finalexam

```
##1


#1-1
epsilon_i<-rnorm(1000)


##1-2
k=1:1000
x1<-sin(2*pi*k/1000)
i=1:1000
x2<-cos(8*pi*i/1000)


##1-3
r=1:1000
y<-5*sin(2*pi*i/1000)+3*cos(8*pi*r/1000)+1.5+epsilon_i
d=1:1000
t<-2*pi*d/1000
plot(t,Y,col='gray60',lwd=5)


##1-4
k=1:1000
x1<-sin(2*pi*k/1000)
i=1:1000
x2<-cos(8*pi*i/1000)
X<-cbind(1,x1,x2)


##1-5
b<-rbind(1.5,5,3)
X<-cbind(1,x1,x2)
X%*%b
plot(t,Y,col='gray60',lwd=5)
lines(t,X%*%b,col='red',lwd=3)


#1-6
cbind(Y)
hatb=(solve(t(X)%*%X))%*%(t(X)%*%y)


#1-7
X%*%hatb
plot(t,Y,col='gray60',lwd=5)
```

```
lines(t,X%*%b,col='red',lwd=3)
lines(t,X%*%hatb,col='blue',lwd=3)


##2


#2-1
x=seq(from=-1.96, to=1.96,by=0.01)
y=1/(sqrt(2*pi))*exp((-1/2)*(x^2))
plot(x,y,type='l')
xx=runif(10000)*1.96
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])
plot(xx,yy,col='gray')
lines(x,y,col='red',lwd=3)
points(xx[tst],yy[tst],col='red')
sum(tst)*2
sum(tst)*2/10000


#2-2
library(tidyverse)
A=rnorm(1000)
A
dim(A)=c(1000)
A
A=as_tibble(A)
A
A %>% filter(A>-1.96&A<1.96)
count(A %>% filter(A>-1.96&A<1.96))


##3


##3-a
rslt<-c()
for(i in 1:10000){
    x=sum(cumprod(rbinom(20,size=1,0.5)))
    xx_=20-x-1
    if(xx_>0) y=rbinom(1,size=xx_,0.95)else y=0
        rslt[i]=+x+y
    }
mean(rslt)
18.1012
```

```r
#3-b
a=rbinom(20,size=1,0.5)
a
cumprod(a)
a=sum(a*cumprod(a))
a
b=rbinom(18,size=1,0.5)
b
cumprod(b)
b=sum(b*cumprod(b))
b
c=rbinom(17,size=1,0.5)
c
cumprod(c)
c=sum(c*cumprod(c))
c
d=rbinom(17,size=1,0.5)
d
cumprod(d)
d=sum(d*cumprod(d))
d
e=rbinom(16,size=1,0.5)
e
cumprod(e)
e=sum(e*cumprod(e))
e
f=rbinom(14,size=1,0.5)
f
cumprod(f)
f=sum(f*cumprod(f))
f
g=rbinom(14,size=1,0.5)
g
cumprod(g)
g=sum(g*cumprod(g))
g
20-6
14


##4
library(tidyverse)
df=read_csv('https://raw.githubusercontent.com/guebin/2021IR/master/_notebooks/covid19.csv')
lst=list(df)
lst
colnames(df)=c('year','month','day','prov','cases')


#4-1 ture=2020,false=2021
df%>%group_by(year==2020)%>%summarise(sum_cases=sum(cases))


#4-2
```

```r
df %>% filter(year==2020 & month==2) %>% filter(day %in% c(1:15))%>%group_by(prov)%>%summarise(sum_cases


#4-3
df %>% filter(year==2020 & month==2) %>% filter(day %in% c(16:29))%>%group_by(prov)%>%summarise(sum_cas
```