R 기말고사 예상문제

오동운(202115809)

```
library(dplyr)
```

```
##
## 다음의 패키지를 부착합니다: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

1-1

```
epsilon <- rnorm(1000)
head(epsilon)
```

```
## [1] -0.36482847 -1.10978335 -0.05485415 -0.62773633  0.34113744 -0.06082655
```

1-2

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

head(x1)
```

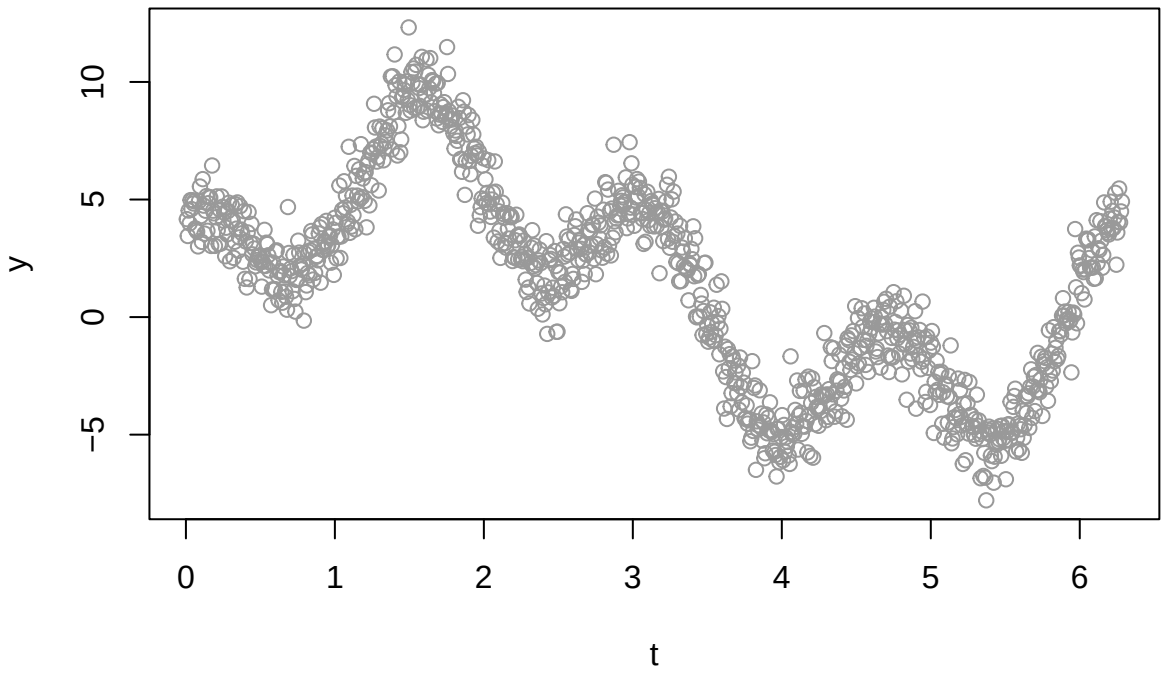
```
## [1] 0.006283144 0.012566040 0.018848440 0.025130095 0.031410759 0.037690183
```

```
head(x2)
```

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

1-3

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



1-4

```
X <- cbind(1,x1,x2)
head(X)
```

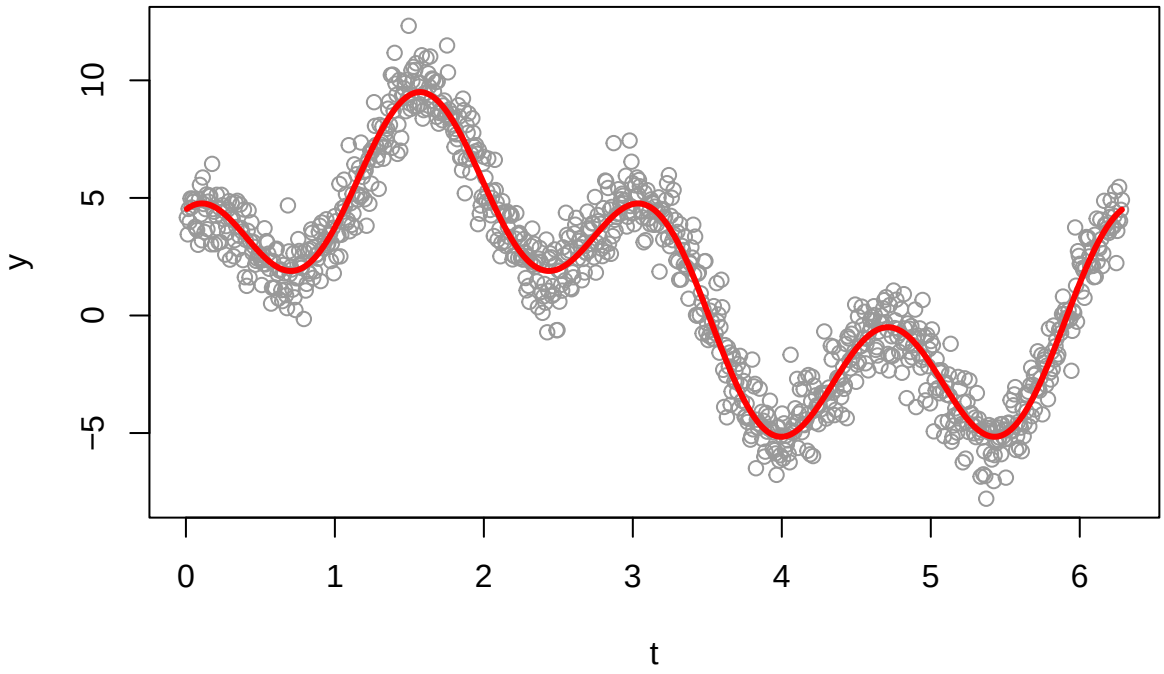
```
##                 x1        x2
## [1,] 1 0.006283144 0.9996842
## [2,] 1 0.012566040 0.9987370
## [3,] 1 0.018848440 0.9971589
## [4,] 1 0.025130095 0.9949510
## [5,] 1 0.031410759 0.9921147
## [6,] 1 0.037690183 0.9886517
```

1-5

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

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

```
XP1 = X %*% beta

plot(t,y,col='gray60')
lines(t,XP1,col='red',lwd='3')
```



1-6

```
beta
```

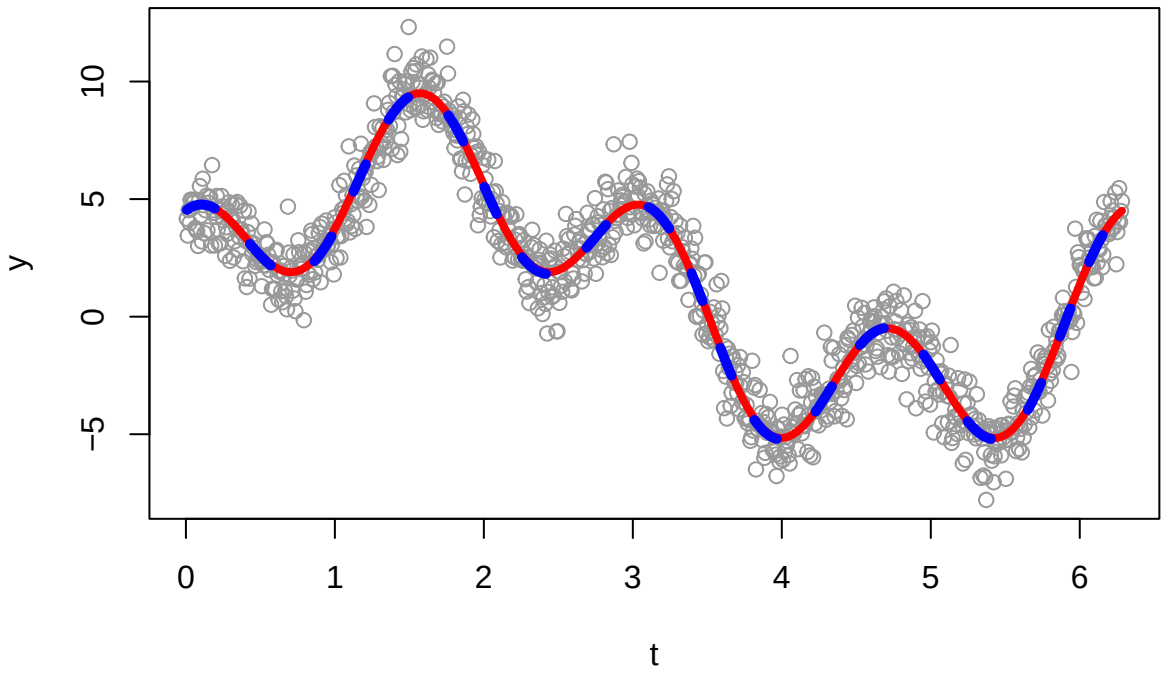
```
##      [,1]
## [1,]  1.5
## [2,]  5.0
## [3,]  3.0
```

```
beta2 = solve(t(X) %*% X )%*%t(X)%*%y
beta2
```

```
##        [,1]
##    1.469555
## x1 4.984470
## x2 3.041008
```

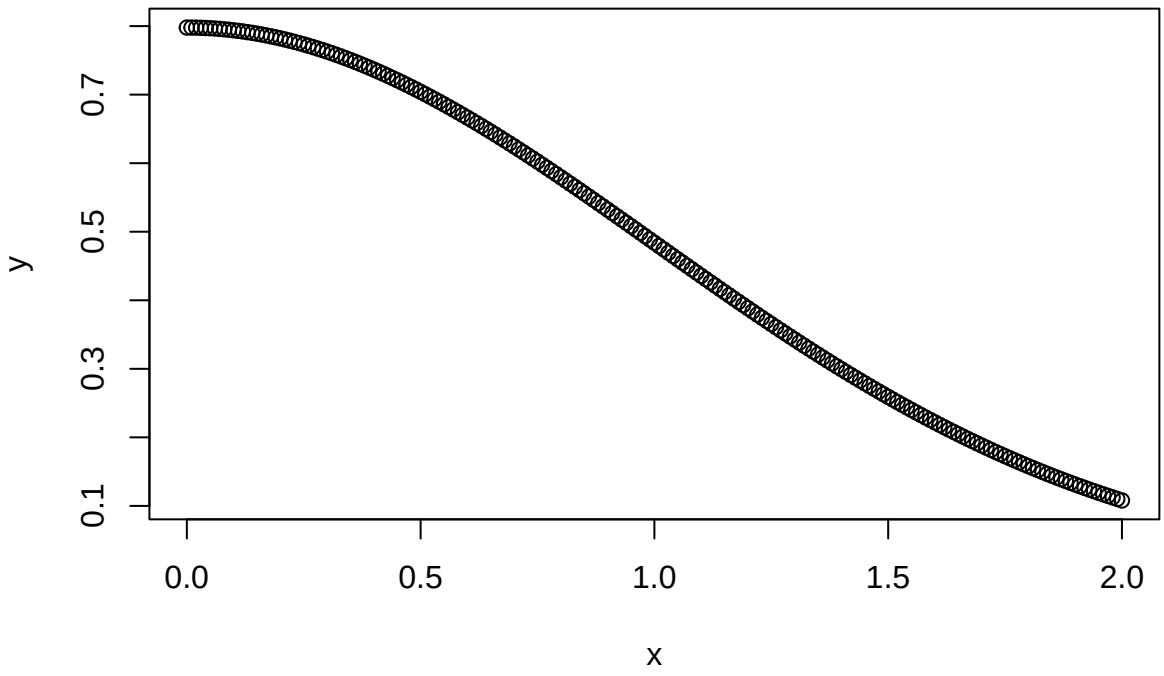
1-7

```
XP2 <- X %*% beta2
plot(t,y,col='gray60')
lines(t,XP1,col='red',lwd=4)
lines(t,XP2,col='blue',lty= 2,lwd=5)
```

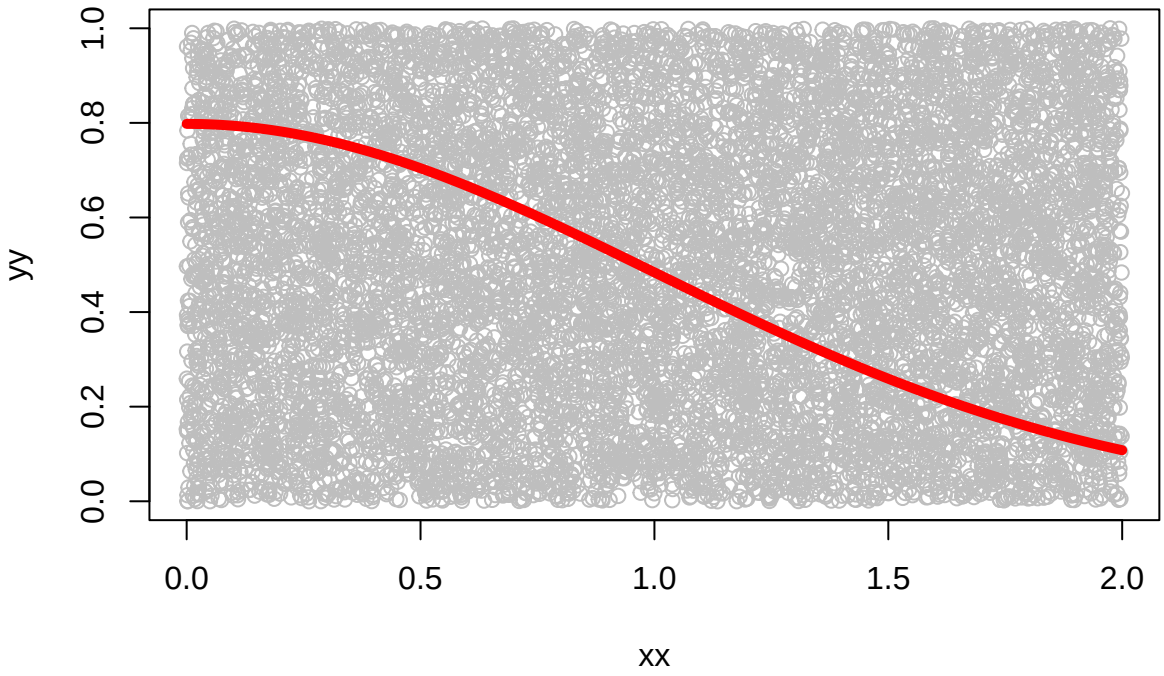


2-1

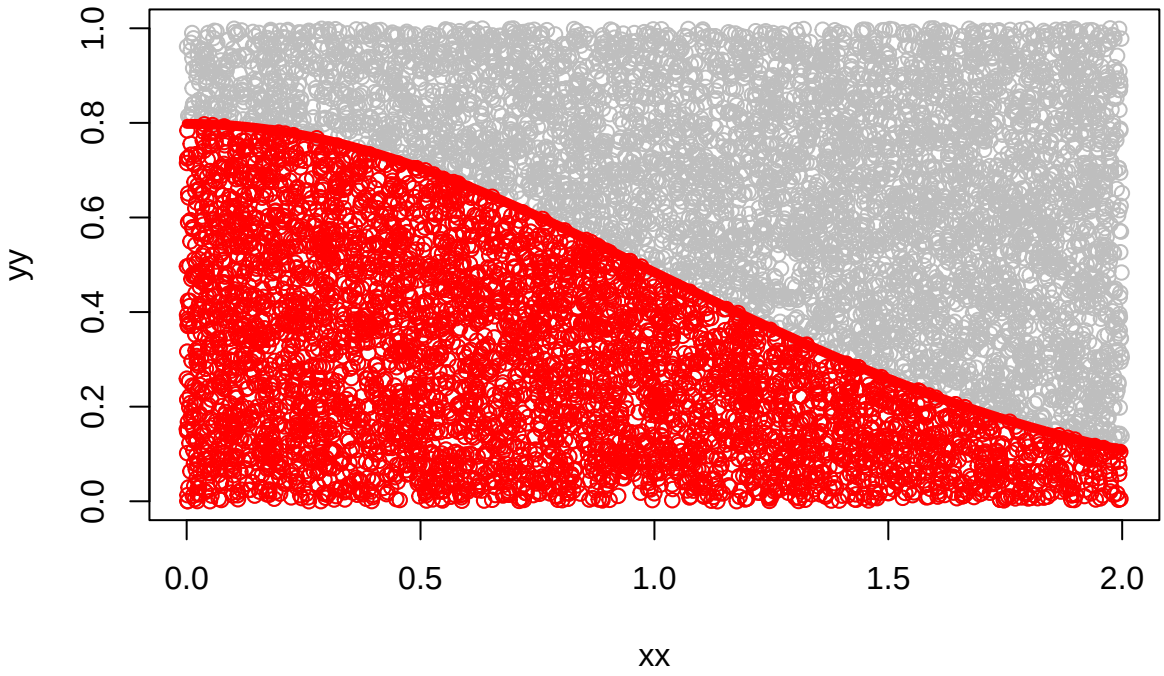
```
x = seq(from=0, to=2, by=0.01)
y = (2*exp(-1/2*x^2))/(2*pi)^(1/2)

plot(x,y)
```




```
xx = runif(n=10000,min=0,max=2)
yy = runif(n=10000,min=0,max=1)

plot(xx,yy,col='gray')
lines(x,y,col='red',lwd=5)
```




```
test<-function(xx,yy){
  yy < (2*exp(-1/2*xx^2)) /(2*pi)^(1/2)
}

tst<-c()
for(i in 1:10000){
  tst[i] <- test(xx[i],yy[i])
}

plot(xx,yy,col='gray')
lines(x,y,col='red',lwd=5)
points(xx[tst],yy[tst],col='red')
```




```
sum(tst)
```

```
## [1] 4780
```

2-2

```
A = rnorm(1000)
sum(abs(A) < 1.96)
```

```
## [1] 956
```

4-1

```
df=read.csv('https://raw.githubusercontent.com/guebin/2021IR/master/_notebooks/covid19.csv',fileEncodin
head(df)
```

```
##   year month day prov cases
## 1 2020     1  20 서울     0
## 2 2020     1  20 부산     0
## 3 2020     1  20 대구     0
## 4 2020     1  20 인천     1
## 5 2020     1  20 광주     0
## 6 2020     1  20 대전     0
```

```
A = df %>% filter(year==2020)



A %>% select(5) %>% sum
```

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

```
B = df %>% filter(year==2021)


B %>% select(5) %>% sum
```

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

4-2

```
DF = df %>% filter(year==2020 & month == 2 & day<=15)
DF %>% group_by(prov) %>% summarise(sum_cases = sum(cases))
```

```
## # A tibble: 18 x 2
##    prov  sum_cases
##    <chr>     <int>
##  1 강원          0
##  2 검역          0
##  3 경기          9
##  4 경남          0
##  5 경북          0
##  6 광주          2
##  7 대구          0
##  8 대전          0
##  9 부산          0
## 10 서울          5
## 11 세종          0
## 12 울산          0
## 13 인천          0
## 14 전남          1
## 15 전북          0
## 16 제주          0
## 17 충남          0
## 18 충북          0
```

```
"경기이다."
```

```
## [1] "경기이다."
```

4-3

```
DF2 = df %>% filter(year==2020 & month == 2 & day%in%c(16:29))
DF2 %>% group_by(prov) %>% summarise(sum_cases = sum(cases))
```

```
## # A tibble: 18 x 2
##    prov  sum_cases
##    <chr>     <int>
##  1 강원          7
##  2 검역          0
##  3 경기         65
##  4 경남         59
##  5 경북        472
##  6 광주          7
##  7 대구       2055
##  8 대전         13
##  9 부산         75
## 10 서울         62
## 11 세종          1
## 12 울산         17
## 13 인천          5
## 14 전남          1
## 15 전북          4
## 16 제주          2
## 17 충남         48
## 18 충북         10
```

```
"대구이다."
```

```
## [1] "대구이다."
```