

# R입문 기말고사

이정민(202115820)

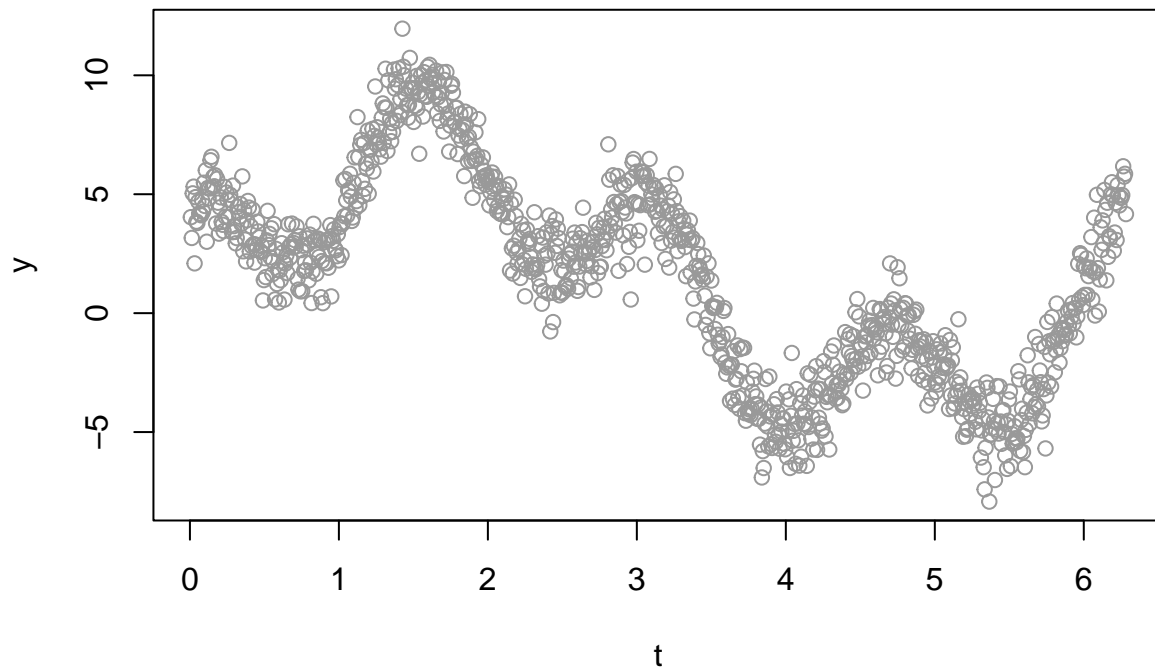
12/20/2021

## 1번 문제풀이

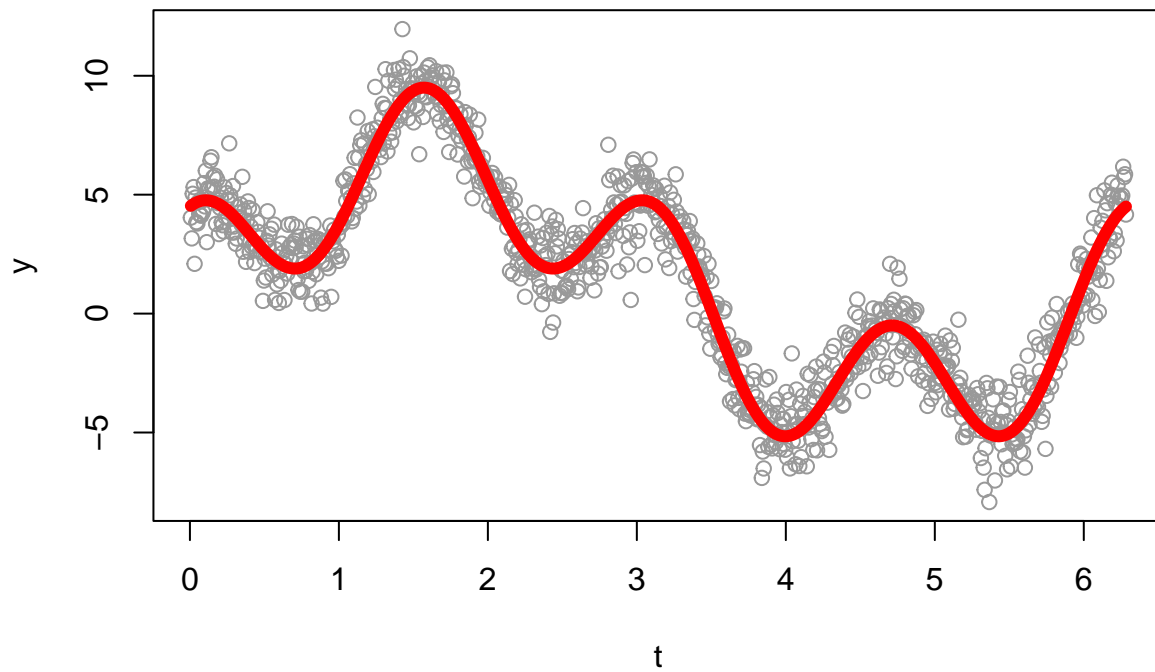
```
#1.
t=c()
for(i in 1:1000){
  t[i] = 2*pi*i/1000
}
#1.1
A_<-rnorm(1000)
```

```
#1.2
x1 = c()
x2 = c()
for(i in 1:1000){
  x1[i]=sin(t[i])
  x2[i]=cos(4*t[i])
}
```

```
#1.3
y=c()
for(i in 1:1000){
  y[i]=1.5+5*x1[i]+3*x2[i]+A_[i]
}
plot(t,y,col='gray60')
```



```
#1.4
X=cbind(1,x1,x2)
#1.5
Beta=rbind(1.5,5,3)
XB=X%*%Beta
plot(t,y,col='gray60')
lines(t,XB, col='red', lwd=6)
```



#1.6

```
xt=solve(t(X)%*%X)
xy=t(X)%*%y
Beta_=xt%*%xy
Beta_
```

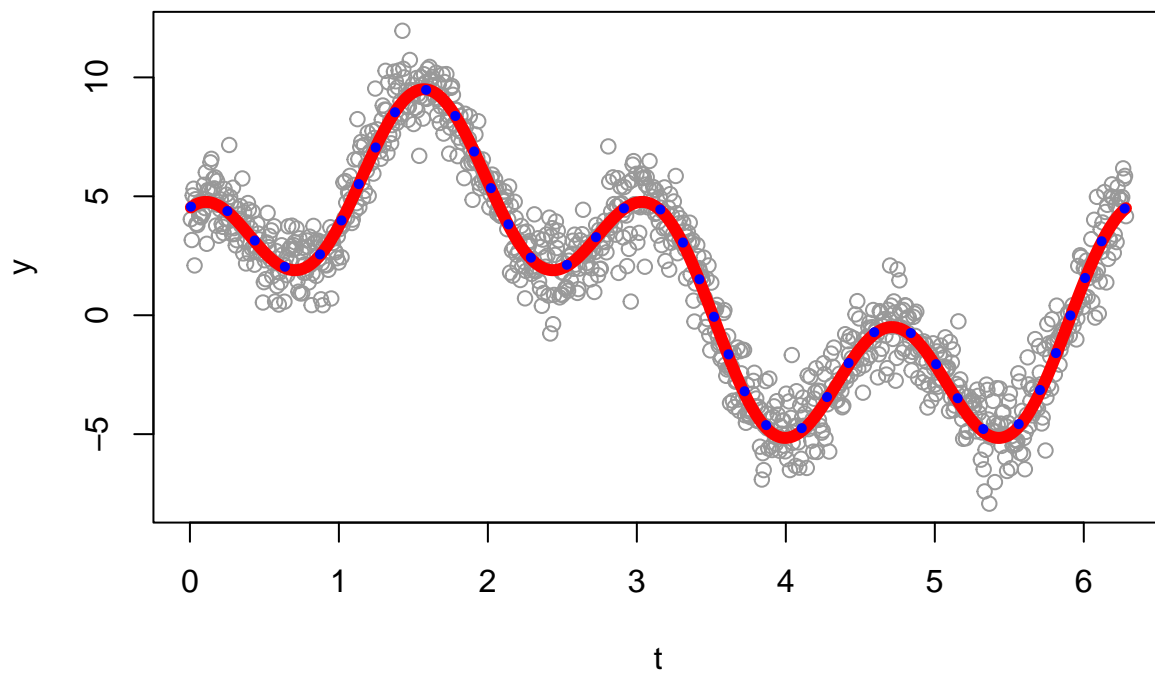
```
##      [,1]
## 1.558046
## x1 4.949198
## x2 2.972754
```

Beta

```
##      [,1]
## [1,] 1.5
## [2,] 5.0
## [3,] 3.0
```

#1.7

```
xb=X%*%Beta_
plot(t,y,col='gray60')
lines(t,XB,col='red',lwd=6)
lines(t,xb,col='blue',lty=3,lwd=5)
```



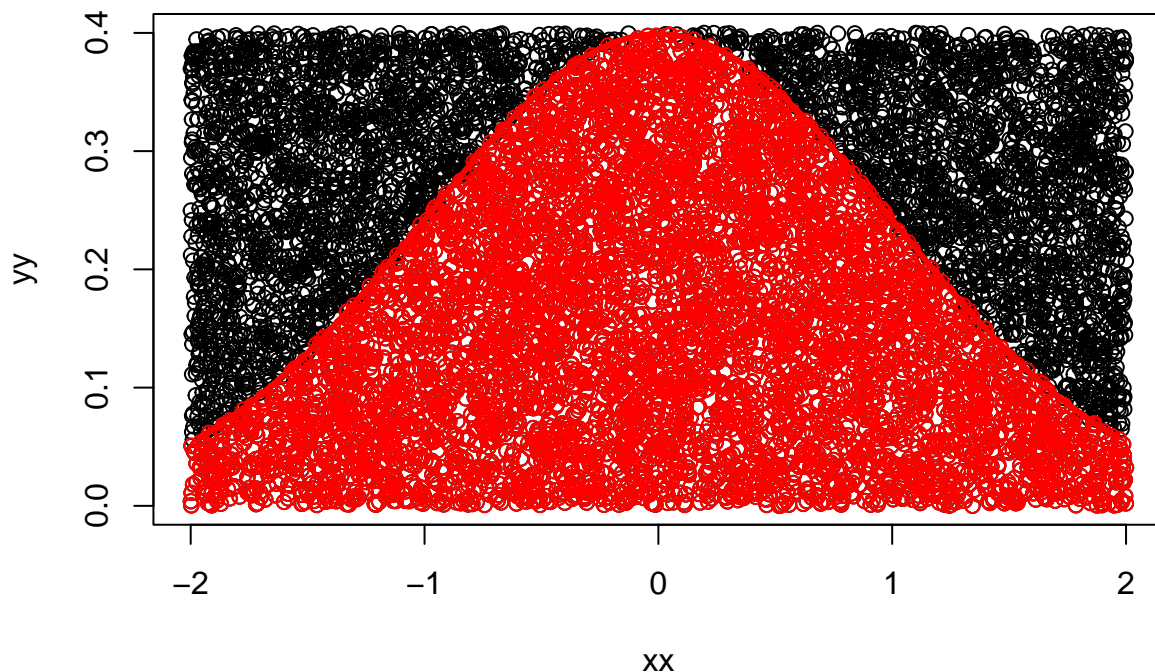
## 2번 문제풀이

### #2.1문제풀이

```
x=seq(from=-1.96, to=1.96, by=0.01)
y=(1/sqrt(2*pi))*exp(-(1/2)*x^2)
xx=runif(10000,min=-2,max = 2)

yy=runif(10000)*0.4
plot(xx,yy)
lines(x,y,col='red',lwd=3)

test = function(xx,yy){
  yy < (1/sqrt(2*pi))*exp(-(1/2)*xx^2)
}
tst = c()
for (i in 1:10000) tst[i] = test(xx[i],yy[i])
points(xx[tst],yy[tst],col='red')
```



```
sum(tst)
```

```
## [1] 5897
```

##  $s/1.6 = 5995/10000$  이므로 넓이는 0.9592이다. 대략 넓이는 1이다

### #2.2문제풀이

```
rn = rnorm(1000)
id=ifelse(-1.96<rn & rn<1.96,1,0)
sum(id)
```

## [1] 953

### 3번 문제풀이

#내림차순

```
k=c()
for(i in 1:10000){
  chance=2
  success=0
  while(TRUE){
    if(chance!=1 && chance!=0){
      if(runif(1)>0.5){
        success=success+1
      }
      else{
        chance=chance-1
        success=success+1
      }
    }

    else if(chance ==1){
      if(runif(1)<0.95){
        success=success+1
      }
      else{
        chance=chance-1
        success=success+1
      }
    }

    else{
      if(runif(1)>0.5){
        success=success+1
      }
      else {
        chance=chance-1
      }
    }
  }

  if(chance==-1){
    k[i]=0
    break
  }
  if(success==20){
    k[i]=1
    break
  }
}
```

```

    }
}

print(sum(k==1)/10000)

```

```
## [1] 0.4336
```

```
# 0.4421
```

```
#오름차순
```

```

k=c()
for(i in 1:10000){
  chance=7
  succeess=0
  while(TRUE){
    if(chance!=0){
      if(runif(1)>0.5){
        succeess=succeess+1
      }

      else{
        chance=chance-1
        succeess=succeess+1
      }
    }
    else{
      if(runif(1)>0.5){
        succeess=succeess+1
      }

      else{
        chance=chance-1
      }
    }
  }

  if(chance== -1){
    k[i]=0
    break
  }
}

```

```

    if(succeess==20){
      k[i]=1
      break
    }
  }
}

print(sum(k==1)/10000)

```

```
## [1] 0.1252
```

```
#0.1362
```

```
#Type A일때 살아남을 확률이 더 높다.
```

```
library(tidyverse)
```

```
## -- Attaching packages ----- tidyverse 1.3.1 --
```

```
## v ggplot2 3.3.5      v purrr  0.3.4
```

```
## v tibble  3.1.6      v dplyr  1.0.7
```

```
## v tidyr   1.1.4      v stringr 1.4.0
```

```
## v readr   2.1.1      v forcats 0.5.1
```

```
## -- Conflicts ----- tidyverse_conflicts() --
```

```
## x dplyr::filter() masks stats::filter()
```

```
## x dplyr::lag()    masks stats::lag()
```

```
df=read_csv('https://raw.githubusercontent.com/guebin/2021IR/master/_notebooks/covid19.c
```

```
## Rows: 12294 Columns: 5
```

```
## -- Column specification -----
```

```
## Delimiter: ","
```

```
## chr (1): prov
```

```
## dbl (4): year, month, day, cases
```

```
##
```

```
## i Use `spec()` to retrieve the full column specification for this data.
```

```
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```



## 4번 문제풀이

#4.1

```
SS_2020 = df[df$year == 2020, ]  
sum(SS_2020$cases)
```

```
## [1] 60726
```

#60726명

```
SS_2021 = df[df$year == 2021, ]  
sum(SS_2021$cases)
```

```
## [1] 396886
```

#396886명

#답 = 60726+396886 = 457612명

```
day15 = SS_2020[SS_2020$month == 2 & SS_2020$day <= 15 & SS_2020$day >= 1, ]  
tapply(day15$cases, day15$prov, sum)
```

```
## 강원 검역 경기 경남 경북 광주 대구 대전 부산 서울 세종 울산 인천 전남 전  
북 제주
```

```
##      0      0      9      0      0      2      0      0      0      5      0      0      0      1      0      0
```

```
## 충남 충북
```

```
##      0      0
```

#경기지역이 가장 많은 확진자수가 나왔다.

```
day29 = SS_2020[SS_2020$month == 2 & SS_2020$day <= 29 & SS_2020$day >= 15, ]  
tapply(day29$cases, day29$prov, sum)
```

```
## 강원 검역 경기 경남 경북 광주 대구 대전 부산 서울 세종 울산 인천 전남 전  
북 제주
```

```
##      7      0      65      59      472      7      2055      13      75      62      1      17      5      1      4      2
```

```
## 충남 충북
```

```
##      48      10
```

#대구지역이 가장 많은 확진자수가 나왔다.