ข้อสอบครั้งที่ 2 (ข้อสอบ Take-home)

วิชา: 2756886 เรื่องคัดเฉพาะทางการวิจัยทางการศึกษา (การใช้โปรแกรม R)

อาจารย์ผู้สอน: อ. ดร. สันทัด พรประเสริฐมานิต

วัน-เวลาทดสอบ: 29-31 มีนาคม พ.ศ. 2559 (60 คะแนน)

# คำสั่ง

1. จงเปิดไฟล์ R ขึ้นมาไฟล์หนึ่ง ตั้งชื่อโดยนำด้วยรหัสประจำตัวนิสิต แล้วตามด้วยคำว่า Test2 เช่น 5947426527Test2.R ไม่ต้องใส่ชื่อนามสกุลของคุณลงในไฟล์คำตอบ
2. จงตอบคำถามทุกข้อ โดยก่อนตอบแต่ละคำข้อคำถามให้ใส่เลขข้อลงใน comment เช่น #### Item 1 หากคำตอบไม่ตรงตำแหน่งของข้อคำถาม ผมขอไม่คิดคะแนนให้
3. เมื่อตอบเสร็จแล้ว ให้ส่งคำตอบของคุณเข้ามาที่ [psunthud@gmail.com](mailto:psunthud@gmail.com) โดยให้แนบไฟล์คำตอบ พร้อมทั้งใส่ชื่อ นามสกุล และเลขประจำตัวนิสิตภายในเนื้อหาของอีเมลล์ และตรวจสอบกับอาจารย์ว่าได้รับไฟล์แล้วก่อนออกจากห้อง
4. ให้ทำ comment ในทุกๆ 4-5 บรรทัด อธิบายว่าคุณกำลังทำอะไรอยู่ หากผมไม่เข้าใจและต้องใช้ความพยายามอย่างที่ไม่สมควรจะเป็น ในการอ่าน code ของคุณ ผมขอหักคะแนนส่วนดังกล่าว
5. การสอบครั้งนี้เป็นการสอบจากการทำที่บ้าน คือ คุณทำของคุณคนเดียวจากที่ใดก็ได้ โดยไม่ต้องมีเพื่อนมานั่งทำด้วยกัน ผมอนุญาตให้คุณใช้ Internet และเปิดหนังสือและเอกสารทุกอย่าง
6. การสอบครั้งนี้ ไม่อนุญาตให้คุณสื่อสารกับผู้อื่น (รวมถึงเพื่อนของคุณ) เกี่ยวกับการสอบนี้ ไม่ว่าวิธีทางใดทั้งสิ้น ทั้งการคุย ส่งไฟล์ แชท โพสเวปบอร์ด ถาม pantip และพวกคุณไม่สามารถแชร์เอกสารเกี่ยวกับการสอบได้ครับ หากผมพบว่าพวกคุณมีลักษณะการเขียน code คล้ายกัน ผมจะปรับ F ทั้งสองคนครับ ดังนั้น คุณต้องระวังไม่ให้เพื่อนเห็น code ของคุณอย่างเด็ดขาดครับ
7. ผมมีสิทธิ์ในการถามซ้ำกับคุณอีกครั้งหนึ่ง เพื่ออธิบาย code ของคุณ หากคุณตอบไม่ได้หรือตอบได้ไม่ชัดเจน ผมถือว่าเป็นตัวบ่งชี้ว่าคุณไม่ได้ทำเอง ผมขออนุญาตสอบคุณใหม่อีกครั้งหนึ่ง และหากคะแนนต่ำลงอย่างเห็นได้ชัด (เช่นต่ำลงมากกว่า 30%) ผมถือว่าคุณอาจจะทุจริตในการสอบ ผมจะใช้คะแนนล่าสุด ผมจะส่งเรื่องให้คณะพิจารณาต่อไป
8. การละเมิดคำสั่งในข้อ 5-7 ถือเป็นการทุจริตการสอบ ต้องได้รับโทษตามระเบียบของจุฬาลงกรณ์มหาวิทยาลัย ซึ่งการลงโทษต่ำสุด คือ ได้ F ในวิชาที่สอบและพิจารณาพักการศึกษาไม่ต่ำกว่า 1 ภาคการศึกษา
9. คุณต้องตรวจสอบคำตอบของคุณก่อนส่งทุกครั้ง ว่าสามารถทำคำสั่งได้ โดยไม่มี error ใดๆ เกิดขึ้น
10. ผมขอไม่รับผิดชอบกับอุบัติเหตุทุกชนิดที่อาจเกิดขึ้นระหว่างสอบ เช่น เครื่องคอมพิวเตอร์เสีย ไฟล์ค้าง และอื่นๆ หากคุณไม่สามารถส่งไฟล์คำตอบได้ ถือว่าคุณสละสิทธิ์การสอบครั้งนี้ ดังนั้นผมขอแนะนำให้คุณบันทึกข้อมูลระหว่างทำบ่อยๆ และ backup ไฟล์ตลอด
11. เมื่อคุณอ่านเอกสารนี้แล้ว คุณไม่โต้แย้งผมภายใน 15:00 ไม่ว่าทางใดทางหนึ่ง ผมถือว่าคุณยอมรับกติกาในข้อที่ผ่านมาโดยอัตโนมัติ

# คำถาม

1. จงสร้างฟังก์ชั่นที่ใส่ตัวเลขในหน่วยเรเดียนแล้วเปลี่ยนให้เป็นองศา พร้อมทั้งแสดงตัวอย่างการใช้ (3 คะแนน)

2. จงหา Bootstrap confidence interval ของส่วนเบี่ยงเบนมาตรฐานของตัวแปรที่หนึ่งในข้อมูล attitude (4 คะแนน)

3. จงสร้างฟังก์ชั่นที่รายงานผลว่าค่าที่ใส่เข้าไปในฟังก์ชั่นนั้น เป็นตัวอักษร (character) ตัวเลข (numeric) หรืออื่นๆ (4 คะแนน)

4. จงสร้างฟังก์ชั่นที่ใส่ด้านประกอบมุมฉากของสามเหลี่ยม แล้วคืนค่าด้านตรงข้ามมุมฉาก ด้วยวิธีของ Pythagoras (3 คะแนน)

5. จงสร้าง barplot เพื่อแสดงค่าเฉลี่ยของตัวแปรตามจากปัจจัยทั้ง 2 ปัจจัยดังต่อไปนี้ (8 คะแนน)

ตารางแสดงค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของความสามารถในการจำ

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ดนตรี | คนฟังเพลงบ่อย | | คนไม่ฟังเพลงบ่อย | |
| *M* | *SD* | *M* | *SD* |
| Classical | 7.00 | 1.58 | 7.00 | 1.58 |
| Pop-dance | 6.00 | 1.58 | 5.50 | 1.58 |
| Heavy-metal | 5.00 | 1.58 | 4.50 | 1.58 |

6. จากข้อมูลในข้อที่ 5 จงสร้าง Line graph ด้วยฟังก์ชั่น plot (8 คะแนน)

7. นักวิจัยต้องการทดสอบประสิทธิภาพของยาลดความดัน นักวิจัยจึงแบ่งผู้ร่วมการทดลองออกเป็น 2 กลุ่ม คือ กลุ่มทานยาจริง (g = 1) และกลุ่มทานยาหลอก (g = 2) และมีการวัดความดันสามครั้งด้วยกัน คือ ก่อนทานยา (pre) หลังทานยาทันที (post) และหลังทานยาสองอาทิตย์ (fu) ข้อมูลสามารถสร้างขึ้นมาได้จาก syntax ดังต่อไปนี้

**set.seed(123321)**

**n <- 20**

**ST <- matrix(0.7 \* 25, 3, 3)**

**diag(ST) <- 25**

**datt <- MASS::mvrnorm(n, c(160, 140, 145), ST)**

**datc <- MASS::mvrnorm(n, c(160, 160, 160), ST)**

**dat2 <- rbind(cbind(1, datt), cbind(2, datc))**

**dat2 <- round(dat2, 0)**

**colnames(dat2) <- c("g", "pre", "post", "fu")**

1) จงตรวจสอบค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของความดันทั้งสามรอบของทั้งสองกลุ่ม พร้อมทั้งสร้างกราฟอธิบายค่าเฉลี่ย (2 คะแนน)

2) จงใช้ ANCOVA ในการเปรียบเทียบคะแนนหลังทานยา (post) ระหว่างกลุ่ม โดยใช้คะแนนก่อนทานยา (pre) เป็นคะแนนควบคุม (2 คะแนน)

3) จงใช้ Mixed-design ANOVA ในการตรวจสอบว่าทั้งสองกลุ่ม ทำให้คะแนนเปลี่ยนแปลงแตกต่างกันหรือไม่ (2 คะแนน)

4) จงใช้ MANOVA ในการตรวจสอบว่าทั้งสองกลุ่ม ทำให้คะแนนเปลี่ยนแปลงแตกต่างกันหรือไม่ (2 คะแนน)

5) จากข้อที่ 4 จงทดสอบ simple main effect ว่าคะแนนที่เปลี่ยนแปลงในแต่ละกลุ่มเป็นอย่างไร ควบคุม alpha level ด้วยวิธีของ Bonferroni (2 คะแนน)

6) จงเปรียบเทียบรายคู่ภายในแต่ละกลุ่มที่คะแนนในข้อที่ 5 ถึงระดับนัยสำคัญ พร้อมทั้งควบคุมระดับ alpha ให้เหมาะสม (2 คะแนน)

7) จงทดสอบว่าคะแนนหลังทานยาทันทีและหลังทานยา 2 สัปดาห์ มีความดันต่ำกว่าคะแนนก่อนทานยาหรือไม่ (2 คะแนน)

8. ท่านต้องการหาปัจจัยที่อธิบายความแตกต่างระหว่างกรกินเผ็ดโดยให้กลุ่มตัวอย่างปรุงรสก๋วยเตี๋ยวที่ตนเองทานประจำแล้ววัดน้ำหนักของพริกป่นที่ใส่ โดยปัจจัยที่ทดสอบมีดังนี้

ก) ภาค (region): 1 = เหนือ, 2 = กลาง, 3 = อีสาน, 4 = ใต้

ข) เพศ (sex): 1 = หญิง, 2 = ชาย

ค) อายุ (age)

**set.seed(123321)**

**region <- as.factor(rep(1:4, each = 100))**

**male <- rep(rep(0:1, each = 50), 4)**

**age <- sample(18:60, 400, replace = TRUE)**

**dregion <- model.matrix(lm(age ~ region))[,2:4]**

**hot <- 2 + 2\*dregion[,1] + 6\*dregion[,2] + 12\*dregion[,3] + 1\*(male - 0.5) - 0.05\*(age - 40) + rnorm(400, 0, 3.5)**

**hot[hot < 0] <- 0**

**hot <- round(hot, 2)**

**summary(lm(hot ~ region + male + age))**

**ex1 <- data.frame(id = sample(1:400, 400), hot, region, sex = male + 1, age)**

**ex1 <- ex1[order(ex1$id), ]**

จงตอบคำถามดังต่อไปนี้

1) ตัวแปรอิสระทั้งหมดสามารถอธิบายความแปรปรวนของการกินเผ็ดได้เท่ากับเท่าไร แตกต่างจาก 0 อย่างมีนัยสำคัญหรือไม่ (2 คะแนน)

2) จงทดสอบจุดตัดแกน Y และความชันของตัวแปรทั้งหมดว่าแตกต่างจาก 0 หรือไม่ พร้อมทั้งแปลความหมาย (2 คะแนน)

3) จงตรวจสอบว่าความแตกต่างระหว่างภาค สามารถอธิบายความแปรปรวนของการกินเผ็ด ได้เพิ่มเติมจากเพศและอายุเท่าไร และค่าที่เพิ่มขึ้นนั้นแตกต่างจาก 0 อย่างมีนัยสำคัญหรือไม่ (2 คะแนน)

4) จงเปลี่ยนค่าศูนย์กลางของอายุ เพื่อหาว่าค่าเฉลี่ยของการกินเผ็ดของของเพศชายภาคใต้อายุ 50 ปี มีการกินเผ็ดแตกต่างจาก 0 อย่างมีนัยสำคัญหรือไม่ (2 คะแนน)

5) จงสร้าง Residual plot และแปลความหมาย (2 คะแนน)

6) จงทดสอบว่าตัวแปรต้นทั้งหมด มีปัญหาเรื่อง Multicollinearity หรือไม่ (2 คะแนน)

7) จงทดสอบว่าเพศและอายุมีปฏิสัมพันธ์กันในการทำนายการกินเผ็ดหรือไม่ (2 คะแนน)

8) จงทดสอบ simple slope ของอายุในแต่ละเพศ (โดยไม่ต้องสนใจผลในข้อที่ 8) และสร้างกราฟเพื่อประกอบการวิเคราะห์ผล (2 คะแนน)