

## Tentamen i Programmering i R, 7.5 hp

---

Skrivtid: 8-12  
Hjälpmedel: Böckerna *R Cookbook* (Teetor) och/eller *The Art of R Programming* (Matloff).  
Böckerna ska vara fria från anteckningar, men får innehålla understrykningar/överstrykningar samt flikar för kapitel/avsnitt.

Jourhavande lärare: Mattias Villani. Tillgänglig regelbundet under tentamen.  
Betygsgränser: Tentamen omfattar totalt 20 poäng. 10 poäng ger Godkänt. 16 poäng ger Väl godkänt.

**Skriv dina lösningar i fullständig och läsbar kod.**

**Lösningen skrivs i en körbar textfil med namnet Main.R.**

**Kommentera direkt i Main.R filen när något behöver förklaras eller diskuteras.**

**Eventuella grafer som skapas under tentans gång behöver INTE skickas in för rättning, det räcker med att skicka in den kod som producerar figurerna.**

---

### 1. Datastrukturer

- (a) Skapa en lista med namnet **staff** innehållande de tre elementen:
  - i. Variabeln **names** som är vektorn ("Mike", "Luke", "Adrian", "Sonja")
  - ii. Variabeln **wage** som är vektorn (24000, 17000, 31000, 36000)
  - iii. Variabeln **manager** som är en vektor med logiska värden (TRUE eller FALSE) som talar om Mike och Sonja är chefer (manager), men Luke och Adrian inte är chefer. 1.5p.
- (b) Skriv kod som väljer ut namnen på de chefer som har en lön (wage) som överstiger 30000. 1p.
- (c) Skriv kod som sänker lönen med 10% på alla icke-chefer som har mer än 20000 i lön. Listan **staff** ska alltså ändras om någon person uppfyller dessa kriterier. 1.5p.

### 2. Uppprepning/Loopar

- (a) Skriv en for-loop som skriver ut talen 3, 7, 12 och 14 på skärmen. 1p.
- (b) Skriv en for-loop som beräknar summan av elementen i vektorn (3, 7, 12, 14). 1.5p.
- (c) Skriv en while-loop som upprepade gånger genererar ett slumptal från en likformig fördelning (uniform distribution) över intervallet [0, 1] och sedan skriver ut meningen 'The random number is smaller than 0.9' på skärmen så länge som det genererade slumptalet är mindre än 0.9. Du ska alltså generera ett nytt slumptal och skriva till skärmen i varje iteration av loopen. Loopen avslutas första gången slumptalet överstiger 0.9. 1.5p.

### 3. Funktioner

- (a) Skriv en funktion som tar en godtycklig datavektor (ett stickprov) som input argument och som returnerar den s k variationskoefficienten. Variationskoefficienten är definierad som

$$\text{Variationskoefficient} = \frac{\text{Standardavvikelse}}{\text{Medelvärde}},$$

där både medelvärde och standardavvikelse beräknas från stickprovet. 1.5p.

- (b) Generera ett slumpmässigt stickprov med 10 observationer från en exponentialfördelning med väntevärde (mean) 2. Använd funktionen i Uppgift 3a för att beräkna variationskoefficienten i det skapade stickprovet. [Hint: **?rexp**] 1p.
- (c) Skriv en kod som använder funktionen i Uppgift 3a för att beräkna variationskoefficienten i 100 olika stickprov av storleken 10 från en exponentialfördelning med väntevärde (mean) 2. Koden ska alltså generera 100 stickprov med 10 dataobservationer i varje stickprov. För varje stickprov beräknas variationskoefficienten. Illustrera fördelningen för variationskoefficienten i dessa 100 stickprov med ett histogram. 1.5p.

#### 4. Regression

- (a) Läs in filen **fishes.txt** som en data frame. Denna data frame ska ha tre kolumner innehållande variablerna: **length**, **age** och **watertemp**. Varje rad innehåller mätningar av en fisks längd, ålder och vattentemperaturen i det kar som fisken odlas i. Filen **fishes.txt** har variablernas namn i första kolumnen och kolumnerna separeras av blanksteg (tomrum). 1p.
- (b) Skatta (estimate) regressionsmodellen

$$\text{length} = \alpha_0 + \alpha_1 \cdot \text{age} + \alpha_2 \cdot \text{watertemp} + \epsilon$$

och spara all resultat i variabeln **regModel**. 1p.

- (c) Plotta regressionsanpassningens residualer (residuals) mot variabeln **age**. Residualerna ska markeras med röda +-tecken i figuren, och *y*-axeln ska ha texten 'Nice residuals' 1p.
- (d) Beskriv kortfattat med ord och eventuellt lämpliga kommandon vilken sorts variabel **regModel** är och hur denna variabel skiljer sig från t ex en vektor eller matris. Notera att även denna uppgift ska enbart skrivas in i filen Main.R. 1p.

#### 5. Sekvenser och Funktionsobjekt

- (a) Rita upp den s k sinus-funktionen (**sin**) genom att först beräkna sinus-funktionen i 100 valda punkter mellan 0 och  $\pi$  (talet pi finns som inbyggd variabel i R). 2p.
- (b) Skriv en egen funktion (döp den till **PlotFunction**) som klarar av att rita upp en given funktion över ett angivet intervall för ett godtyckligt antal funktionsvärden. Användaren av funktionen ska alltså själv kunna välja vilken funktion som ska ritas upp, vilket intervall funktionen ska rita upp över och hur många funktionsvärden som ska användas. Användaren ska inte behöva ändra i din kod för **PlotFunction**. Du kan anta att den funktion som användaren vill rita upp är vektoriserad. Visa hur anropet till **PlotFunction()** ser ut för att lösa Uppgift 5a 2p.