

BÀI THỰC HÀNH SỐ 2

1. Lệnh điều kiện: *if*

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

2. 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*.

3. 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  
}
```

4. 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

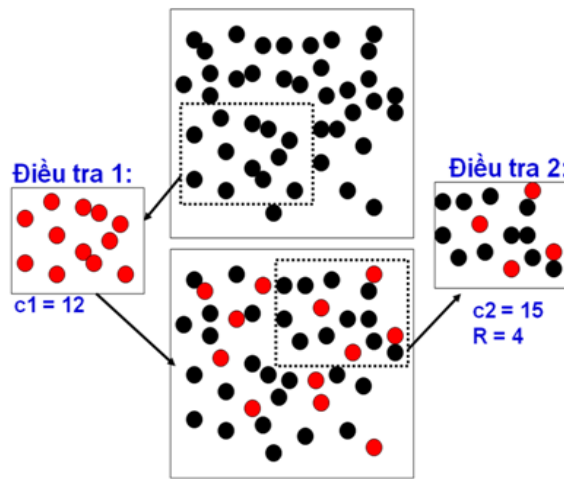
- Sắp xếp dữ liệu theo thứ tự tăng dần (từ nhỏ đến lớn).
- Tính chỉ số i :

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

- 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 .

6/ Phương pháp Capture – Recapture thường được dùng để ước lượng kích cỡ của quần thể (Ví dụ: Số cá trong một hồ nước). Phương pháp này được trình bày ngắn gọn như sau:



$$\frac{c_1}{N} \approx \frac{R}{c_2}$$

Nên ta có thể dùng N^* để ước tính kích thước quần thể:

$$N^* = \frac{c_1 c_2}{R}$$

Để tăng tính chính xác thì ta điều chỉnh công thức trên, để N^* trở thành một ước lượng không chệch của N

$$N^* = \frac{(c_1 + 1)(c_2 + 1)}{R + 1} - 1$$

$$Var(N^*) = \frac{(c_1 + 1)(c_2 + 1)(c_1 - R)(c_2 - R)}{(R + 1)^2(R + 2)}$$

Vậy khoảng tin cậy 95% của kích cỡ quần thể là

$$\left(N^* - 1.96\sqrt{Var(N^*)}, N^* + 1.96\sqrt{Var(N^*)}\right)$$

Hãy viết hàm mô phỏng phương pháp trên. Hàm gồm có:

- Input: Kích cỡ thật sự của quần thể, số cá thể lấy lần 1, số cá thể lấy lần 2.
- Output: Kích cỡ quần thể được ước tính bằng phương pháp C-R; khoảng tin cậy 95%.