้หมดไปได้ แต่สามารถลดให้เหลือน้อยได้
2. ขั้นตอนการออกแบบงานวิจัยและการวิเคราะห์ข้อมูล
3. ทำคู่มือการเก็บข้อมูลให้เป็นมาตรฐานและอบรมทีมวิจัย
4. ใช้โมเดลทางสถิติ Multiple logistic regression
5. กำหนดเป็นช่วงเวลาที่สัมผัสปัจจัยเสี่ยงให้สั้นลงเช่น ใน 7 วันที่ผ่านมา โดยเฉลี่ยท่านได้สูบบุหรี่เป็นกี่มวน เป็นต้น

เฉลยบทที่ 5

1. ชีวสถิติ หมายถึง ข้อมูลทางสถิติหรือข้อมูลสรุปที่ได้จากการวิเคราะห์หรือประมวลผลข้อมูลในงานวิจัยทางด้านวิทยาศาสตร์สุขภาพ
2. สามารถทำได้ เช่น อายุ ข้อมูลเดิมเป็นแบบอัตราส่วน สามารถแบ่งเป็นกลุ่มหรือช่วงได้ เช่น 21-30, 31-40, 41-50, 51-60 เป็นต้น
3. เพราะการศึกษาบางชนิดไม่สามารถเก็บข้อมูลของประชากรแบบอนันท์ (Infinite population) ได้ และเป็นการประหยัด งบประมาณ กำลังคน และระยะเวลา
4. ไม่เหมาะสมเพราะค่าสถิติคำนวณจากกลุ่มตัวอย่างแล้วนำไปอ้างอิงสู่ประชากร
5. มาตราวัดแบบอันดับหรือระดับ (Ordinal scale)

เฉลยบทที่ 6

1. โปรแกรม R เป็นโปรแกรมหลักในการวิเคราะห์ข้อมูล และโปรแกรม RStudio เป็นโปรแกรมเสริมที่ช่วยให้การทำการเขียนคำสั่งวิเคราะห์ข้อมูลง่ายขึ้น
2. Source
3. setwd()
4. ใช้สำหรับการวิเคราะห์ข้อมูลทางระบาดวิทยา
5. cut()

เฉลยบทที่ 7

ข้อที่ 1

1. ทดสอบการแจกแจงของข้อมูล

จากการทดสอบการแจกแจงพบว่าข้อมูลมีการแจกแจงแบบปกติ ดังนั้นจึงใช้สถิติ Unpaired t-test

2. การทดสอบทางสถิติ

จากการทดสอบความเท่ากันของความแปรปรวนได้ค่าความน่าจะเป็น (p-value = 0.033) ดังนั้นในการทดสอบ Unpaired t-test นี้จะใช้การทดสอบแบบความแปรปรวนไม่เท่ากัน ซึ่งได้ค่าความน่าจะเป็นเท่ากับ 0.118 สามารถสรุปผลการศึกษาได้ว่า ค่าเฉลี่ยของ LDL-C ระหว่างกลุ่มที่รับประทานไขมันแบบปกติ กับ กลุ่มที่รับประทานไขมันต่ำ มีค่าแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ (p-value = 0.118)

ข้อที่ 2

1. ทดสอบการแจกแจง

ผลการทดสอบการแจกแจงพบว่าข้อมูลมีการแจกแจงปกติทั้ง 3 กลุ่ม (p-value > 0.05)

2. ทดสอบความแปรปรวน

ผลการทดสอบข้อมูลมีความแปรปรวนเท่ากัน (p-value = 0.637)

3. ทดสอบ ANOVA

ผลการทดสอบพบว่าค่าเฉลี่ยของคะแนนพฤติกรรมหลังการได้รับโปรแกรมทั้ง 3 กลุ่มมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ (p-value = 0.001)

4. การเปรียบเทียบรายคู่

การวิเคราะห์การวิเคราะห์ข้อมูลพบว่ากลุ่มที่ อ้วน-ปกติ และ น้ำหนักต่ำกว่าเกณฑ์-คนปกติ มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ สรุปผลการศึกษา จากการทดสอบค่าเฉลี่ยของคะแนนคุณภาพชีวิตทั้งสามกลุ่มมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ (p-value = 0.001) ทั้งนี้เมื่อทำการทดสอบรายคู่แล้วพบว่า กลุ่มอ้วน กับ กลุ่มปกติ มีค่าแตกต่างกัน 8.3 (95% CI: 2.7, 13.9) และ กลุ่มน้ำหนักต่ำกว่าเกณฑ์ กับ กลุ่มคนปกติมีค่าแตกต่างกัน 7.9 (95% CI: 2.3, 13.5)

เฉลยบทที่ 8

ข้อที่ 1. จากโจทย์สามารถเขียนตารางการณ์จรได้ดังนี้

OC	MI	
	Yes	No
OC	13	4,987
Non-OC	7	9,993
Total	20	14,980

จากนั้นทำการวิเคราะห์ข้อมูลโดยใช้โปรแกรม R

จากการวิเคราะห์ข้อมูลพบว่าการใช้ OC มีความสัมพันธ์กับการเกิด MI อย่างมีนัยสำคัญทางสถิติ (p-value = 0.003) โดยที่ผู้ที่ใช้ OC มีโอกาสเสี่ยงต่อการเกิด MI เป็น 3.72 เท่าเมื่อเทียบกับผู้ที่ไม่ได้ใช้ OC (95% CI: 1.48, 9.33)

ข้อที่ 2.จากโจทย์สร้างเป็นตารางการณ์จรได้ดังนี้

Cause of death	Diet	
	High salt	Low salt
Non CVD	2	23
CVD	5	30
Total	7	53

ทำการวิเคราะห์ข้อมูลโดยใช้โปรแกรม R

ผลการทดสอบทางสถิติโดยใช้ Chi-square test พบว่า การประมาณค่าดังกล่าวอาจไม่เหมาะสมเพราะว่า มีค่า expected value น้อยกว่า 5 มากกว่า 20% ของจำนวนเซลล์ทั้งหมด (มีคำเตือนจากโปรแกรมโดยมีข้อความดังนี้: Warning message: In chisq.test(ctbl2) : Chi-squared approximation may be incorrect) ดังนั้นจึงใช้สถิติ Fisher's exact test โดยได้ค่าความน่าจะเป็นเท่ากับ 0.688 สรุปได้ว่าการรับประทานอาหารเค็มไม่มีความสัมพันธ์กับการเกิดโรค CVD (p-value = 0.688)

ข้อที่ 3. จากข้อมูลดังกล่าวใช้โปรแกรม R และใช้ชุดคำสั่ง epiDIsplay เพื่อที่สามารถวิเคราะห์ข้อมูลโดยใช้ mhor ได้

ผลการวิเคราะห์พบว่า การสูบบุหรี่มีลักษณะเป็นตัวแปรกวน (p-value = 0.071) แปลผล หลังจากควบคุมปัจจัยด้านการสูบบุหรี่แล้วพบว่า การได้รับควันบุหรี่มือสองมีความสัมพันธ์กับการเกิดโรคมะเร็งปอดเป็น 1.63 เท่าเมื่อเทียบกับคนที่ไม่ได้รับควันบุหรี่มือสอง (95% CI: 1.26, 2.09)

เฉลยบทที่ 9

1. Linear regression
2. สามารถควบคุมตัวแปรกวนได้หลายตัวแปร
3. One-way ANOVA
4. -1 ถึง 1
5. 3 วิธีคือ Backward, Forward, และ Stepwise

เฉลยบทที่ 10

1. ได้ไม่จำกัด ทั้งนี้ขึ้นอยู่กับขนาดตัวอย่าง
2. อาจทำการปรับค่าของข้อมูลเช่นข้อมูลตัวเนื่องอาจจะทำการจัดกลุ่ม แล้วนำเข้าในโมเดล หรือจากตัวแปรกลุ่มมีจำนวนหลายกลุ่มอาจพิจารณายุบรวมกลุ่มให้น้อยลง แล้วทำการฟิตโมเดลใหม่
3. Multiple logistic regression
4. tableStack()
5. นำตัวแปรนั้นกลับเข้าไปในโมเดล

เฉลยบทที่ 11

1. Cox proportional hazards model มีการวัดผลลัพธ์ทางการศึกษาที่คำนึงถึงเวลาแต่ Multiple logistic regression model ไม่คำนึงถึงเวลา (Time-to event)
2. Hazard ratio ได้จากการศึกษาแบบ Prospective study โดยใช้การวิเคราะห์แบบ Survival analysis, Odds ratio ได้จากรูปแบบการศึกษาแบบ Case-control, Cross sectional study หรือ Cohort study ในกรณีโรคเกิดขึ้นน้อย (Rare event) และ Relative risk ได้มาจากการคำนวณในรูปแบบการศึกษาแบบ Cohort study
3. Loss to follow-up
4. Censoring เกิดขึ้น
5. Simple cox proportional hazard model