

**BÀI TẬP**

# **THỐNG KÊ ỨNG DỤNG**

***Buổi 1***

**ĐỒ LÊN**

## ***YÊU CẦU BUỔI HỌC:***

- 1) Cài phần mềm R và sử dụng các chức năng cơ bản.**
- 2) Vẽ các bảng tần số bằng R.**
- 3) Vẽ các loại đồ thị bằng R.**

## Ví dụ

Giả sử ta có dữ liệu về số môn còn nợ của 30 bạn sinh viên lớp TH04.1 như sau:

5	1	3	2	0	1	2	4	2	4	4	2	5
5	1	5	4	4	2	5	5	0	1	0	5	0
0	0	3	5									

**Yêu cầu: Vẽ bảng tần số**

## Ví dụ

Giả sử có tập dữ liệu về tuổi của người lao động trong một phân xưởng như sau:

48	30	35	31	21	28	34	43	36	45	41
33	47	47	30	47	44	45	32	46	47	23
30	23	49	20	24	20	40	50			

**Yêu cầu: 1) Phân tổ (theo 2 cách)**

**2) Vẽ bảng tần số**

## Lập bảng tần số bằng R

Trường hợp dữ liệu cần lập bảng tần số là định tính hoặc định lượng nhưng số các biểu hiện không nhiều, ta lập bảng tần số, bảng tần suất, bảng tần số tích lũy, bảng tần suất tích lũy bằng các lệnh sau:

<code>table(x)</code>	bảng tần số
<code>prop.table(table(x))</code>	bảng tần suất
<code>cumsum(table(x))</code>	bảng tần số tích lũy
<code>cumsum(prop.table(table(x)))</code>	bảng tần suất tích lũy

trong đó

x: vec tơ dữ liệu

## Ví dụ

Bảng số liệu về số hộp sữa bán được của một cửa hàng

	9	10	10	11	12	12	
10		11	10	9	11	9	12
	9	12	10	9	12	10	11
	11	12	11	10	12	10	11

***Yêu cầu: Lập bảng tần số***

Khi lập bảng dữ liệu định lượng nhiều biểu hiện, trước khi lập bảng tần số, ta tiến hành phân tổ bằng lệnh sau:

```
cut(x, breaks, right=TRUE, include.lowest=FALSE)
```

trong đó

x

: vector dữ liệu

breaks: một vector gồm các điểm chia tổ hoặc một số cho biết số tổ cần chia

right: TRUE (FALSE) nếu các tổ có dạng (a,b] ([a,b))

include.lowest: TRUE (FALSE) đưa giá trị nhỏ nhất (lớn nhất) vào tổ đầu tiên (cuối cùng) khi right=T (right=F).

## Ví dụ

Giả sử ta có bảng dữ liệu sau về giới tính và nghề nghiệp của 10 người được điều tra.

<i>TT</i>	<i>Giới tính</i>	<i>Nghề nghiệp</i>
1	Nam	CongNhan
2	Nam	QuanLy
3	Nam	CongNhan
4	Nam	NVVP
5	Nam	NVVP
6	Nam	NVVP
7	Nu	QuanLy
8	Nam	CongNhan
9	Nu	NVVP
10	Nam	CongNhan



## Lệnh Hist

```
hist(x, breaks = "Sturges", include.lowest = TRUE, right = TRUE, col =  
, border = , labels = FALSE, main = "", xlab = "", ylab = "")
```

trong đó

- x: vector dữ liệu dạng số dùng để vẽ biểu đồ,
- breaks: như lệnh cut
- right: như lệnh cut
- include.lowest: như lệnh cut
- xlab: tên trục ngang
- ylab: tên trục dọc
- main: tên biểu đồ
- col: màu sắc các thanh

## Lệnh plot

```
plot(x, y, type = "b", main = "", sub = "", xlab = "", ylab = "")
```

trong đó

- x, y: hoành độ, tung độ của các điểm trong đồ thị,
- type = "b": chỉ kiểu vẽ trong đồ thị là điểm nối các đoạn thẳng
- xlab: tên trục ngang
- ylab: tên trục dọc
- main, sub: tên chính, tên phụ của đồ thị
- col: màu sắc các thanh

## Example

Giả sử ta có điểm quá trình môn thống kê của 30 sinh viên như sau:

6.9 3.2 6.4 7.2 7.2 4.8 6.1 7.4 7.8 8.0 6.0 6.6 6.5 7.2 2.2  
8.7 7.3 7.8 4.7 4.8 5.7 6.1 6.5 8.7 5.0 7.8 6.5 5.6 7.9 8.8

***Yêu cầu: Vẽ biểu đồ thân lá***

## Hàm stem()

```
stem(x, scale = 1)
```

trong đó

- x: vector dữ liệu dạng số;
- scale: tham số điều chỉnh chiều dài của biểu đồ.

## Ví dụ

Theo ước tính, lượng khách quốc tế đến Việt Nam vào tháng 2 năm 2015:

Phương tiện	Lượng khách	Tỉ lệ	Góc (độ)
Đường biển	7 021	0.009287	3.34
Đường bộ	138 146	0.182733	65.78
Đường không	610 834	0.80798	290.87

**Yêu cầu: Vẽ biểu đồ hình tròn + biểu đồ thanh**

Cho bảng số liệu về diện tích các vùng trong lãnh thổ Việt Nam.

Vùng	Diện tích ( $m^2$ )	Tỉ lệ (%)	Góc (độ)
ĐB Sông Hồng	14862.4	4.50	16.15
Đông Bắc Bộ	64025.2	19.3	69.59
Tây Bắc Bộ	37533.9	11.3	40.80
Bắc Trung Bộ	51551.9	15.6	56.03
Nam Trung Bộ	33166.0	10.0	36.05
Tây Nguyên	54659.6	16.5	59.41
Đông Nam Bộ	34807.8	10.5	37.83
ĐB Sông Cửu Long	40604.8	12.3	44.13

***Yêu cầu: Vẽ biểu đồ hình tròn + biểu đồ thanh***

## Hàm pie()

`pie(x, labels = names(x), col = NULL, border = NULL, lty = NULL, main = NULL)`

trong đó

- `x`: vector dữ liệu dạng số thể hiện giá trị của mỗi hình quạt trong biểu đồ;
- `labels`: tham số chỉ tên những hình quạt trong biểu đồ;
- `col`: tham số chỉ màu của các hình quạt;
- `boder`: tham số chỉ màu của các đường danh giới giữa các hình quạt;
- `main`, `sub`: tham số chỉ tiêu đề và tiêu đề phụ của biểu đồ.

## Hàm barplot()

```
barplot(height, names.arg = NULL, beside = FALSE, horiz = FALSE, col = "", border = "", main = "", sub = "", xlab = "", ylab = "")
```

trong đó

- height: vector hoặc ma trận dữ liệu dạng số thể hiện giá trị của mỗi hình quạt trong biểu đồ;
- name.arg: tham số chỉ tên viết dưới mỗi thanh hoặc nhóm thanh trong biểu đồ;
- beside: tham số dạng logic, nếu = TRUE thì các cột của biểu đồ được vẽ chồng lên nhau, nếu = FALSE thì các cột được vẽ cạnh nhau;
- horiz: tham số dạng logic, nếu = TRUE thì các cột được vẽ song song với trục nằm ngang với cột đầu tiên ở dưới cùng, nếu = FALSE thì các cột được vẽ vuông góc với trục nằm ngang với cột đầu tiên nằm bên trái;



## Example

Giả sử ta có thu nhập và chi tiêu của 20 người như sau.

TN	19	17	18	23	14	11	17	26	19	19	8	26
CT	12	7	10	10	6	5	10	15	11	9	3	11

***Yêu cầu: Vẽ biểu đồ tán xạ***

## Hàm plot()

```
plot (x, y, type = "p", main = "", xlab = "", ylab = "", bty = "l")
```

trong đó

- x,y: vector tọa độ của biến định lượng thứ nhất, thứ hai;
- bty="l": tham số cho kiểu bao quanh biểu đồ giống hai trục tọa độ Oxy.











