

สถิติในเวชปฏิบัติ

รศ.นพ.ชำแก้ว หวานวารี

วัตถุประสงค์ของวิชาสถิติในเวชปฏิบัติของรายวิชา RAID 511

1. ระบุความหมายของศัพท์ทางสถิติตามเกณฑ์มาตรฐานแพทยสภา
2. เข้าใจวิธีการทางสถิติที่ใช้ในเวชปฏิบัติ

1. Central tendency (การวัดแนวโน้มเข้าสู่ส่วนกลาง)

เป็นการบอกค่ากลางของชุดข้อมูลว่าเป็นเท่าใด ซึ่งตัวชี้วัดที่นิยมได้แก่ mean, median และ mode

2. Variability (ความแปรปรวน) บางครั้งเรียกว่า spread หรือ dispersion

เป็นการบอกว่า แต่ละค่าของชุดข้อมูลแตกต่างกันมากน้อยเท่าใด ซึ่งตัวชี้วัดที่นิยมได้แก่

2.1 Variance มีค่าเท่ากับ

$$\sigma^2 = \frac{\sum(X - \mu)^2}{N}$$

2.2 Standard deviation (S.D.) มีค่าเท่ากับ รากที่สองของ variance

2.3 Range มีค่าเท่ากับค่า maximum ลบด้วยค่า minimum

2.4 Interquartile มีค่าเท่ากับค่า percentile ที่ 75 ลบด้วยค่า percentile ที่ 25

3. Probability (ความน่าจะเป็น)

3.1 Probability ของการเกิดเหตุการณ์ใดเหตุการณ์หนึ่งเป็นดังนี้

“probability of A or B = probability of A + probability of B – probability of both A and B”

ถ้าเหตุการณ์ A กับเหตุการณ์ B เกิดร่วมกันไม่ได้ ค่า probability of both A and B จะเท่ากับศูนย์

3.2 Probability ของการเกิดเหตุการณ์ร่วมเป็นดังนี้

“probability of A and B = probability of A x probability of B” (ถ้า A และ B เกิดอิสระต่อกัน)

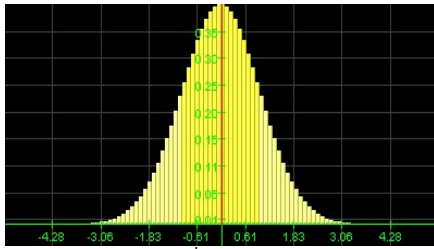
คำถาม คู่สมรสคูหนึ่งที่ไม่มีอาการผิดปกติแต่อย่างใด มีลูกคนแรกเป็น cystic fibrosis ซึ่งเป็น autosomal recessive ถ้าคู่สมรสคูนี้มีลูกคนที่สอง โอกาสที่ลูกคนหลังจะเป็นโรคนี้อีกเท่าใด

4. Distribution (การกระจาย) การกระจายของข้อมูลมีหลายแบบ ที่สำคัญได้แก่ normal distribution เป็นการกระจายของ continuous variables แบบโค้งระฆังคว่ำ ซึ่งค่า mean, median, mode จะเท่ากัน (ดูรูปที่ 1)

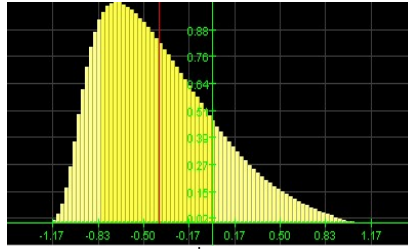
คำถาม กรณีต่อไปนี้ ค่า mean, median, mode ค่าใดจะมากกว่า

ก. หาก plot curve แล้วพบว่าเบ้ (skewed) ไปทางขวา (ดูรูปที่ 2).....

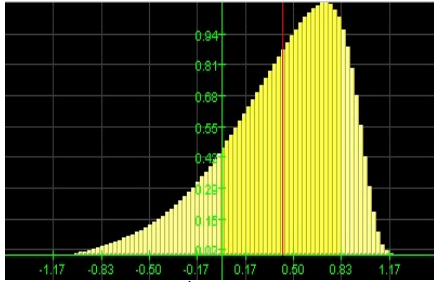
ข. หาก plot curve แล้วพบว่าเบ้ (skewed) ไปทางซ้าย (ดูรูปที่ 3)



รูปที่ 1



รูปที่ 2



รูปที่ 3

รูปที่ 1 Normal (Gaussian) distribution
 รูปที่ 2 Skewed to the right distribution (positively skewed)
 รูปที่ 3 Skewed to the left distribution (negatively skewed)

5. Scale of measurement (ชนิดตัวแปร)

- 1.1 Nominal : บอกลักษณะ เช่น เพศ blood group การเป็นหรือไม่เป็นโรค
- 1.2 Ordinal : บอกลำดับ เช่น severity (mild, moderate, severe) ระดับการศึกษา
- 1.3 Ratio : บอกปริมาณ มีศูนย์ที่แท้จริง (ศูนย์คือไม่มีปริมาณ) เช่น น้ำหนัก ส่วนสูง Hb FBS
- 1.4 Interval : : บอกปริมาณ แต่มีศูนย์ที่ไม่แท้จริง เช่น อุณหภูมิองศาเซลเซียส

Nominal และ ordinal variables จัดให้อยู่ในกลุ่ม discrete หรือ categorical variables
 Ratio และ interval variables จัดให้อยู่ในกลุ่ม continuous variables

คำถาม ตัวแปรต่อไปนี้เป็นชนิดใด (เลือก 1.1 – 1.4)

- ระดับไขมันในเลือด การควบคุมโรคเบาหวานได้ / ไม่ได้
- สาเหตุการตาย สูบบุหรี่ (≤10 มวน/วัน , 11-20 มวน/วัน , >20มวน/วัน)

6. Statistical inference (การสรุปค่าในประชากรจากค่าในกลุ่มตัวอย่างที่สุ่มมาศึกษา)

ในการศึกษาหรือวิจัยย่อมไม่สามารถทำในประชากรทั้งหมดได้ เพราะสิ้นเปลืองทรัพยากรมาก จึงมักทำการสุ่มตัวอย่างมาศึกษา เมื่อได้ค่าในกลุ่มตัวอย่าง (sample estimate) จึงต้องใช้วิธีการสรุปผลว่าค่าในประชากร (population estimate) น่าจะเป็นเท่าไร วิธีการดังกล่าวคือ inferential statistics

7. Hypothesis generation and testing (การสร้างและทดสอบสมมติฐาน)

การศึกษาหรือวิจัยในทางการแพทย์จะมีการตั้งสมมติฐานเสมอ โดยทั่วไปจะตั้งไว้ 2 สมมติฐานดังนี้

-Null hypothesis คือสมมติฐานของความไม่แตกต่าง หรือตั้งไว้ว่า ไม่มีความแตกต่างระหว่างกลุ่ม

-Alternative hypothesis คือสมมติฐานทางเลือก หรือตั้งไว้ว่า มีความแตกต่างระหว่างกลุ่ม

เมื่อตั้งสมมติฐานแล้ว จึงไปเก็บข้อมูลกลุ่มตัวอย่าง นำผลไปวิเคราะห์ทางสถิติเพื่อทดสอบสมมติฐานว่าจะรับหรือปฏิเสธสมมติฐาน ความหมายก็คือ เนื่องจากผลที่ได้จากการศึกษาเป็นผลในตัวอย่าง (sample) หากพบว่ามีความแตกต่างระหว่างกลุ่มก็ยังไม่ทราบว่าเป็นประชากร (population) ที่สุ่มตัวอย่างมา จะมีความแตกต่างระหว่างกลุ่มหรือไม่

8. Test statistics (ค่าสถิติ)

เป็นค่าสถิติที่คำนวณได้จากกลุ่มตัวอย่าง เช่น ค่า mean, median, mode, ค่า t จากการทำ t-test, ค่า chi-square จากการทำ chi-square test เป็นต้น ซึ่งค่าเหล่านี้จะนำไปสรุปผลในประชากรต่อไป

9. Statistical significance (นัยสำคัญทางสถิติ)

เมื่อทำการทดสอบสมมติฐาน จะคำนวณได้ค่าที่บอกว่ามีนัยสำคัญทางสถิติหรือไม่ จากคำกล่าวที่ว่า “การเกิดเหตุการณ์ที่มีโอกาสเกิดขึ้นได้ยาก ถือว่ามีนัยสำคัญทางสถิติ (occurrence of rare event is significant)” โดยทั่วไปมักกำหนดว่า หากเหตุการณ์นั้นเกิดน้อยกว่าร้อยละ 5 ถือว่ามีนัยสำคัญทางสถิติ ในภาษาทางสถิติคือ $p\text{-value} < .05$ ($p = \text{probability value}$)

ในโปรแกรมสถิติทั่วไป เมื่อใส่ข้อมูลดิบเข้าไป เครื่องคอมพิวเตอร์จะคิด p-value ให้ ซึ่งสามารถแปลผลได้ว่ามีนัยสำคัญทางสถิติหรือไม่ เช่น ถ้ากำหนดนัยสำคัญทางสถิติไว้ที่ .05 เมื่อเครื่องคำนวณได้ค่า p-value เป็น .02 ก็หมายความว่ามีความนัยสำคัญทางสถิติ หากได้ค่าเท่ากับ .05 หรือมากกว่าจะถือว่าไม่มีนัยสำคัญทางสถิติ อย่างไรก็ตามอาจกำหนดนัยสำคัญที่ระดับอื่น ๆ ได้ เช่น .01 เป็นต้น

วิธีการทดสอบสมมติฐานที่ใช้บ่อยในเวชปฏิบัติได้แก่ t-test, chi-square และ correlation

t-test เป็นการทดสอบว่าค่าเฉลี่ยระหว่างกลุ่ม 2 กลุ่มแตกต่างกันหรือไม่

ตัวอย่าง : การศึกษาเปรียบเทียบความสามารถของยาลดน้ำหนักรักษา 2 ชนิดในการลดน้ำหนักให้ผู้ป่วย

Null hypothesis : ยาทั้ง 2 ชนิดมีความสามารถในการลดน้ำหนักไม่แตกต่างกัน

เมื่อเก็บข้อมูลโดยชั่งน้ำหนักของผู้ป่วย 2 กลุ่ม (ตามชนิดของยา) ซึ่งเป็น continuous จึงใช้ t-test ในการทดสอบสมมติฐาน เช่น ผลปรากฏว่า $p\text{-value} = 0.02$ แปลผลว่ามีนัยสำคัญทางสถิติ (หากกำหนดระดับนัยสำคัญทางสถิติไว้ก่อนที่ระดับ .05) นั่นคือ ปฏิเสธ null hypothesis หรือ ยาทั้ง 2 ตัวมีผลแตกต่างกัน

Chi-square test เป็นการทดสอบว่ามีความสัมพันธ์ระหว่างกลุ่มหรือไม่โดยตัวแปรต้องเป็น discrete
ตัวอย่าง : การศึกษาเปรียบเทียบอัตราการหายจากโรคด้วยการรักษาด้วยยาใหม่เปรียบเทียบกับยาเก่า

	หาย	ไม่หาย
ยาใหม่	80	20
ยาเก่า	60	40

Chi-square = 9.52 p-value = 0.002
แปลผลว่า ผลการรักษาด้วยยาใหม่แตกต่างจากยาเก่า
ส่วนขนาดความแตกต่าง = %

Correlation เป็นการศึกษาความสัมพันธ์ระหว่างตัวแปรที่เป็น continuous ทั้งคู่

ตัวอย่าง : การศึกษาในผู้ป่วยเบาหวานเพื่อดูว่า FBS สัมพันธ์กับ HbA_{1c} หรือไม่
เนื่องจากทั้ง 2 ตัวแปรเป็น continuous จึงใช้ correlation ซึ่งจะให้ค่าสถิติ r (correlation coefficient) ค่า r จะมีค่าได้ระหว่าง -1 ถึง +1 ถ้าค่า r สูงใกล้ 1 แสดงว่าสัมพันธ์กันดี แต่ถ้าเป็นลบ แสดงว่าสัมพันธ์ผกผันกัน และต้องดูที่ p-value ด้วยว่ามีนัยสำคัญทางสถิติหรือไม่

11. Confidence interval (ช่วงแห่งความเชื่อมั่น)

ในการศึกษาหรือวิจัยเรามักจะทำในกลุ่มตัวอย่าง ดังนั้น ค่าที่ได้จึงเป็นค่าในกลุ่มตัวอย่าง แต่เราอยากรู้ค่านั้นในประชากรที่สุ่มตัวอย่างมา จึงต้องหาค่าดังกล่าวซึ่งค่านี้นี้ก็คือ confidence interval

ค่าในกลุ่มตัวอย่างเรียกว่า *point estimate* ค่าในประชากรเรียกว่า *population estimate*

ตัวอย่าง : อัตราชุกของโรคเบาหวานในกลุ่มตัวอย่าง = 10% (สำรวจ 1000 คน พบ 100 คน) คำนวณค่า 95 C.I. ได้เท่ากับ 8.2% – 12.1% แปลผลว่า อัตราชุกของโรคเบาหวานในประชากรที่ถูกสุ่มตัวอย่างมาน่าจะมีค่าอยู่ระหว่าง 8.2% ถึง 12.1% ด้วยความเชื่อมั่น 95%

การแปลผล C.I. ของการศึกษาอื่นๆ ไม่ว่าจะ เป็น relative risk, odds ratio, correlation coefficient, RRR, ARR, NNT, etc ก็เป็นไปในทำนองเดียวกัน

การศึกษาที่หาค่า relative เช่น R.R., O.R. ถ้า 95% C.I. คร่อม 1 แสดงว่า p-value \geq .05 (non-significant)

การศึกษาที่หาค่า relative เช่น R.R., O.R. ถ้า 95% C.I. ไม่คร่อม 1 แสดงว่า p-value < .05 (significant)

การศึกษาที่หาค่า absolute เช่น ARR., difference ถ้า 95% C.I. คร่อม 0 แสดงว่า p-value \geq .05

การศึกษาที่หาค่า absolute เช่น ARR., difference ถ้า 95% C.I. ไม่คร่อม 0 แสดงว่า p-value < .05

12. Sample size and Power

จำนวนตัวอย่างมีความสำคัญในการศึกษาวิจัยทางการแพทย์ หากจำนวนตัวอย่างมากก็จะเป็นตัวแทนที่ดีของประชากร (ถ้าสุ่มตัวอย่างอย่างไม่อคติ) เมื่อจำนวนตัวอย่างน้อย อาจไม่มี power พอที่จะตรวจสอบความแตกต่างระหว่างกลุ่ม โดยทั่วไป มักจะกำหนดให้มี power อย่างน้อย 80% จึงจะถือว่าการศึกษานั้นมีอำนาจเพียงพอในการตรวจสอบความแตกต่าง หากคำนวณ power แล้วพบว่ามีความน้อยกว่า 80% จะถือว่า power ไม่เพียงพอ เวลาที่มี negative study (สรุปผลว่าไม่มีความแตกต่าง) คำถามแรกที่เกิดขึ้นคือมี power เพียงพอหรือไม่ ซึ่ง power จะขึ้นอยู่กับ sample size เป็นหลัก หากจำนวนตัวอย่างมาก ย่อมมีความสามารถในการตรวจสอบความแตกต่างได้ดีขึ้น

การคำนวณ sample size และ power สามารถคำนวณโดยใช้ software ทางสถิติ

13. Type I (alpha) and Type II (beta) errors

ผลการศึกษาวิจัยย่อมมีความผิดพลาดได้ ความผิดพลาดมี 2 แบบได้แก่

7.1 สรุปผลว่ามีความแตกต่างในขณะที่ความจริงไม่มี เรียกว่า เกิด Type I (alpha) error

7.2 สรุปผลว่าไม่มีความแตกต่างในขณะที่ความจริงมี เรียกว่า เกิด Type II (beta) error

ก่อนการศึกษาวิจัย มักจะกำหนด alpha error ไว้ที่ .05 และกำหนด beta error ไว้ที่ .2 นั่นคือยอมรับความผิดพลาดแบบ Type I ร้อยละ 5 และยอมรับความผิดพลาดแบบ Type II ร้อยละ 20

Power จะมีค่าเท่ากับ $1 - \text{beta error}$ หมายความว่า เมื่อกำหนด beta error ไว้ที่ .2 แสดงว่ามี power เท่ากับ .8 หรือร้อยละ 80 นั่นเอง (งานวิจัยโดยทั่วไปควรกำหนด power ไว้ไม่ต่ำกว่าร้อยละ 80)

Reference : David Coggon. Statistics in Clinical practice. BMJ Books 2003 (Commed library)