

Bài 2.

Lập trình cơ bản trong R

1. Các phép toán và các hàm cơ bản trong R

1.1. Các phép toán trong R

+, -, *, /

1.2. Các phép toán logic:

! : NOT

&: AND

|: OR

<, >, <=, >=, ==, != (khác)

&&, ||

xor(x,y); isTRUE(x)

1.3 Các hàm toán học thường dùng trong R:

log(x) : logarit cơ số e

log10(x), log(x,n): logarit cơ số n

exp(x): e^x

sqrt(x): căn bậc 2 của x

factorial(x): $x!$

choose(n,k): tổ hợp n chập k

floor(x): giá trị nguyên $\leq x$ (sàn của x)

ceiling(x): giá trị nguyên $\geq x$ (trần của x)

trunc(x): làm tròn tới giá trị nguyên gần nhất giữa x và 0.

round(x, digits=n): làm tròn x đến n chữ số

signif(x, digits=n): hiển thị x dưới dạng dấu chấm thập phân, n tổng chữ số hiển thị

sin(x), cos(x), tan(x)

abs(x): $|x|$

$x \%/\% y$: lấy phần nguyên của phép chia x/y

$x \% \% y$: lấy phần dư của phép chia x/y

2. Lệnh điều kiện: if

```
if (bieu_thuc_dk)  bieu_thuc_1 else bieu_thuc_2
```

3. Vòng lặp: for, while và repeat

```
> for( bien_chay in bieu_thu_dk) bieu_thuc_lenh
```

```
> repeat bieu_thuc_lenh
```

```
> while bieu_thuc_dk    bieu_thuc_lenh
```

Ngắt vòng lặp: *break*, *next*.

4. Script và hàm:

4.1. Script:

- Tập hợp các đoạn lệnh
- Tạo script: vào File -> New script
- Gọi lại script: > source('ten_script.r')

4.2. Hàm:

Cú pháp:

```
> ten_ham <- function(tham_so1, tham_so2, ...) {cac_bieu_thuc_lenh}
```

Ví dụ: viết hàm tính sai số chuẩn,

```
stderr <- function(x) {  
  se <- sd(x)/length(x);  
  se  
}
```

5. Bài tập:

1/ Tạo một vec-tơ X chứa n phần tử (n: tự cho). Viết hàm tính tổng tích lũy đến vị trí thứ i của X.

2/ Thể tích hình cầu với bán kính r là: $V = 4\pi r^3/3$. Hãy viết hàm xây dựng 1 dataframe để tính thể tích hình cầu với bán kính tương ứng là 3,4,5,...,20. Cột radius lưu bán kính và cột volume lưu thể tích.

3/ Trong file data01.xls, dùng lệnh if và vòng lặp for để tạo biến Index theo yêu cầu sau:

Nếu Age <=60 thì Index = 0; 60 < Age <=70 thì Index = 1; 70 < Age <=80 thì Index = 2 và Age >80 thì Index = 3.

4/ File data11.xls chứa số liệu về chiều cao của 1 loại cây trồng theo bảng tần số dạng khoảng. Thực hiện các bước sau:

- a. Đọc số liệu từ data11.xls và gán vào 1 dataframe.
- b. Viết một hàm tính tham số là các biến trong dataframe vừa nhập, xuất ra các giá trị sau: chiều cao bé nhất, lớn nhất của cây, trung bình mẫu, phương sai mẫu hiệu chỉnh.

5/ Cho vec-tơ X chứa n giá trị quan sát, phân vị thứ p được xác định như sau

- Sắp xếp dữ liệu theo thứ tự tăng dần (từ nhỏ đến lớn).
- Tính chỉ số i :

$$i = \left(\frac{p}{100} \right) n$$

- Nếu i không phải là số nguyên, làm tròn i . Phân vị thứ p chính là giá trị nằm ở vị trí thứ i đã được làm tròn.
Nếu i nguyên, phân vị thứ i chính là giá trị trung bình của 2 giá trị nằm ở vị trí thứ i và thứ $i + 1$.

Hãy viết hàm **phanvi(X, P)** cho kết quả là phân vị thứ p từ vec-tơ X .