Assignment 2
2301520 FUNDAMENTALS of AMCS
(Due: July 26, 2019)

Instructions: Use Mathematica to answer the following questions.

1. จงสร้างข้อมูลสุ่มต่อไปนี้

   ```
   Table[{x,(1+Random[])x^4+x Random[]},{x,-1,1,0.05}]
   ```

   และให้ fit ข้อมูลเข้ากับโมเดลต่อไปนี้

   (a) $y = ax^4 + bx^3 + cx^2 + dx + e$
   (b) $y = ax^4 + bx$

2. จงสร้างข้อมูลสุ่มต่อไปนี้

   ```
   Table[{x,x+(4+Random[])Sin[x]+(1+Random[])Sin[5x]},{x,0,10,0.1}]
   ```

   และให้ fit ข้อมูลเข้ากับโมเดล $y = ax + b\sin(x) + c\sin(5x)$

3. จงสร้างข้อมูลสุ่มต่อไปนี้

   ```
   Table[{x,3 Exp[-0.8 x] Sin[4x+10]+Random[]},{x,0,4,0.1}]
   ```

   และให้ fit ข้อมูลเข้ากับโมเดลที่เหมาะสม

4. ถ้า N เป็นจำนวนสัตว์ทั้งหมด M เป็นจำนวนสัตว์ที่ถูกทำเครื่องหมายไว้ และ n เป็นจำนวนของสัตว์ที่ถูกสุ่มจับขึ้นมา แล้วจำนวนของสัตว์ที่ถูกทำเครื่องหมายจากสัตว์ที่สุ่มจับได้ จะมีการแจกแจงแบบ `HypergeometricDistribution[n,M,N]` สมมุติว่าเราทำเครื่องหมายสัตว์ชนิดนั้นไว้ 10 ตัว (จำนวนสัตว์ทั้งหมด N ยังไม่ทราบ)แล้วปล่อย จากนั้นเราสุ่มจับสัตว์นั้นมาได้ 8 ตัว แล้วพบว่า 3 ตัวถูกทำเครื่องหมายไว้ จงหาว่าจำนวนสัตว์ N ที่น่าจะเป็นไปได้ที่สุดเป็นเท่าใด ด้วยการใช้ Mathematica คำนวณค่าความน่าจะเป็นสำหรับ N หลาย ๆค่าแล้วดูว่าค่าไหนที่ทำให้ความน่าจะเป็นดังกล่าวมีค่ามากที่สุด คุณอาจจะใช้คำสั่ง `Table[`$expr$`, {`i`,` i_{min}`,`i_{max}`}]` และ `BarChart` เพื่อช่วยในการตัดสินใจ

5. ข้อมูลจากไฟล์ "sampledat.txt" เป็นข้อมูลของกลุ่มตัวอย่างจากประชากรกลุ่มหนึ่ง ให้ import ข้อมูลจากไฟล์ดังกล่าว แล้วให้คำนวณค่า sample mean, sample variance และ confidence interval ของ sample mean และ sample variance ที่ระดับ confidence level ที่ 95% และ 99%

6. ถ้า X และ Y เป็นตัวแปรสุ่มต่อเนื่องที่เป็นอิสระต่อกัน โดยที่มีค่า mean และ variance เป็นจำนวน finite ให้ c และ d เป็นค่าคงที่แล้วจะได้

$$E[cX + dY] = cE[X] + dE[Y]$$
$$Var(cX + dY) = c^2Var(X) + d^2Var(Y)$$

   ให้ลองตรวจสอบเอกลักษณ์ดังกล่าวโดยสร้างตัวอย่างข้อมูล 1000 ตัวอย่างของ X จากการแจกแจง Gamma(3,4) และ ตัวอย่างข้อมูล 1000 ตัวอย่างของ Y จากการแจกแจง Uniform(0,2) ให้ $c = 2$ และ $d = 3$ แล้วใช้ตัวอย่างข้อมูลที่สร้างได้คำนวณปริมาณทั้งสองข้างของสมการทั้งสองข้างต้นแล้วเปรียบเทียบกัน

7. Central limit theorem (CLT) กล่าวว่า

พิจารณาลำดับของตัวแปรสุ่มที่เป็นอิสระต่อกันและมีการแจกแจงเหมือนกัน (independent and identically distributed หรือ iid) โดยการแจกแจงดังกล่าวมีค่าเฉลี่ยและความแปรปรวนเป็นค่าจำกัด ผลรวมของตัวแปรสุ่มเหล่านั้นในจำนวนที่มากพอ จะมีการแจกแจงเข้าใกล้การแจกแจงแบบปกติ

ให้ลองตรวจสอบทฤษฎีบทนี้โดยสร้างตัวแปรสุ่มที่มีการแจกแจงอย่างใดอย่างหนึ่ง คำนวณผลรวมของตัวแปรสุ่มจำนวนหนึ่ง แล้วเขียน histogram ของผลรวมนั้น

Submission:

1. Name your homework file in the following format: `Assignment02-YourName.nb`. For example, if your name is Somsri, your file name will be `Assignment02-Somsri.nb`. (No need to put in your last name.)
2. Email your file to me at `boonyarit.i@chula.ac.th` with subject line in this format: `2301520 Assignment02-Yourname`. For example, if your name is Somchai, your subject line will be `2301520 Assignment02-Somchai`. (No need to put in your last name.)