การทดสอบสำหรับตัวอย่าง 2 ชุด ที่มีความสัมพันธ์กัน

208348 : สถิตินอนพาราเมตริก

โดย ...ผศ. ดร. สุคนธ์ ประสิทธิ์วัฒนเสรี
ภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยเชียงใหม่

Outline

- Two related samples
- การทดสอบโดยใช้เครื่องหมาย (Sign test)
- การทดสอบการเปลี่ยนแปลงอย่างมีนัยสำคัญของแมคนีมาร์ (McNemar's test)
- การทดสอบอันดับที่มีเครื่องหมายกำกับของวิลคอกซอน (Wilcoxon signed-rank test)
- ช่วงความเชื่อมั่นของผลต่างมัธยฐาน
- การทดสอบการสุ่มของฟีชเชอร์ (Fisher randomization test)
- การทดสอบสัมประสิทธิ์สหสัมพันธ์แบบอันดับ
- Kendall rank correlation coefficient

2

ปัญหาวิจัย : เปรียบเทียบความพึงพอใจยาสระผม A กับ B

- วิธีการศึกษา 1
  - สุ่มตัวอย่างมา 2 กลุ่มอย่างอิสระ
  - สุ่มเลือกชนิดยาสระผมให้แต่ละกลุ่มทดลองใช้
  - วัดความพึงพอใจของแต่ละกลุ่ม

ความพึงพอใจ	
ยาสระผม A	ยาสระผม B
$x_1, x_2, \ldots, x_{n1}$	$y_1, y_2, \ldots, y_{n2}$

ปัญหา : หาก กลุ่มที่ 1 เป็นผู้ชายทั้งหมด
กลุ่มที่ 2 เป็นผู้หญิงทั้งหมด
ความน่าเชื่อถือของผลการศึกษา ?????

- วิธีการศึกษา 2
  - สุ่มตัวอย่างมา 1 กลุ่ม
  - ให้แต่ละคนทดลองยาสระผมทั้ง 2 ชนิด (โดยสุ่มชนิดก่อน-หลัง)
  - วัดความพึงพอใจของแต่ละคน

	คนที่	1 2 … n
ความพึงพอใจ	A	x_1 x_2 … x_n
	B	y_1 y_2 … y_n

ผลการวิเคราะห์จะเป็นการเปรียบเทียบความพึงพอใจของผู้บริโภคที่มีต่อยาสระผมชนิด A กับชนิด B

3

ตัวอย่าง 2 ชุดที่มีความสัมพันธ์กัน (2 related samples)

- 2 วิธีในการจัดตัวอย่างที่มีความสัมพันธ์กัน
  - Repeated measures design
    ใช้ตัวอย่างเดียววัดค่าหลายครั้ง
  - Matched pair samples
    เลือกตัวอย่างมา 1 คู่ที่มีลักษณะคล้ายคลึงกัน

4

ตัวอย่าง 2 ชุดที่มีความสัมพันธ์กัน (2 related samples)

- ตัวอย่าง Matched pair samples
  - เปรียบเทียบความแตกต่างระหว่างวิธีการสอน 2 วิธี ทำการเลือกนศ.มาเป็นคู่ ๆ ที่มีลักษณะใกล้เคียงกัน (อายุ, ชั้นปี, IQ, ...) เพื่อทดลองสอน

คู่ที่	1	2	3	...	n
วิธีการสอน 1	x_1	x_2	x_3	...	x_n
วิธีการสอน 2	y_1	y_2	y_3	...	y_n

5

ตัวอย่าง 2 ชุดที่มีความสัมพันธ์กัน (2 related samples)

- ตัวอย่าง Repeated measures design
  - เปรียบเทียบน้ำหนักก่อนและหลังใช้ยาลดความอ้วน

คนที่	1	2	3	...	n
ก่อนใช้ยา	x_1	x_2	x_3	...	x_n
หลังใช้ยา	y_1	y_2	y_3	...	y_n

6

สถิติพาราเมตริกทดสอบความแตกต่าง

- 2 samples independent
  - z test
  - t test
- 2 related samples
  - Paired t test

    ข้อตกลงเบื้องต้น (assumption)
    - ตัวแปรที่ศึกษาต้องเป็นตัวแปรปริมาณ
    - ตัวแปรต้องมีการแจกแจงปกติ

เมื่อลักษณะข้อมูลที่ศึกษาไม่เป็นไปตาม assumption ควรใช้ **สถิตินอนพาราเมตริก**เพื่อทดสอบความแตกต่าง

7

การทดสอบโดยใช้เครื่องหมาย (sign test)

- จุดเด่น
  - เป็นการทดสอบที่ง่ายและสะดวก
  - ใช้กับข้อมูลในมาตรวัดเรียงลำดับ อันตรภาค และอัตราส่วน
  - ใช้เปรียบเทียบความแตกต่างระหว่างค่ามัธยฐาน
- จุดด้อย
  - มีกำลังในการทดสอบไม่สูงเมื่อเทียบกับวิธีการทดสอบอื่น ๆ

8

Sign Test

- Data
  - 2 related samples (X_i, Y_i)

Pairs	1	2	...	n′
X	X_1	X_2	...	$X_{n'}$
Y	Y_1	Y_2	...	$Y_{n'}$

  - การเปรียบเทียบข้อมูล
    - ถ้า $X_i > Y_i$ ให้เครื่องหมาย " + "
    - ถ้า $X_i < Y_i$ ให้เครื่องหมาย " - "
    - ถ้า $X_i = Y_i$ (tied) ให้เครื่องหมาย " 0 " (ไม่นำมาใช้วิเคราะห์ข้อมูล)

9

Sign Test

- Assumptions
  - ตัวแปรแต่ละคู่ (X_1, Y_1), (X_2, Y_2), ..., ($X_{n'}$, $Y_{n'}$) เป็นอิสระกัน
  - ค่าตัวแปร X และ Y ต้องอยู่ในมาตรเรียงลำดับ
  - ลักษณะของ (X_i, Y_i) เป็นแบบนัยภายใน (internally consistent) นั่นคือ
    - หากกำหนดให้
      โอกาสที่ค่า X มากกว่าค่า Y จะสูงกว่าโอกาสที่ค่า X น้อยกว่าค่า Y
      เขียนแทนด้วย $P(+) > P(-)$
      ทุกคู่ระหว่าง (X_i, Y_i) จะอยู่ภายใต้ลักษณะดังกล่าว
    - $P(+) < P(-)$ ทุกคู่ระหว่าง (X_i, Y_i) จะอยู่ภายใต้ลักษณะดังกล่าว
    - $P(+) = P(-)$ ทุกคู่ระหว่าง (X_i, Y_i) จะอยู่ภายใต้ลักษณะดังกล่าว

10

Sign Test

- Statistical Hypothesis
  - การทดสอบ 2 ทาง (2 tailed test)

    $H_0 : P(X_i > Y_i) = P(X_i < Y_i)$ vs. $H_1 : P(X_i > Y_i) \neq P(X_i < Y_i)$

    สำหรับทุกคู่ i

    หรือ $H_0 : P(+) = P(-)$ vs. $H_1 : P(+) \neq P(-)$

    หรือ $H_0 : E(X) = E(Y)$ vs. $H_1 : E(X) \neq E(Y)$
  - การทดสอบทางเดียวด้านมาก

    $H_0 : P(X_i > Y_i) \leq P(X_i < Y_i)$ vs. $H_1 : P(X_i > Y_i) > P(X_i < Y_i)$

    สำหรับทุกคู่ i

    หรือ $H_0 : P(+) \leq P(-)$ vs. $H_1 : P(+) > P(-)$

    หรือ $H_0 : E(X) \leq E(Y)$ vs. $H_1 : E(X) > E(Y)$

11

Sign Test

- Statistical Hypothesis (ต่อ)
  - การทดสอบทางเดียวด้านน้อย

    $H_0 : P(X_i > Y_i) \geq P(X_i < Y_i)$ vs. $H_1 : P(X_i > Y_i) < P(X_i < Y_i)$

    สำหรับทุกคู่ i

    หรือ $H_0 : P(+) \geq P(-)$ vs. $H_1 : P(+) < P(-)$

    หรือ $H_0 : E(X) \geq E(Y)$ vs. $H_1 : E(X) < E(Y)$

- Statistical Test

  T = จำนวนคู่ที่มีสัญลักษณ์ " + "

12

Sign Test

- Critical Regions
  - เมื่อกำหนดระดับนัยสำคัญ α
  - n = จำนวนคู่ทั้งหมดที่ไม่ใช่ tie ($n \leq n'$)

  - **กรณีการทดสอบ 2 ทาง** (2 tailed test)
    จะกำหนดค่าวิกฤตทั้งด้านซ้าย ($W_{\alpha/2}$) และขวา ($W_{1-\alpha/2}$) นั่นคือ
    เขตวิกฤต : $T \leq W_{\alpha/2}$ หรือ $T \geq W_{1-\alpha/2}$
    - **สำหรับตัวอย่างขนาดเล็ก ($n \leq 20$)**
      ใช้ตารางทวินามที่ p = 1/2 ทำการพิจารณาค่า
      $P(X \leq t) = \alpha/2$ จะได้ว่า $W_{\alpha/2} = t$ และ $W_{1-\alpha/2} = n - t$
    - **สำหรับตัวอย่างขนาดใหญ่**
      ประมาณค่า $W_{\alpha/2} = np + z_{\alpha/2}\sqrt{npq} = ½[n + z_{\alpha/2}\sqrt{n}\,]$ และ
      $W_{1-\alpha/2} = ½[n + z_{1-\alpha/2}\sqrt{n}\,]$
    - **หรือใช้ตาราง Sign test**

13

Table 1 : Binomial Distribution (p. 1 – 12)

n	y	p = .50	.55	…
1	0	.5000	…	
	1	1.0000	…	
2	0	.2500	…	
	1	.7500	…	
	2	1.0000	…	
…				
7	0	.0078	…	
	1	.0625	…	
	…	…	…	
	7	1.0000	…	

n	y	p = .50	.55	…
8	0	.0039	…	
…	…	…		
9	0	.0020	…	
…				
10	0	.0010	…	
	1	.0107	…	
	2	.0547	…	
	3	.1719	…	
	4	.3770	…	
	5	.6230	…	
	6	.8281	…	

14

Table 21 : Critical values for Sign test (p.48)

n	1%	5%	10%	25%	n	1%	5%	10%	25%
1					46	13	15	16	18
2					47	14	16	17	19
3				0	48	14	16	17	19
4				0	49	15	17	18	19
5			0	0	50	15	17	18	20
6		0	0	1	51	15	18	19	20
7		0	0	1	52	16	18	19	21
8	0	0	1	1	53	16	18	20	21
9	0	1	1	2	54	17	19	20	22
10	0	1	1	2	55	17	19	20	22
…					…				

15

Sign Test

- Critical Regions (ต่อ)
  - **กรณีการทดสอบทางเดียวด้านมาก**
    จะกำหนดค่าวิกฤตเฉพาะด้านขวา ($W_{1-\alpha}$) นั่นคือ
    $\boxed{\text{เขตวิกฤต : } T \geq W_{1-\alpha}}$
    - **สำหรับตัวอย่างขนาดเล็ก ($n \leq 20$)**
      ใช้ตารางทวินามที่ $p = 1/2$ ทำการพิจารณาค่า
      $P(X \leq t) = \alpha$ จะได้ว่า $W_{\alpha} = t$ และ $W_{1-\alpha} = n - t$
    - **สำหรับตัวอย่างขนาดใหญ่**
      ประมาณค่า $W_{1-\alpha} = ½[n + z_{1-\alpha}\sqrt{n}\,]$
    - **หรือใช้ตาราง Sign test**

16

Sign Test

- Critical Regions (ต่อ)
  - **กรณีการทดสอบทางเดียวด้านน้อย**

    จะกำหนดค่าวิกฤตเฉพาะด้านซ้าย (W_α) นั่นคือ

    เขตวิกฤต : $T \le W_\alpha$
    - **สำหรับตัวอย่างขนาดเล็ก ($n \le 20$)**

      ใช้ตารางทวินามที่ $p = 1/2$ ทำการพิจารณาค่า

      $P(X \le t) = \alpha$ จะได้ว่า $W_\alpha = t$
    - **สำหรับตัวอย่างขนาดใหญ่**

      ประมาณค่า $W_\alpha = ½[n + z_\alpha\sqrt{n}\,]$
    - **หรือใช้ตาราง Sign test**

17

Sign Test

- ทดลองหาค่าวิกฤต
  - จงหาค่าวิกฤตของการทดสอบ 2 ทาง เมื่อกำหนด $\alpha = 0.05$ และ $n = 10$

  - จงหาค่าวิกฤตของการทดสอบทางเดียวด้านมาก เมื่อกำหนด $\alpha = 0.05$ และ $n = 10$

18

Sign Test

- **การตัดสินใจ**

  จะปฏิเสธ H_0 เมื่อค่า T_{cal} ที่คำนวณได้ตกอยู่ในเขตวิกฤต

ตัวอย่าง 1

บริษัทขายไก่ย่างแห่งหนึ่งต้องการที่จะทดสอบดูว่าการใช้เครื่องปรุงในการทำไก่ย่างที่แตกต่างกัน 2 วิธี ลูกค้าจะชอบทั้งสองสูตรแตกต่างกันหรือไม่ จึงได้เตรียมไก่ย่างที่ปรุงโดยใช้เครื่องปรุงทั้ง 2 สูตร แยกต่างหากจากกัน แล้วจึงสุ่มตัวอย่างลูกค้ามา 20 คน โดยให้ชิมไก่ย่างอย่างละชิ้น และให้บอกว่าชอบชิ้นไหนมากกว่ากัน ชิ้นที่ชอบมากให้อันดับ 1 และชิ้นที่ชอบน้อยให้อันดับ 2 ถ้าชอบทั้ง 2 ชิ้นเท่า ๆ กัน ให้อันดับ 1 ทั้งคู่ ผลปรากฏดังนี้

สูตร ก (X)	1	1	1	1	1	2	1	2	1	2	2	2	1	1	2	1	1	1	1	1
สูตร ข (Y)	2	2	2	2	1	1	2	1	2	1	1	1	2	2	1	1	2	2	2	2

สูตร ก (X)	1	1	1	1	1	2	1	2	1	2	2	2	1	1	2	1	1	1	1	1
สูตร ข (Y)	2	2	2	2	1	1	2	1	2	1	1	1	2	2	1	1	2	2	2	2
Compare	+	+	+	+	0	-	+	-	+	-	-	-	+	+	-	0	+	+	+	+

- พิจารณาลักษณะข้อมูล
  - เป็นข้อมูล 2 ชุดที่มีความสัมพันธ์กัน (แบบ Repeated measures)
  - ข้อมูลอยู่ในมาตรเรียงลำดับ (1 = ชอบมาก, 2 = ชอบน้อย)
  - ต้องการเปรียบเทียบความชอบของลูกค้าระหว่างสูตร ก กับ ข ว่าแตกต่างกันหรือไม่
- Hypothesis

  $H_0 : P(+) = P(-)$

  $H_1 : P(+) \neq P(-)$
- Statistical Test

  T_{cal} = จำนวนคู่ที่ "+"

  = 12 **[n'=20, n=18]**

- **Hypothesis**

  $H_0 : P(+) = P(-)$

  $H_1 : P(+) \neq P(-)$
- **Statistical Test**

  T_{cal} = จำนวนคู่ที่ "**+**"

  = 12 [n' = 20, n = 18]
- **Critical Region** กำหนด $\alpha = 0.05$

  ทดสอบ 2 ทาง, n = 18 ใช้ตาราง Sign test

  ดังนั้น $W_{\alpha/2} = 4$ และ $W_{1-\alpha/2} = 18 - 4 = 14$

  เขตวิกฤต คือ $T \leq 4$ หรือ $T \geq 14$
- **Conclusion**

  จาก T_{cal} ไม่ตกในเขตวิกฤต จึง **Accept H_0** นั่นคือ

  ที่ $\alpha = 0.05$ ลูกค้าชอบทั้ง 2 สูตรไม่แตกต่างกัน

22

ตัวอย่าง 2

ในการสัมภาษณ์บิดามารดาของนักเรียนกลุ่มหนึ่ง โดยตั้งคำถามว่าลูกของตนเองเป็นคนดีเพียงใด ให้ตอบโดยเลือกมาตราส่วนประมาณค่า (Rating scales) ที่มี 5 ระดับ ซึ่ง 5 หมายถึงดีที่สุดและลดหลั่นลงไปถึง 1 หมายถึงดีน้อยที่สุด ผลการสอบถามพ่อและแม่ของนักเรียน (โดยใช้ rate ลูกคนเดียวกัน) ปรากฏดังนี้

คนที่	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Rate ของพ่อ	3	1	5	5	5	4	5	5	5	5	4	4	5	5	3	2	5
Rate ของแม่	3	2	3	2	2	5	3	5	3	1	2	3	3	3	3	3	3

ที่ระดับนัยสำคัญ .05 จงทดสอบว่าการ rate ของพ่อสูงกว่าของแม่จริงหรือไม่

23

คนที่	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Rate ของพ่อ	3	1	5	5	5	4	5	5	5	5	4	4	5	5	3	2	5
Rate ของแม่	3	2	3	2	2	5	3	5	3	1	2	3	3	3	3	3	3
Compare	0	-	+	+	+	-	+	0	+	+	+	+	+	+	0	-	+

- พิจารณาลักษณะข้อมูล
  - เป็นข้อมูล 2 ชุดที่มีความสัมพันธ์กัน (Repeated measures)
  - ข้อมูลอยู่ในมาตรเรียงลำดับ (5 = ดีสุด, ..., 1 = ดีน้อยสุด)
  - ต้องการทดสอบว่า การ rate ของพ่อสูงกว่าของแม่หรือไม่
- ให้ X = Rate ของพ่อ, Y = Rate ของแม่

- Hypothesis

  $H_0 : P(+) \leq P(-)$

  $H_1 : P(+) > P(-)$

- Statistical Test

  T_{cal} = จำนวนคู่ที่ "+"

  = 11 **[n'=17, n=14]**

- **สมมติฐาน**

  $H_0 : P(+) \leq P(-)$

  $H_1 : P(+) > P(-)$

- **ตัวสถิติทดสอบ**

  T_{cal} = จำนวนคู่ที่ "**+**"

  = 11 [n′ = 17, n = 14]

- **เขตวิกฤต** กำหนด $\alpha = 0.05$

  ทดสอบทางเดียวด้านมาก, n = 14 ใช้ตาราง Sign test

  ดังนั้น $W_{\alpha} = 3$ และ $W_{1-\alpha} = 14 - 3 = 11$

  เขตวิกฤต คือ $T \geq 11$

- **สรุปผล**

  จาก T_{cal} ตกในเขตวิกฤต จึง **Reject H_0** นั่นคือ

  ที่ $\alpha = 0.05$ การ rate ของพ่อสูงกว่าของแม่

25

ตัวอย่าง 3

บริษัทอุตสาหกรรมแห่งหนึ่งต้องการศึกษาดูว่า การนำระบบจ่ายค่าตอบแทนแผนใหม่เข้ามาใช้แทนแผนเก่าจะทำให้ผลิตภาพของคนงานดีขึ้นหรือไม่ จึงสุ่มตัวอย่างคนงานมา 25 คน และบันทึกปริมาณการผลิตรายสัปดาห์ก่อนและหลังจากที่ได้นำระบบการจ่ายค่าตอบแทนแผนใหม่เข้ามาใช้ ผลการสำรวจปรากฏดังตารางต่อไปนี้ โดยให้

X แทนปริมาณสินค้าที่ผลิตได้หลังใช้ระบบการจ่ายค่าตอบแทนแผนใหม่

Y แทนปริมาณสินค้าที่ผลิตได้ก่อนใช้ระบบจ่ายค่าตอบแทนแผนใหม่

X	79	87	70	93	85	75	80	71	80	88	82	71	75	85
Y	83	85	75	91	80	75	90	65	78	85	83	75	78	81

X	86	85	82	87	78	84	85	81	76	80	82
Y	82	88	85	81	78	81	70	80	78	80	79

จงทดสอบสมมติฐานที่ระดับนัยสำคัญ .05

26

X	79 87 70 93 85 75 80 71 80 88 82 71 75 85
Y	83 85 75 91 80 75 90 65 78 85 83 75 78 81

X	86 85 82 87 78 84 85 81 76 80 82
Y	82 88 85 81 78 81 70 80 78 80 79

- พิจารณาลักษณะข้อมูล
  - เป็นข้อมูล 2 ชุดที่มีความสัมพันธ์กัน (Repeated measures)
  - ข้อมูลอยู่ในมาตรอัตราส่วน
  - ต้องการทดสอบว่า การนำระบบใหม่มาใช้จะทำให้ผลิตภาพของคนงานดีขึ้นหรือไม่

X	79 87 70 93 85 75 80 71 80 88 82 71 75 85
Y	83 85 75 91 80 75 90 65 78 85 83 75 78 81

Compare - + - + + 0 - + + + - - - +

X	86 85 82 87 78 84 85 81 76 80 82
Y	82 88 85 81 78 81 70 80 78 80 79

Compare + - - + 0 + + + - 0 +

- Hypothesis

  $H_0 : P(+) \leq P(-)$

  $H_1 : P(+) > P(-)$

- Statistical Test

  T_{cal} = จำนวนคู่ที่ "+"

  = 13 **[n'=25, n=22]**

- **สมมติฐาน**

  $H_0 : P(+) \leq P(-)$

  $H_1 : P(+) > P(-)$

- **ตัวสถิติทดสอบ**

  T_{cal} = จำนวนคู่ที่ "**+**"

  = 13 [n′ = 25, n = 22]

- **เขตวิกฤต** กำหนด $\alpha = 0.05$

  ทดสอบทางเดียวด้านมาก, n = 22 ใช้ตาราง Sign test

  ดังนั้น $W_{\alpha} = 6$ และ $W_{1-\alpha} = 22 - 6 = 16$

  เขตวิกฤต คือ $T \geq 16$

- **สรุปผล**

  จาก T_{cal} ตกในเขตวิกฤต จึง **Accept H_0** นั่นคือ

  ที่ $\alpha = 0.05$ ระบบใหม่ไม่ช่วยให้ผลิตภาพของคนงานดีขึ้น

29