

R입문 기말고사

정덕중(201922037)

12/21/2021

1번문제풀이

(1)

```
epsilon=rnorm(1000)
head(epsilon)
```

```
## [1]  0.3282909 -0.1561083 -0.9119773 -1.6974335  0.7960057  2.6838088
```

(2)

```
t<-c()
for(i in 1:1000) t[i]=2*pi*i/1000
x_1<-c()
x_1 = sin(t)
x_2<-c()
x_2 = cos(4*t)
```

- x_1

```
head(x_1)
```

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

- x_2

```
head(x_2)
```

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

(4)

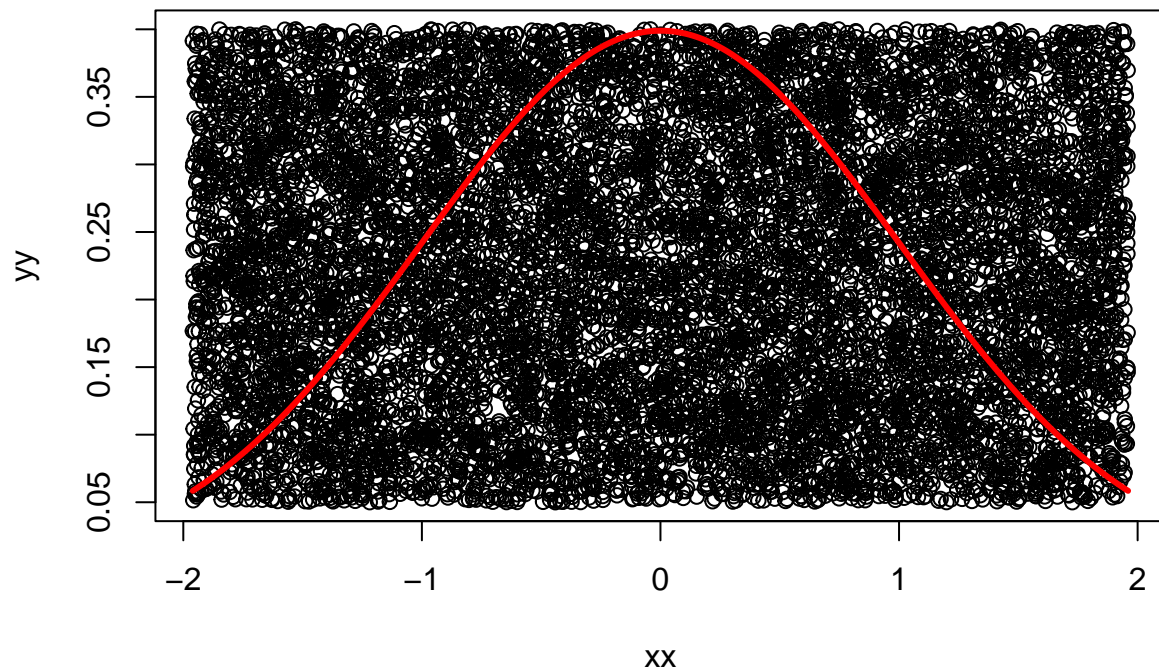
```
head(cbind(1,x_1,x_2))
```

```
##           x_1      x_2
## [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
```

2번문제풀이

(1)

```
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=-1.96,max=1.96)
yy= runif(10000,min=0.05,max=0.40)
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])
head(tst)
```

```
## [1] FALSE FALSE TRUE FALSE TRUE FALSE
```

```
sum(tst)
```

```
## [1] 5459
```

(2)

```
x <- rnorm(1000)
sum(x<1.96 & x>-1.96)
```

```
## [1] 950
```

3번문제풀이

```
library(tidyverse)
```

```
## Warning: 패키지 'tidyverse'는 R 버전 4.1.2에서 작성되었습니다
```

```
## -- 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
```

```
## Warning: 패키지 'ggplot2'는 R 버전 4.1.2에서 작성되었습니다
```

```
## Warning: 패키지 'tibble'는 R 버전 4.1.2에서 작성되었습니다
```

```
## Warning: 패키지 'tidyr'는 R 버전 4.1.2에서 작성되었습니다
```

```
## Warning: 패키지 'readr'는 R 버전 4.1.2에서 작성되었습니다
```

```
## Warning: 패키지 'purrr'는 R 버전 4.1.2에서 작성되었습니다
```

```
## Warning: 패키지 'dplyr'는 R 버전 4.1.2에서 작성되었습니다
```

```
## Warning: 패키지 'stringr'는 R 버전 4.1.2에서 작성되었습니다
```

```
## Warning: 패키지 'forcats'는 R 버전 4.1.2에서 작성되었습니다
```

```
## -- 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.
```

```
head(df)
```

```
## # A tibble: 6 x 5
```

```
##   year month   day prov  cases
```

```
##   <dbl> <dbl> <dbl> <chr> <dbl>
```

```
## 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
```

4번 문제풀이

(1)

- 2020년의 확진자(cases)총합

```
df %>% filter(year == '2020') %>% summarise(sum_value1=sum(cases))
```

```
## # A tibble: 1 x 1
##   sum_value1
##       <dbl>
## 1       60726
```

- 2021년의 확진자(cases)총합

```
df %>% filter(year == '2021') %>% summarise(sum_value2=sum(cases))
```

```
## # A tibble: 1 x 1
##   sum_value2
##       <dbl>
## 1      396886
```

#(2) - 2020년2월1일~2020년2월15일 기간 서울에서 발생한 확진자의 합

```
df %>% filter(year == '2020' & month == '2' & prov == '서울' )%>% filter(day == 1:15)%>%
```

```
## Warning in day == 1:15: 두 객체의 길이가 서로 배수관계에 있지 않습니다
```

```
## # A tibble: 1 x 1
##   sum_서울
##       <dbl>
## 1         5
```

- 2020년2월1일~2020년2월15일 기간 부산에서 발생한 환자수

```
df %>% filter(year == '2020' & month == '2' & prov == '부산' ) %>% filter(day == 1:15) %>%
```

```
## Warning in day == 1:15: 두 객체의 길이가 서로 배수관계에 있지 않습니다
```

```
## # A tibble: 1 x 1
##   sum_부산
##       <dbl>
## 1         0
```

- 2020년2월1일~2020년2월15일 기간 대구에서 발생한 환자수

```
df %>% filter(year == '2020' & month == '2' & prov == '대구' ) %>% filter(day == 1:15) %>% summarise(sum_대구 = sum(case_when(day == 1:15 ~ 1, TRUE ~ 0)))

## Warning in day == 1:15: 두 객체의 길이가 서로 배수관계에 있지 않습니다

## # A tibble: 1 x 1
##   sum_대구
##   <dbl>
## 1       0
```

- 2020년2월1일~2020년2월15일 기간 인천에서 발생한 환자수

```
df %>% filter(year == '2020' & month == '2' & prov == '인천' ) %>% filter(day == 1:15) %>% summarise(sum_인천 = sum(case_when(day == 1:15 ~ 1, TRUE ~ 0)))

## Warning in day == 1:15: 두 객체의 길이가 서로 배수관계에 있지 않습니다

## # A tibble: 1 x 1
##   sum_인천
##   <dbl>
## 1       0
```

- 2020년2월1일~2020년2월15일 기간 광주에서 발생한 환자수

```
df %>% filter(year == '2020' & month == '2' & prov == '광주' ) %>% filter(day == 1:15) %>% summarise(sum_광주 = sum(case_when(day == 1:15 ~ 1, TRUE ~ 0)))

## Warning in day == 1:15: 두 객체의 길이가 서로 배수관계에 있지 않습니다

## # A tibble: 1 x 1
##   sum_광주
##   <dbl>
## 1       2
```

- 2020년2월1일~2020년2월15일 기간 울산에서 발생한 환자수

```
df %>% filter(year == '2020' & month == '2' & prov == '울산' ) %>% filter(day == 1:15) %>% summarise(sum_울산 = sum(case_when(day == 1:15 ~ 1, TRUE ~ 0)))

## Warning in day == 1:15: 두 객체의 길이가 서로 배수관계에 있지 않습니다

## # A tibble: 1 x 1
##   sum_울산
##   <dbl>
## 1       0
```

- 2020년2월1일~2020년2월15일 기간 세종에서 발생한 환자수

```
df %>% filter(year == '2020' & month == '2' & prov == '세종' ) %>% filter(day == 1:15) %>% summarise(sum_세종 = sum(case_when(day == 1:15 ~ 1, TRUE ~ 0)))

## Warning in day == 1:15: 두 객체의 길이가 서로 배수관계에 있지 않습니다

## # A tibble: 1 x 1
##   sum_세종
##   <dbl>
```

```
## 1      0
• 2020년2월1일~2020년2월15일 기간 경기에서 발생한 환자수
df %>% filter(year == '2020' & month == '2' & prov == '경기' ) %>% filter(day == 1:15) %
## Warning in day == 1:15: 두 객체의 길이가 서로 배수관계에 있지 않습니다
## # A tibble: 1 x 1
##   sum_경기
##   <dbl>
## 1      9
• 2020년2월1일~2020년2월15일 기간 강원에서 발생한 환자수
df %>% filter(year == '2021' & month == '2' & prov == '강원' ) %>% filter(day == 1:15) %
## Warning in day == 1:15: 두 객체의 길이가 서로 배수관계에 있지 않습니다
## # A tibble: 1 x 1
##   sum_강원
##   <dbl>
## 1     78
• 2020년2월1일~2020년2월15일 기간 충북에서 발생한 환자수
df %>% filter(year == '2020' & month == '2' & prov == '충북' ) %>% filter(day == 1:15) %
## Warning in day == 1:15: 두 객체의 길이가 서로 배수관계에 있지 않습니다
## # A tibble: 1 x 1
##   sum_충북
##   <dbl>
## 1      0
• 2020년2월1일~2020년2월15일 기간 충남에서 발생한 환자수
df %>% filter(year == '2020' & month == '2' & prov == '충남' ) %>% filter(day == 1:15) %
## Warning in day == 1:15: 두 객체의 길이가 서로 배수관계에 있지 않습니다
## # A tibble: 1 x 1
##   sum_충남
##   <dbl>
## 1      0
• 2020년2월1일~2020년2월15일 기간 전북에서 발생한 환자수
df %>% filter(year == '2020' & month == '2' & prov == '전북' ) %>% filter(day == 1:15) %
## Warning in day == 1:15: 두 객체의 길이가 서로 배수관계에 있지 않습니다
## # A tibble: 1 x 1
```

```
## sum_전북
##      <dbl>
## 1      0
```

- 2020년2월1일~2020년2월15일 기간 전남에서 발생한 환자수

```
df %>% filter(year == '2020' & month == '2' & prov == '전남' ) %>% filter(day == 1:15) %>% summarise(sum_전남 = sum(case_when(
  prov == '전남' ~ 1,
  TRUE ~ 0)))

## Warning in day == 1:15: 두 객체의 길이가 서로 배수관계에 있지 않습니다

## # A tibble: 1 x 1
##   sum_전남
##     <dbl>
## 1       1
```

- 2020년2월1일~2020년2월15일 기간 경북에서 발생한 환자수

```
df %>% filter(year == '2020' & month == '2' & prov == '경북' ) %>% filter(day == 1:15) %>% summarise(sum_경북 = sum(case_when(
  prov == '경북' ~ 1,
  TRUE ~ 0)))

## Warning in day == 1:15: 두 객체의 길이가 서로 배수관계에 있지 않습니다

## # A tibble: 1 x 1
##   sum_경북
##     <dbl>
## 1       0
```

- 2020년2월1일~2020년2월15일 기간 경남에서 발생한 환자수

```
df %>% filter(year == '2020' & month == '2' & prov == '경남' ) %>% filter(day == 1:15) %>% summarise(sum_경남 = sum(case_when(
  prov == '경남' ~ 1,
  TRUE ~ 0)))

## Warning in day == 1:15: 두 객체의 길이가 서로 배수관계에 있지 않습니다

## # A tibble: 1 x 1
##   sum_경남
##     <dbl>
## 1       0
```

- 2020년2월1일~2020년2월15일 기간 제주에서 발생한 환자수

```
df %>% filter(year == '2020' & month == '2' & prov == '제주' ) %>% filter(day == 1:15) %>% summarise(sum_제주 = sum(case_when(
  prov == '제주' ~ 1,
  TRUE ~ 0)))

## Warning in day == 1:15: 두 객체의 길이가 서로 배수관계에 있지 않습니다

## # A tibble: 1 x 1
##   sum_제주
##     <dbl>
## 1       0
```

강원이 78명으로 가장 많다.

(3)

- 2020년2월16일~2020년2월29일 기간 서울에서 발생한 확진자의 합

```
df %>% filter(year == '2020' & month == '2' & prov == '서울' & cases )
```

```
## # A tibble: 14 x 5
##   year month   day prov  cases
##   <dbl> <dbl> <dbl> <chr> <dbl>
## 1  2020     2     2  서울     1
## 2  2020     2     5  서울     2
## 3  2020     2     6  서울     2
## 4  2020     2    16  서울     2
## 5  2020     2    19  서울     2
## 6  2020     2    21  서울     5
## 7  2020     2    22  서울     2
## 8  2020     2    23  서울     3
## 9  2020     2    24  서울     4
## 10 2020     2    25  서울     7
## 11 2020     2    26  서울     8
## 12 2020     2    27  서울    10
## 13 2020     2    28  서울     7
## 14 2020     2    29  서울    12
```

```
2+2+5+2+3+4+7+8+10+7+12
```

```
## [1] 62
```

- 2020년2월16일~2020년2월29일 기간 부산에서 발생한 환자수

```
df %>% filter(year == '2020' & month == '2' & prov == '부산' & cases ) %>% summarise(sum
```

```
## # A tibble: 1 x 1
##   sum_value
##   <dbl>
## 1       75
```

- 2020년2월16일~2020년2월29일 기간 대구에서 발생한 환자수

```
df %>% filter(year == '2020' & month == '2' & prov == '대구' & cases ) %>% summarise(sum
```

```
## # A tibble: 1 x 1
##   sum_value
##   <dbl>
## 1    2055
```

- 2020년2월16일~2020년2월29일 기간 인천에서 발생한 환자수

```
df %>% filter(year == '2020' & month == '2' & prov == '인천' & cases ) %>% summarise(sum
```

```
## # A tibble: 1 x 1
##   sum_value
##   <dbl>
## 1         5
```

- 2020년2월16일~2020년2월29일 기간 광주에서 발생한 환자수

```
df %>% filter(year == '2020' & month == '2' & prov == '광주' )
```

```
## # A tibble: 29 x 5
##   year month   day prov cases
##   <dbl> <dbl> <dbl> <chr> <dbl>
## 1  2020     2     1  광주     0
## 2  2020     2     2  광주     0
## 3  2020     2     3  광주     0
## 4  2020     2     4  광주     1
## 5  2020     2     5  광주     1
## 6  2020     2     6  광주     0
## 7  2020     2     7  광주     0
## 8  2020     2     8  광주     0
## 9  2020     2     9  광주     0
## 10 2020     2    10  광주     0
## # ... with 19 more rows
```

- 2020년2월16일~2020년2월29일 기간 울산에서 발생한 환자수

```
df %>% filter(year == '2020' & month == '2' & prov == '울산' & cases ) %>% summarise(sum
```

```
## # A tibble: 1 x 1
##   sum_value
##   <dbl>
## 1        17
```

- 2020년2월16일~2020년2월29일 기간 세종에서 발생한 환자수

```
df %>% filter(year == '2020' & month == '2' & prov == '세종' & cases ) %>% summarise(sum
```

```
## # A tibble: 1 x 1
##   sum_value
##   <dbl>
## 1         1
```

- 2020년2월16일~2020년2월29일 기간 경기에서 발생한 환자수

```
df %>% filter(year == '2020' & month == '2' & prov == '경기' )
```

```
## # A tibble: 29 x 5
##   year month   day prov cases
##   <dbl> <dbl> <dbl> <chr> <dbl>
```

```
## 1 2020 2 1 경기 1
## 2 2020 2 2 경기 2
## 3 2020 2 3 경기 0
## 4 2020 2 4 경기 0
## 5 2020 2 5 경기 2
## 6 2020 2 6 경기 0
## 7 2020 2 7 경기 0
## 8 2020 2 8 경기 0
## 9 2020 2 9 경기 3
## 10 2020 2 10 경기 1
## # ... with 19 more rows
```

- 2020년2월16일~2020년2월29일 기간 강원에서 발생한 환자수

```
df %>% filter(year == '2020' & month == '2' & prov == '강원' )
```

```
## # A tibble: 29 x 5
##   year month   day prov cases
##   <dbl> <dbl> <dbl> <chr> <dbl>
## 1 2020 2 1 강원 0
## 2 2020 2 2 강원 0
## 3 2020 2 3 강원 0
## 4 2020 2 4 강원 0
## 5 2020 2 5 강원 0
## 6 2020 2 6 강원 0
## 7 2020 2 7 강원 0
## 8 2020 2 8 강원 0
## 9 2020 2 9 강원 0
## 10 2020 2 10 강원 0
## # ... with 19 more rows
```

- 2020년2월16일~2020년2월29일 기간 충북에서 발생한 환자수

```
df %>% filter(year == '2020' & month == '2' & prov == '충북' & cases ) %>% summarise(sum_
```

```
## # A tibble: 1 x 1
##   sum_value
##   <dbl>
## 1 10
```

- 2020년2월16일~2020년2월29일 기간 충남에서 발생한 환자수

```
df %>% filter(year == '2020' & month == '2' & prov == '충남' & cases ) %>% summarise(sum_
```

```
## # A tibble: 1 x 1
##   sum_value
##   <dbl>
## 1 48
```

- 2020년2월16일~2020년2월29일 기간 전북에서 발생한 환자수

```
df %>% filter(year == '2020' & month == '2' & prov == '전북' & cases ) %>% summarise(sum_value)
```

```
## # A tibble: 1 x 1
```

	sum_value
1	4

- 2020년2월16일~2020년2월29일 기간 전남에서 발생한 환자수

```
df %>% filter(year == '2020' & month == '2' & prov == '전남' )
```

```
## # A tibble: 29 x 5
```

	year	month	day	prov	cases
1	2020	2	1	전남	0
2	2020	2	2	전남	0
3	2020	2	3	전남	0
4	2020	2	4	전남	0
5	2020	2	5	전남	0
6	2020	2	6	전남	1
7	2020	2	7	전남	0
8	2020	2	8	전남	0
9	2020	2	9	전남	0
10	2020	2	10	전남	0

```
## # ... with 19 more rows
```

- 2020년2월16일~2020년2월29일 기간 경북에서 발생한 환자수

```
df %>% filter(year == '2020' & month == '2' & prov == '경북' & cases ) %>% summarise(sum_value)
```

```
## # A tibble: 1 x 1
```

	sum_value
1	472

- 2020년2월16일~2020년2월29일 기간 경남에서 발생한 환자수

```
df %>% filter(year == '2020' & month == '2' & prov == '경남' & cases ) %>% summarise(sum_value)
```

```
## # A tibble: 1 x 1
```

	sum_value
1	59

- 2020년2월16일~2020년2월29일 기간 제주에서 발생한 환자수

```
df %>% filter(year == '2020' & month == '2' & prov == '제주' & cases ) %>% summarise(sum_value)
```

```
## # A tibble: 1 x 1
##   sum_value
##       <dbl>
## 1         2
```

가장많은 곳은 대구로 2055명이다

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

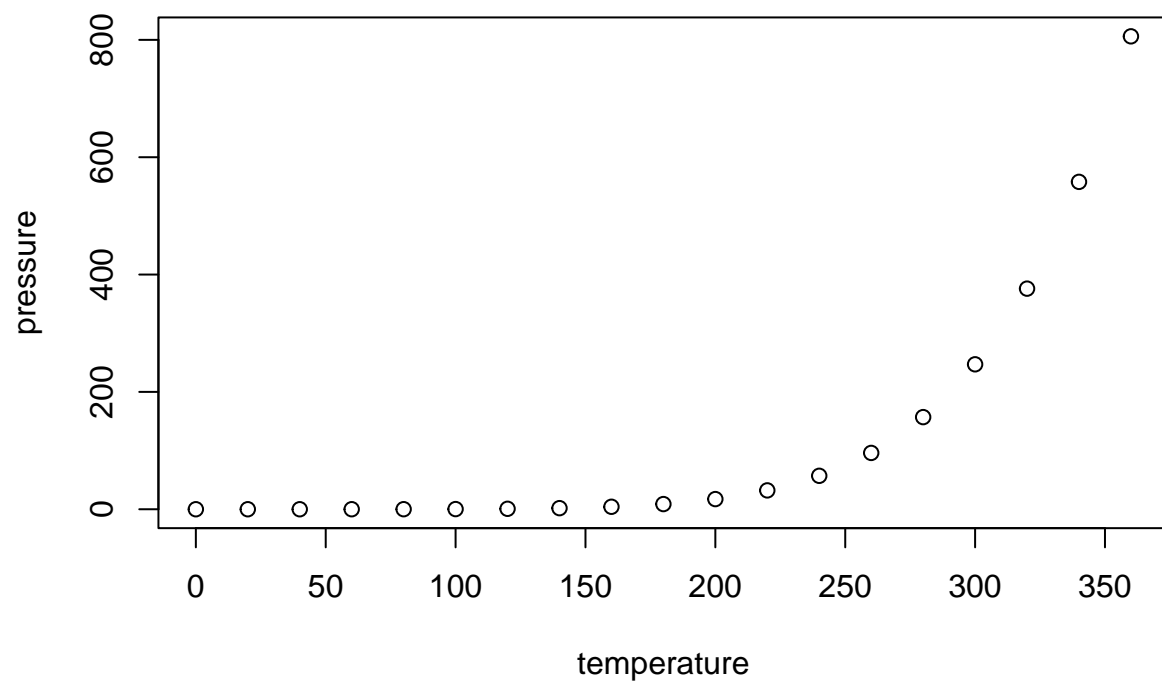
When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
summary(cars)
```

```
##      speed      dist
##  Min.   : 4.0   Min.   :  2.00
##  1st Qu.:12.0   1st Qu.: 26.00
##  Median :15.0   Median : 36.00
##  Mean   :15.4   Mean    : 42.98
##  3rd Qu.:19.0   3rd Qu.: 56.00
##  Max.   :25.0   Max.    :120.00
```

Including Plots

You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.