

แบบฝึกหัดในคาบเรียนครั้งที่ 10

SEL TOP ED RES (Spring 16)

แบบฝึกหัดที่ 10.1

1. จงใช้คำสั่ง `rbinom(100, 1, 0.2)` เพื่อสร้าง **vector** ของตัวเลข 1 และ 0 รวม 100 ตัว โดยที่สัดส่วนของเลข 1 ในประชากรเท่ากับ 0.2
2. จงใช้ **for-loop** เพื่อสร้าง **vector** ดังกล่าว 1000 ครั้ง และหาสัดส่วนของเลข 1 ในแต่ละ **vector**
3. จงหาค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของค่าสัดส่วน
4. จงสร้าง **histogram** ของค่าสัดส่วนทั้ง 1000 ค่า

แบบฝึกหัดที่ 10.2

1. จงจัดเก็บข้อมูลในการสุ่มแต่ละครั้งในข้อที่ 1 ของแบบฝึกหัด 10.1 ใน `binary.1` โดยใช้คำสั่ง `sapply`
2. จากข้อมูลในข้อที่ 1 จงใช้คำสั่ง `sapply` เพื่อหาสัดส่วนของเลข 1 จากการสุ่มแต่ละครั้ง
3. จากข้อมูลในข้อที่ 1 จงใช้คำสั่ง `lapply` เพื่อหา **chi-square: goodness-of-fit** เพื่อทดสอบว่าค่าสัดส่วนที่สุ่มมาได้ แตกต่างจาก .2 อย่างมีนัยสำคัญหรือไม่
4. จากผลการวิเคราะห์ในข้อที่ 3 จงใช้คำสั่ง `sapply` เพื่อดึง **p-value** ออกมาจากผลการวิเคราะห์แต่ละชิ้น
5. จงหา **Type I error** ของการทดสอบจากข้อที่ 4

แบบฝึกหัดที่ 10.3

1. จากผลการวิเคราะห์ในแบบฝึกหัดที่ 10.2 จงวิเคราะห์ผลโดยใช้ **parallel processing** โดยใช้ `parLapply` หรือ `mclapply` เพียงแค่ครั้งเดียว กล่าวคือทำกระบวนการข้อที่ (1) ถึง (4) ภายในฟังก์ชันที่ถูกกระจายด้วย **parallel processing**

การบ้านที่ 10

1. จงใช้คำสั่ง `rnorm` เพื่อสร้างข้อมูลมา แล้วหาค่าส่วนเบี่ยงเบนมาตรฐาน (**sd**) จำนวน 1000 ครั้ง แล้วทดสอบว่าค่าเฉลี่ยของส่วนเบี่ยงเบนมาตรฐานนั้น แตกต่างจากค่าพารามิเตอร์เท่าไร
 - a. ใช้ **for loop** เพื่อประมวลผล
 - b. ใช้ `sapply` หรือ `lapply` เพื่อประมวลผล
 - c. ใช้ `mclapply` หรือ `parLapply` เพื่อประมวลผล

2. ในการทดสอบความแม่นยำของช่วงเชื่อมั่น นักวิจัยจะทดสอบว่าช่วงเชื่อมั่นนั้นคลุมค่าพารามิเตอร์หรือไม่ ให้ใช้ script ดังต่อไปนี้ เพื่อทดสอบความแม่นยำของช่วงเชื่อมั่นของสัมประสิทธิ์ถดถอย

```
n <- 100
x <- rnorm(n, 0, 1)
y <- 0.5*x + rnorm(n, 0, 0.8)
cf <- confint(lm(y ~ x))
(cf[2, 1] < 0.5) & (0.5 < cf[2, 2])
```

กล่าวคือ ให้หาว่าในจำนวน 1,000 ครั้ง มีกี่ครั้งที่ช่วงเชื่อมั่นคลุมค่าพารามิเตอร์

- ใช้ for loop เพื่อประมวลผล
- ใช้ sapply หรือ lapply เพื่อประมวลผล
- ใช้ mclapply หรือ parLapply เพื่อประมวลผล