

# Finance Analytics: R 프로그래밍 기초 PART 2

권태연

한국외대 국제금융학과

# Library의 설치

R-base에 모든 함수를 넣지 못한다. 특정 함수를 실행하기 위해서는 특정 library를 설치하여야 한다.

- (ex) 다변량 정규분포로부터 랜덤 sample추출하기
- (ex) 역감마분포로부터 랜덤 sample추출하기

# 흐름 제어문

## 조건문 사용하기 : if() & ifelse() 문

R에서 if()문은 어떤 명령문이 실행되어야 하는지를 제어해 주는 매우 유용한 방법이다.

괄호 안 (condition) 조건이 만족하면, 뒤 따르는 중괄호 안 {statement}의 명령문을 실행한다.

- 조건부 실행 : if and ifelse

- if문은 길이가 1인 논리형 벡터에 적용가능하다.

```
if(condition 1) {  
  statement 1  
} else if (condition 2) {  
  statement 2  
} else {  
  statement 3  
}  
ifelse(condition, TURE, FALSE)
```

## 예제 if() & ifelse() 문

```
> x <- 3
> if (x > 2) y <- 2 * x else y <- 3 * x
> y <- ifelse(x>2, 2*x, 3*x)
> x = -3
> if (x < -1) {
+   y = -1
+ } else if (x < 0) {
+   y = 0
+ } else {
+   y = 1
+ }
> y
```

# 루프(loop) 사용하기

통계적인 계산에서는 특정한 작업을 주어진 수만큼 반복하거나 어떤 특정 조건을 만족하는 한 어떤 작업을 계속 수행하도록 해야 하는 경우가 존재한다. 이러한 상황을 다루기 위해 R에서는 다양한 함수를 제공한다.

- `for(name in vector) {commands}` : `name`이라는 이름의 변수가 `vector`의 각 구성요소의 값을 순차적으로 가지게 됨, 매번 중괄호 안의 `command`를 실행
- `while(condition) {commands}` : execute a block of commands until the condition is no longer satisfied
- `repeat {statements}`
- `break` : 즉각적으로 루프를 빠져나가도록 해 준다.
- `next` : 즉각적으로 루프의 다음 단계로 넘어가도록 한다.

## 연습문제 for()

x벡터값에 행번호를 넣어보자.

```
>x<-rep(1,100)
>for ( i in 1:100){
+ x[i]<-i
+ }
> x
```

## 연습문제 for()

1부터 100까지 정수를 차례대로 더해보자

```
>x<-rep(1,100)
>for ( i in 2:100){
+ x[i]<-x[i-1]+i
+ }
> x
```



## 실습 for()

피보나치 수열: 수학에서 잘 알려진 수열 중 하나로써 다음과 같은 점화식이다.

$$x_n = x_{n-2} + x_{n-1}, \quad n = 3, 4, 5, \dots,$$

$x_1 = 1$ 과  $x_2 = 1$ 는 초기에 주어져야 한다.

## 연습 : For and IF

- BMI data를 이용하여 체질량지수(BMI:Body Mass Index)를 계산하고 이를 이용하여 비만의 정도를 구분하고자 한다.
  - 다음은 6명의 몸무게weight= (60,72,57,90,95,72)와
  - 키 height=(1.75, 1.80, 1.65, 1.90, 1.74, 1.91)이다.
  - $BMI = Weight(kg) \div [Height(m)]^2$  (예:  $23.3 = 80.5 / 1.86^2$ ) :  
(소수점 첫째 자리까지)
  - Body Mass Index(BMI) 구분
    - BMIID=1 if  $\leq 18.5$
    - BMIID=2 if  $18.5 < BMI \leq 24.9$
    - BMIID=3 if  $24.9 < BMI \leq 29.9$
    - BMIID=4 if  $29.9 < BMI \leq 30$
    - BMIID=5 if  $30 < BMI > 30$
- BMIID=1,5만으로 이루어진 data1과 BMIID=3,4만으로 이루어진 data2를 생성하는 프로그램을 작성하시오.

## 실습: pizzadata

pizzadata.csv의 변수 age1에는 실제 개인의 연령이 기록되어 있다. 다음과 같은 연령그룹 변수를 만들고 이를 이용하여 회귀분석을 하기 위해 더미변수를 생성하시오. 그리고 생성한 연령그룹변수에 대한 더미변수와 income변수를 이용하여 회귀분석을 하세요.

- AGE\_G=1, 18-25세
- AGE\_G=2, 26-35세
- AGE\_G=3, 36세 이상

- ✓ For문과 IF문을 사용하여 만드는 방법
- ✓ For문과 IF문을 사용하지 않고 만드는 방법

✓ 위와 같은 분석에서 연령을 설명변수로 고려할 때, 연령그룹으로 더미변수화 하여 고려하는 것과 연령을 정량변수로 고려하는 것의 차이는 무엇인지 생각해보자.

1. 연령을 정량변수로 하였을 때 연령에 대한 회귀계수의 의미
2. 연령을 더미변수화 하였을 때 회귀계수의 의미

## glm() 함수

생성된 AGE\_G변수를 이용하여 더미변수를 만들지 않고 바로 회귀분석하기 위해서는 lm()대신 glm()함수를 사용한다.  
glm()함수의 사용방법은 lm()함수와 동일하다.

```
> model3<-glm(pizza~income+as.factor(AGE_G),data=data1)
> summary(model3)
```

## 예제 while()

1부터 5보다 작은값까지 print 하시오.

```
> ## WHILE##  
> x <- 1  
> while(x < 5) {  
+   x <- x+1  
+   print(x)}  
[1] 2  
[1] 3  
[1] 4  
[1] 5
```

## 예제: break()

Let's break the loop when  $x=3$ :

```
> x <- 1
> while(x < 5) {
x <- x+1
if (x == 3) break
print(x)
}
[1] 2
```

## 예제: next()

Let's skip one step when  $x=3$ :

```
> x <- 1
> while(x < 5) {
x <- x+1
if (x == 3) next
print(x)}
[1] 2
[1] 4
[1] 5
```



# 사용자 정의 함수

# 사용자 정의 함수

함수는 특정한 동작을 일반적인 형태로 정의해 놓은 지시문의 집합이다.  
반복적으로 사용되는 루틴이 함수로 제공되면 편리하다.  
R은 다른 프로그래밍 언어와 같이 사용자가 직접 함수를 정의하여 사용할 수 있다.

- 함수의 일반적인 형태는 다음과 같다.

```
f=function(x,y) { expression involving x, y, ... }
```

- f: 함수명.
  - function: 명령문.
  - x, y: 입력인자 목록. 목록은 비어 있을 수도 있다.
  - {}: 단일 또는 여러 개의 지시문들.
  - > f(x,y)
- 입력인자(인수)는 기본값을 가질 수 있다.
    - f=function(x,y=1): y는 기본값이 1이 된다는 뜻이다.  
> f(x)
    - 기본값을 사용하고 싶지 않을 경우 함수 호출 시 해당 인수를 명시해 줘야 한다.  
> f(x,y=2)

# 사용자 정의 함수 :반환값

- 함수의 반환값은 어떤 R 객체도 가능하다. (상수, 행렬 벡터)등
- `return( )` 함수를 사용하여 값을 반환한다.
- 반환값이 여러개 있을 때는 `list( )` 를 사용한다.

## 예제: 사용자 정의 함수

예제: 다음의 함수

$$y = x^2 + 10$$

를 작성하고  $x=10$ 일때  $y$ 값을 구해보자.

```
> TY<-function(x){  
+ z<-x^2+10  
+ return(z)}  
> TY(10)  
[1] 110
```

# 사용자 정의 함수 실습

매년 고정된 이자수익이  $i$ 인 은행계좌에 매년  $R$ 달러씩 저금을 한다고 생각해 보자.  $n$ 년후 내 계좌의 잔고를 계산하는 함수를 작성하여 보자. 만약 10년간 매년 400 달러 씩 저축하고 연 이자가 5%라면 누적 총계는 얼마인지 구하라. (Annuity)