

# R입문 기말고사

조운호(201822032)

```
library(tidyverse)
```

## 1번

```
### 1)
e=c()
for(i in 1:1000){
  e[i] = rnorm(1)
}
head(e)
```

```
## [1]  0.9497786  0.1874779  0.6281326 -1.9025805  0.7670813  0.5950860
```

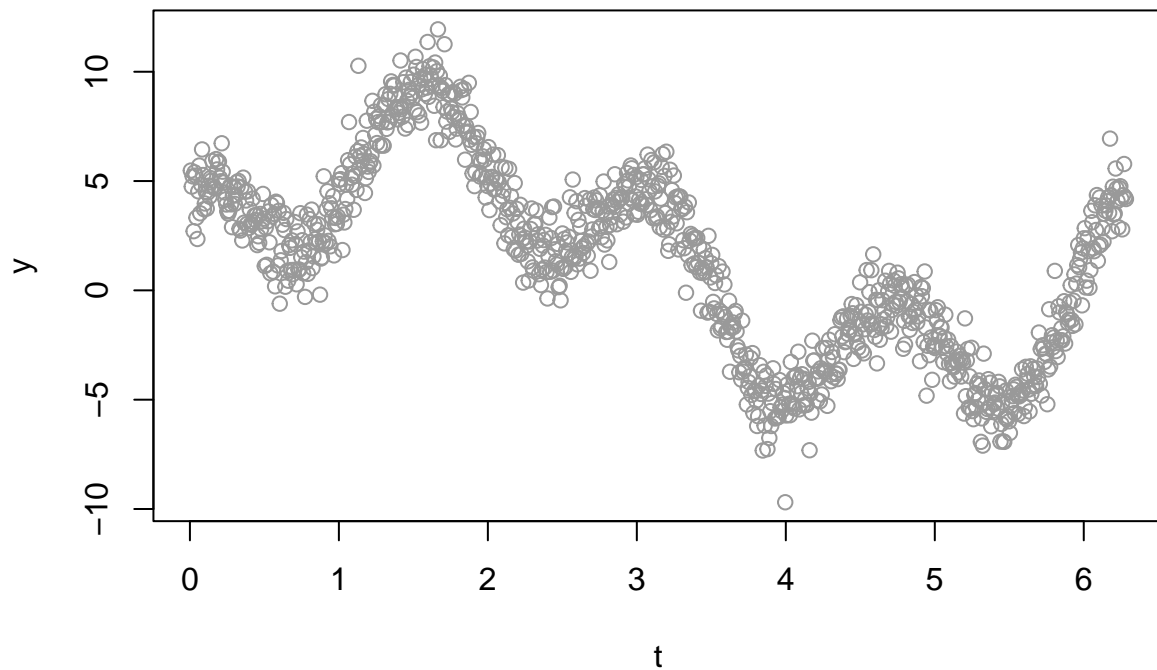
```
### 2)
t=c()
for(i in 1:1000){
  t[i]=(2*pi/1000)*i
}
x1=c()
for(i in 1:1000){
  x1[i]=sin(t[i])
}
x2=c()
for(i in 1:1000){
  x2[i]=cos(t[i]*4)
}
head(x1)
```

```
## [1] 0.006283144 0.012566040 0.018848440 0.025130095 0.031410759 0.037690183
```

```
head(x2)
```

```
## [1] 0.9996842 0.9987370 0.9971589 0.9949510 0.9921147 0.9886517
```

```
### 3)
y=c()
for(i in 1:1000){
  y[i]= 1.5 +5*x1[i] +3*x2[i] + e[i]
}
plot(t,y,col="gray60")
```



```
### 4)
```

```
X=cbind(rep(1,1000),x1,x2)
head(X)
```

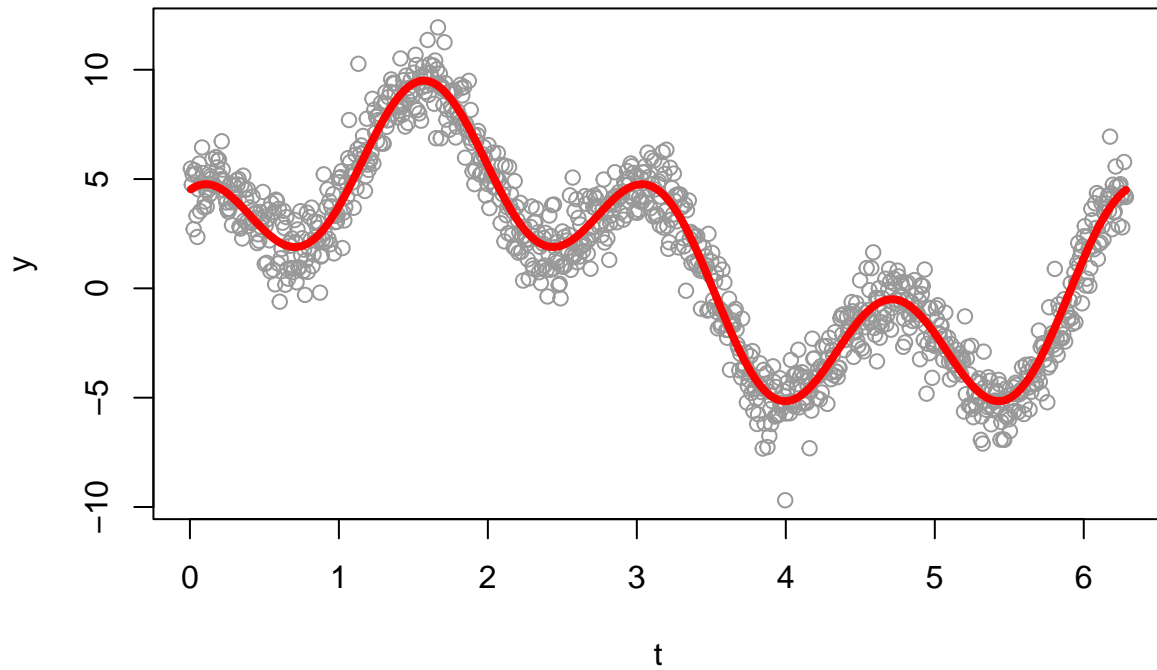
```
##           x1           x2
## [1,] 1 0.006283144 0.9996842
## [2,] 1 0.012566040 0.9987370
## [3,] 1 0.018848440 0.9971589
## [4,] 1 0.025130095 0.9949510
## [5,] 1 0.031410759 0.9921147
## [6,] 1 0.037690183 0.9886517
```

```
### 5)
```

```
beta = rbind(1.5,5,3)
beta
```

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

```
Xbeta=array(X %*% beta)
plot(t,y,col="gray60")
lines(t,Xbeta,col="red",lwd=4)
```

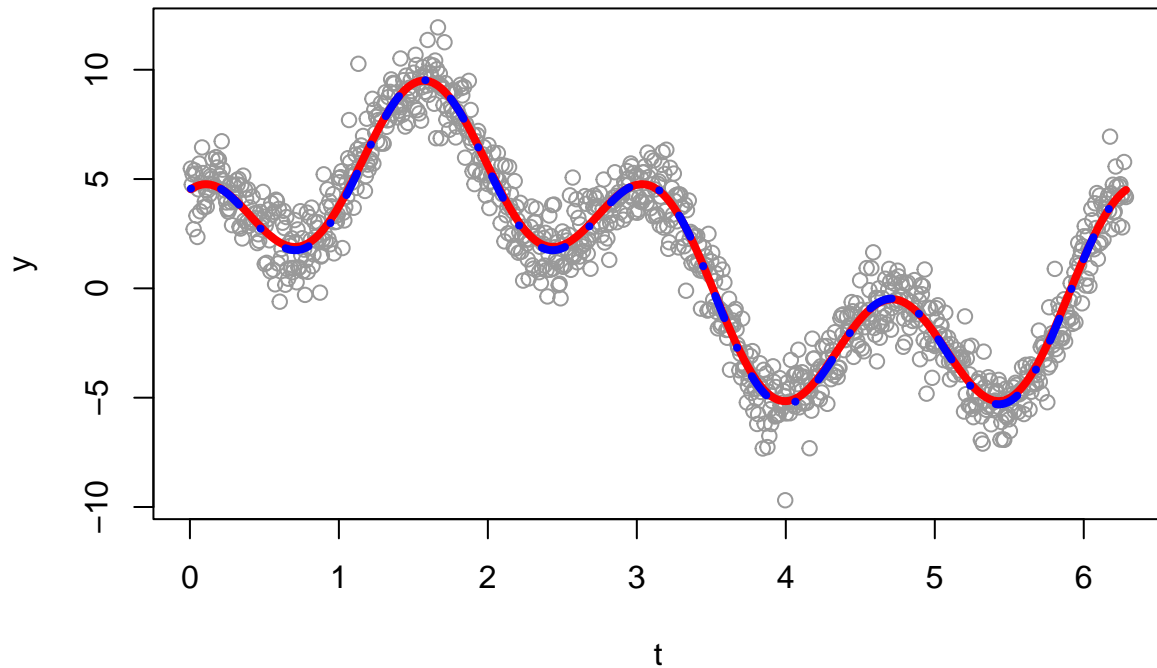


```
### 6)
beta_hat=solve(t(X) %*% X) %*% t(X) %*% y
beta_hat
```

```
##      [,1]
## 1.439189
## x1 4.997604
## x2 3.089615
```

$\hat{\beta}$ 은  $\beta$ 와 거의 유사하다.

```
### 7)
Xbeta_hat=array(X %*% beta_hat)
plot(t,y,col="gray60")
lines(t,Xbeta,col="red",lwd=4)
lines(t,Xbeta_hat,col="blue",lty=4,lwd=4)
```

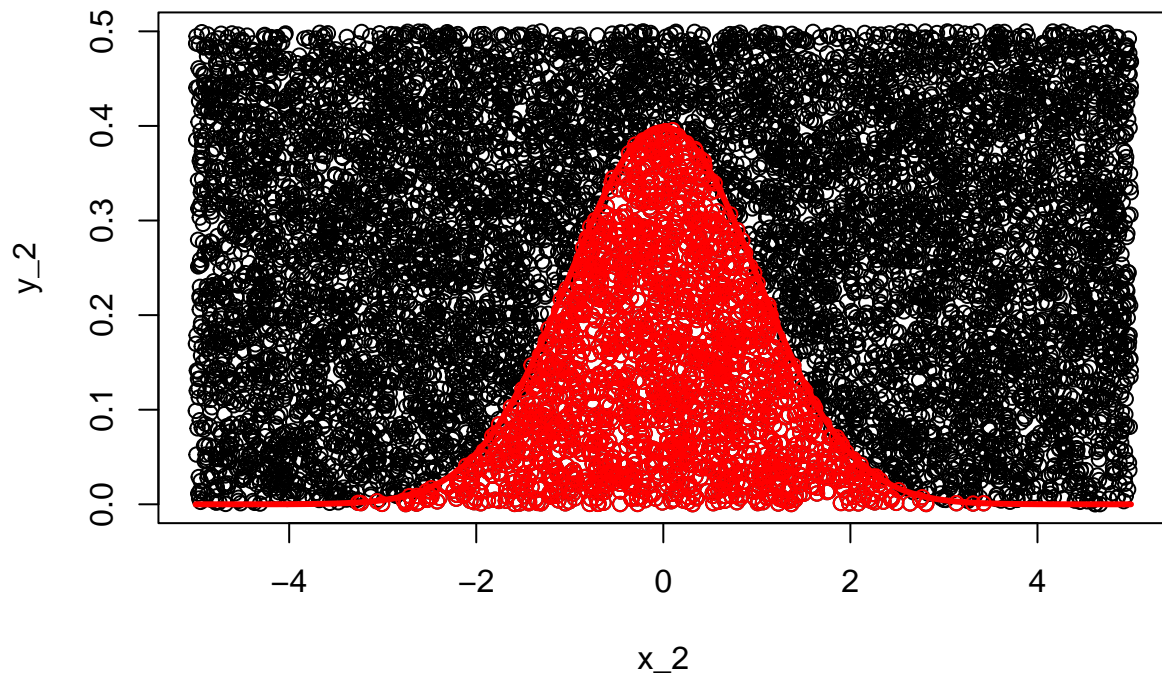


## 2번

```
### 1)
x_2=runif(10000,-5,5)
y_2=runif(10000)*0.5
x_21=seq(from=-5, to=5, by=0.01)
y_21=(1/sqrt(2*pi))*exp((-1/2)*(x_21**2))
check=function(x,y){
  y<(1/sqrt(2*pi))*exp((-1/2)*(x**2))
}
tst = c()
for (i in 1:10000) tst[i] = check(x_2[i],y_2[i])
sum(tst)
```

## [1] 2013

```
plot(x_2,y_2)
lines(x_21,y_21,col='red',lwd=3)
points(x_2[tst],y_2[tst],col='red')
```



빨간 점들 중에서 -1.96~1.96 사이의 점의 개수 비율

```
sum(x_2[tst]>=-1.96 & x_2[tst]<=1.96) / sum(tst)
```

```
## [1] 0.9483358
```

```
### 2)
x_norm=c()
count=0
for(i in 1:1000){
  x_norm[i]=rnorm(1)
  if(x_norm[i]>=-1.96 & x_norm[i]<=1.96){
    count = count +1
  }
}
count
```

```
## [1] 957
```

### 3번

```
### type A 생존확률
nor1=c()
spec=c()
nor8=c()
count=c()
for(i in 1:10000){
  nor1[i]=sum(cumprod(rbinom(20,size=1,0.5))) # 일반인 10이 해결한 수
```

```

spec[i]=sum(cumprod(rbinom(20-nor1[i]-1,size=1,0.95))) # 장인이 해결한 수
if(spec[i]>=(20-nor1[i]-1-1)){
  count[i] = 1
}else{
  nor8[i]=sum(cumprod(rbinom(20-nor1[i]-1-spec[i]-1,size=1,0.5)))
  if(nor8[i]== (20-nor1[i]-1-spec[i]-1)){
    count[i] = 1
  }else{
    count[i] = 0
  }
}
}
sum(count)/10000

```

```
## [1] 0.4404
```

```
### type b 생존 확률
```

```

type_b=rbinom(10000, 20, 0.5)
sum(type_b<=7)/10000

```

```
## [1] 0.1296
```

type A가 많이 유리하다

## 4번

```

df=read_csv('https://raw.githubusercontent.com/guebin/2021IR/master/_notebooks/covid19.csv') head(df)
### 1) df %>% group_by(year) %>% summarise(sum=sum(cases))

```

## 2)

```

df %>% filter(year==2020 & month==2 & day<=15) %>% group_by(prov) %>% summarise(prov_cases=sum(cases))
%>% arrange(desc(prov_cases)) #### 경기

```

## 3)

```

df %>% filter(year==2020 & month==2 & day>15) %>% group_by(prov) %>% summarise(prov_cases=sum(cases))
%>% arrange(desc(prov_cases))

```

대구