

การกำหนดขนาดตัวอย่างสำหรับสถิติวิเคราะห์ t-test for two samples

โดยการวิเคราะห์อำนาจการทดสอบด้วยโปรแกรม R

เขมิกา อูระวงศ์

คณะวิทยาศาสตร์ มหาวิทยาลัยสงขลานครินทร์

khamika.u@gmail.com; jantra.s@hotmail.com

ตัวอย่าง 1 เมื่อกำหนดอำนาจการทดสอบ (Power of the test)

ในการบรรจุขวดน้ำดื่มยี่ห้อ A และ ยี่ห้อ B ต้องการทดสอบว่าปริมาตรของน้ำขวดทั้งสองยี่ห้อ มีปริมาตรแตกต่างกันหรือไม่ ($H_0 : \mu_1 = \mu_2$ $H_1 : \mu_1 \neq \mu_2$) จากข้อมูลสำรวจนั้ร่งพบว่า ค่าเฉลี่ยปริมาตรของน้ำขวดยี่ห้อ A และ B เท่ากับ 355 และ 345 มิลลิลิตร ตามลำดับ ส่วนเบี่ยงเบนมาตรฐานของปริมาตรน้ำทั้งสองกลุ่ม (SD) ประมาณ 5 มิลลิลิตรใกล้เคียงกัน

1.1 ผู้วิจัยควรกำหนดขนาดตัวอย่างน้ำขวดทั้งสองยี่ห้อเท่าไร? เมื่อกำหนดอำนาจการทดสอบ 0.9

1.2 อำนาจการทดสอบของการศึกษาเป็นเท่าไร? เมื่อใช้ขนาดตัวอย่างแต่ละกลุ่มเท่ากันเป็น 9 ขวด

วิธีทำ 1.1 ผู้วิจัยควรกำหนดขนาดตัวอย่างน้ำขวดทั้งสองยี่ห้อเท่าไร? เมื่อกำหนดอำนาจการทดสอบ 0.9

กรณีนี้ เป็นการวิเคราะห์ข้อมูลเพื่อทดสอบความแตกต่างระหว่างค่าเฉลี่ยประชากรสองกลุ่มอิสระกัน ใช้สถิติทดสอบ t-test for two independent samples การกำหนดขนาดตัวอย่างด้วยการกำหนดอำนาจการทดสอบ (Power of the test) บนโปรแกรม R มีขั้นตอนดังนี้

ขั้นที่ 1. กำหนดค่าต่างๆ ดังนี้

1.1 กำหนดอำนาจการทดสอบ 0.80-0.90

เช่น กำหนด 0.90-> power=0.9

1.2 กำหนดระดับนัยสำคัญ 0.01- 0.05

เช่น กำหนด 0.05 -> sig.level=0.05

ขั้นที่ 2. คำนวณค่า design effect (d) ดังนี้

$$d = \frac{\bar{x}_1 - \bar{x}_2}{SD} = \frac{355 - 345}{5} = \frac{10}{5} = 2$$

เมื่อ $\bar{x}_1 - \bar{x}_2$ คือ ความแตกต่างระหว่าง ค่าเฉลี่ยตัวอย่าง 2 กลุ่มภายใต้สมมติฐาน H_0

ขั้นที่ 3. คำสั่งใน R และผลลัพธ์ของการวิเคราะห์

คำสั่งใน R:

```
>install.packages("pwr")
```

```
>library(pwr)
```

```
>pwr.t.test(n = NULL, d = 2, power=0.9, sig.level=0.05, type="two.sample", alternative="two.sided")
```

หมายเหตุ

- type = "two.sample" "one.sample" หรือ "paired" สำหรับข้อมูลตัวอย่าง 2 กลุ่มเป็นอิสระกัน 1 กลุ่ม หรือ ข้อมูลแบบจับคู่ ตามลำดับ
- alternative = "two.sided" "less" หรือ "greater" สำหรับการทดสอบสมมติฐานสองทาง ทางเดียว เครื่องหมายน้อยกว่า หรือทางเดียวเครื่องหมายมากกว่า ตามลำดับ

ผลลัพธ์:

Two-sample t test power calculation

n = 6.386753

d = 2

sig.level = 0.05

power = 0.9

alternative = two.sided

NOTE: n is number in *each* group

การแปลผล:

$n = 6.386753 \approx 7$ (หมายเหตุ กรณีค่าไม่เป็นจำนวนเต็ม ให้ปัดขึ้นทุกกรณี)

หมายความว่า ผู้วิจัยควรสุ่มตัวอย่างขนาดน้ำดื่มยี่ห้อ A และ B อย่างน้อยกลุ่มละ 7 ขวด สำหรับการศึกษานี้ จะมั่นใจได้ว่ามีอำนาจการทดสอบ 0.90

วิธีทำ 1.2 อำนาจการทดสอบของการศึกษาเป็นเท่าไร? เมื่อใช้ขนาดตัวอย่างแต่ละกลุ่มเท่ากันเป็น 7 ขวด กำหนดค่าต่างๆ ทำนองเดียวกันกับ 1.1. และใช้คำสั่งใน R และผลลัพธ์ของการวิเคราะห์ดังนี้

คำสั่งใน R:

```
>install.packages("pwr")
>library(pwr)
>pwr.t.test(n = 7, d = 2, power=NULL, sig.level=0.05, type="two.sample",alternative="two.sided")
```

หมายเหตุ

- type = "two.sample" "one.sample" หรือ "paired" สำหรับข้อมูลตัวอย่าง 2 กลุ่มเป็นอิสระกัน 1 กลุ่ม หรือ ข้อมูลแบบจับคู่ ตามลำดับ
- alternative = "two.sided" "less" หรือ "greater" สำหรับการทดสอบสมมติฐานสองทาง ทางเดียว เครื่องหมายน้อยกว่า หรือทางเดียวเครื่องหมายมากกว่า ตามลำดับ

ผลลัพธ์:

Two-sample t test power calculation

n = 7

d = 2

sig.level = 0.05

power = 0.9290

alternative = two.sided

NOTE: n is number in *each* group

การแปลผล:

หมายความว่า การศึกษานี้ จะมีอำนาจการทดสอบ 0.9290 เมื่อกำหนดขนาดตัวอย่างขวดบรรจุน้ำยี่ห้อละ 7 ขวด ที่ effect size เท่ากับ 2

ตัวอย่าง 2 กรณีสถิติวิเคราะห์ t-test for two dependent samples (paired t-test)

ผู้วิจัยคาดว่าน้ำหนักของข้างป่าจะเพิ่มขึ้นหลังกินอาหารสูตรใหม่ ($H_0 : \mu_d = 0$ $H_1 : \mu_d > 0$) จึงทำการศึกษาข้างป่า 3 เชือก ด้วยการชั่งน้ำหนักก่อนและหลังกินอาหารสูตรใหม่ ได้น้ำหนักที่ลดลงไปเฉลี่ย 5 กิโลกรัม ตามลำดับ และส่วนเบี่ยงเบนมาตรฐาน (SD) เท่ากับ 4 กิโลกรัม

- 2.1 ผู้วิจัยควรกำหนดขนาดตัวอย่างข้างเท่าไร? เมื่อกำหนดอำนาจการทดสอบ (Power of the test) 0.8
- 2.2 อำนาจการทดสอบของการศึกษาเป็นเท่าไร? เมื่อใช้ขนาดตัวอย่างข้าง 3 เชือก

วิธีทำ 2.1 ผู้วิจัยควรกำหนดขนาดตัวอย่างข้างเท่าไร? เมื่อกำหนดอำนาจการทดสอบ (Power of the test) 0.8

กรณีนี้ เป็นการวิเคราะห์ข้อมูลเพื่อทดสอบการเปลี่ยนแปลงของค่าเฉลี่ยประชากร ใช้สถิติทดสอบ paired t-test การกำหนดขนาดตัวอย่างด้วยการกำหนดอำนาจการทดสอบ (Power of the test) บนโปรแกรม R มีขั้นตอนดังนี้

ขั้นที่ 1. กำหนดค่าต่างๆ ดังนี้

- 1.1 อำนาจการทดสอบ 0.80-0.90 เช่น กำหนด 0.90-> power=0.9
- 1.2 ค่าเฉลี่ยของความแตกต่างระหว่างน้ำหนักหลังและก่อนการทดลอง เช่น กำหนด 5 กิโลกรัม
-> e= 5
- 1.3 กำหนดระดับนัยสำคัญ 0.01- 0.05 เช่น กำหนด 0.05 -> sig.level=0.05

ขั้นที่ 2. คำนวณค่า design effect (d) ดังนี้

$$d = e/SD = 5/4$$

ขั้นที่ 3. คำสั่งใน R และผลลัพธ์ของการวิเคราะห์

```
>install.packages("pwr")
>library(pwr)
>pwr.t.test(n = NULL, d = 5/4, power=0.9, sig.level=0.05, type="paired", alternative="greater")
```

หมายเหตุ

- type = "two.sample" "one.sample" หรือ "paired" สำหรับข้อมูลตัวอย่าง 2 กลุ่มเป็นอิสระกัน 1 กลุ่ม หรือ ข้อมูลแบบจับคู่ ตามลำดับ
- alternative = "two.sided" "less" หรือ "greater" สำหรับการทดสอบสมมติฐานสองทาง ทางเดียว เครื่องหมายน้อยกว่า หรือทางเดียวเครื่องหมายมากกว่า ตามลำดับ

ผลลัพธ์:

Paired t test power calculation

n = 7.0726

d = 1.25

sig.level = 0.05

power = 0.9

alternative = greater

NOTE: n is number of *pairs*

การแปลผล:

$n = 7.0726 \approx 8$ (หมายเหตุ กรณีค่าไม่เป็นจำนวนเต็ม ให้ปัดขึ้นทุกกรณี)

หมายความว่า ผู้วิจัยควรกำหนดข้างปาดอย่างน้อย 8 เชือก จะมีโอกาสหรืออำนาจการทดสอบ 0.9 (หรือร้อยละ 90) ที่จะปฏิเสธ H_0 เมื่อสมมติว่า H_1 จริง ที่ระดับนัยสำคัญ 0.05

วิธีทำ 2.2 อำนาจการทดสอบของการศึกษาเป็นเท่าไร? เมื่อใช้ขนาดตัวอย่างข้าง 3 เชือก

การกำหนดค่าต่างๆ ทำนองเดียวกับข้อ 2.1 โดยคำสั่งใน R และผลลัพธ์ของการวิเคราะห์ดังนี้

```
>install.packages("pwr")
```

```
>library(pwr)
```

```
>pwr.t.test(n = 3, d = 5/4, power=NULL, sig.level=0.05, type="paired", alternative="greater ")
```

หมายเหตุ

- type = "two.sample" "one.sample" หรือ "paired" สำหรับข้อมูลตัวอย่าง 2 กลุ่มเป็นอิสระกัน 1 กลุ่ม หรือ ข้อมูลแบบจับคู่ ตามลำดับ
- alternative = "two.sided" "less" หรือ "greater" สำหรับการทดสอบสมมติฐานสองทาง ทางเดียว เครื่องหมายน้อยกว่า หรือทางเดียวเครื่องหมายมากกว่า ตามลำดับ

ผลลัพธ์:

Paired t test power calculation

n = 3

d = 1.25

sig.level = 0.05

power = 0.4230

alternative = greater

NOTE: n is number of *pairs*

การแปลผล:

power = 0.4230

หมายความว่า จะมีโอกาสหรืออำนาจการทดสอบ 0.4230 (หรือร้อยละ 42.30) ที่จะปฏิเสธ H_0 เมื่อสมมติว่า H_1 จริง ที่ระดับนัยสำคัญ 0.05 เมื่อกำหนดการศึกษาข้างปาดัวอย่าง 3 เชือก

เอกสารอ้างอิง/อ่านเพิ่มเติม:

วินัย โพธิ์สุวรรณ (2562). เอกสารอบรมเชิงปฏิบัติการหลักสูตร “สถิติและการวางแผนการวิจัยที่ใช้สัตว์ (Statistics and Research Designs Using Animals) รุ่นที่ 9”. สนับสนุนโดย สถาบันพัฒนาการดำเนินการต่อสัตว์เพื่องานทางวิทยาศาสตร์ (สพสว.) สำนักงานคณะกรรมการวิจัยแห่งชาติ (วช.)