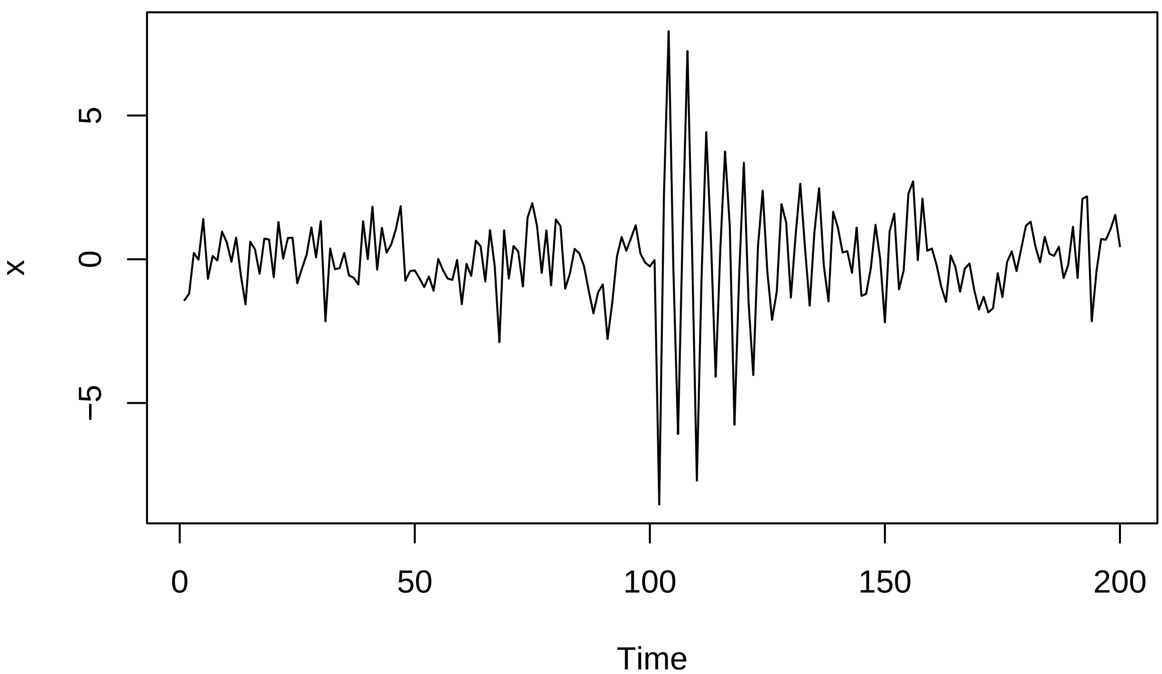
Trabalho 1 séries temporais

Daniel Krügel

2023-09-06

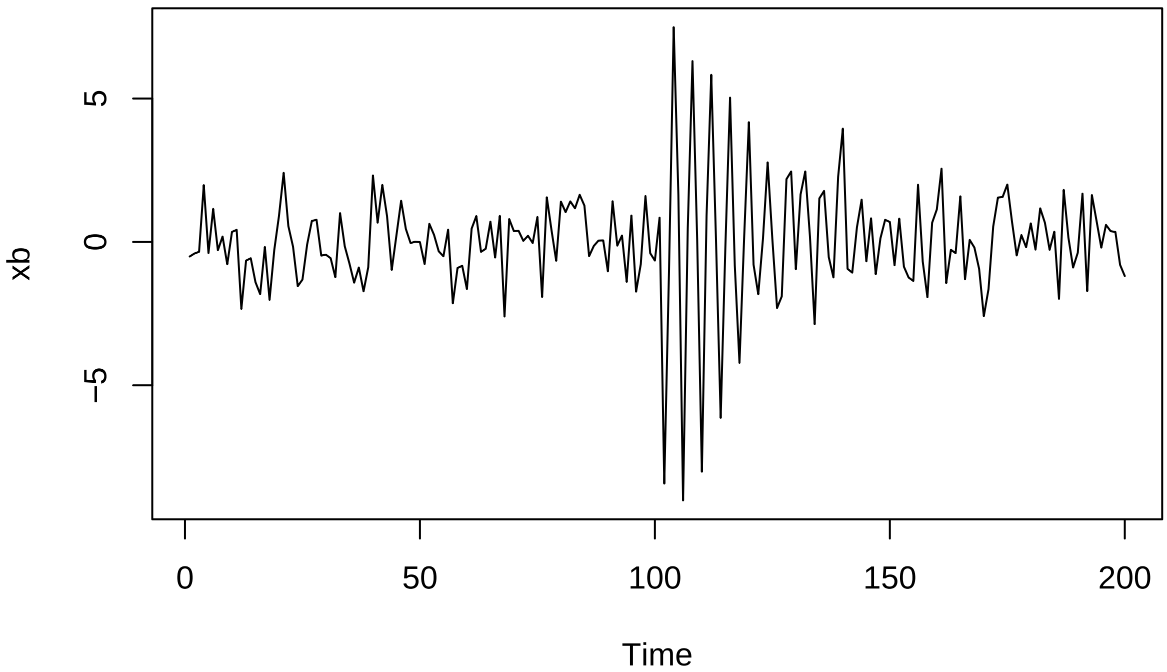
Questão 1

```
set.seed(156) # Escolhendo uma seed para reprodutibilidade
#a)
s = c(rep(0,100), 10*exp(-(1:100)/20)*cos(2*pi*(101:200)/4))
x = s + rnorm(200) # Adicionando componente aleatório
plot.ts(x)
```

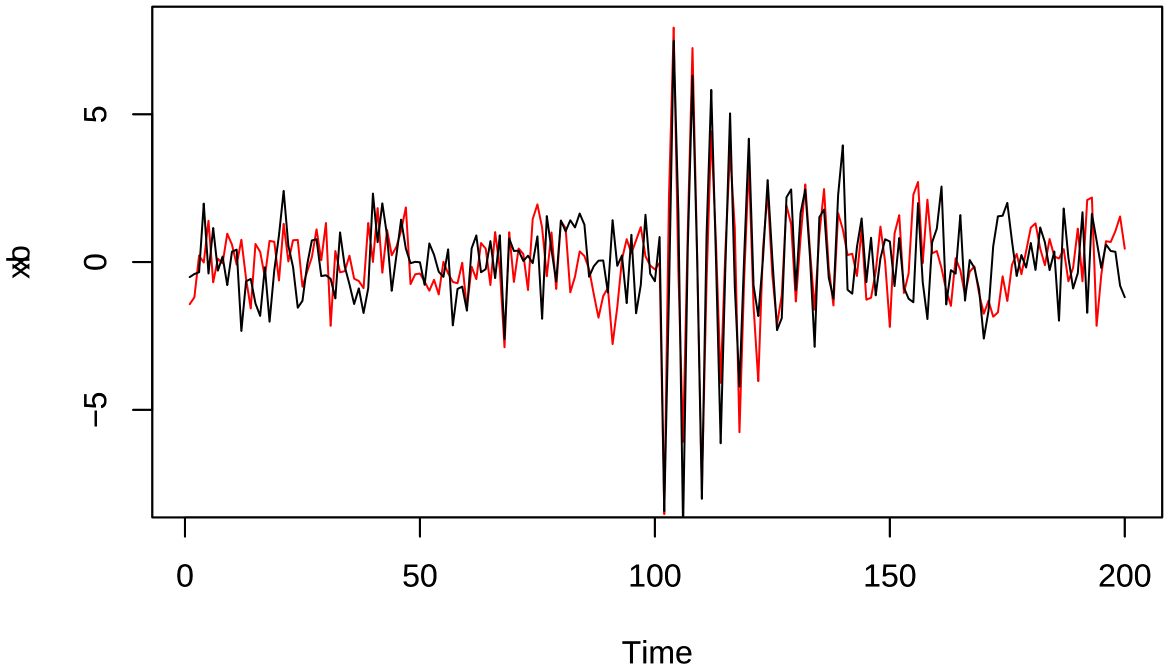


b)

```
#b)
sb = c(rep(0,100), 10*exp(-(1:100)/200)*cos(2*pi*(101:200)/4))
xb = s + rnorm(200)
plot.ts(xb)
```




```
a1 <- plot.ts(x, col = "red", ylim = c(-8,8))
par(new = T)
b1 <- plot.ts(xb, col = "black", ylim = c(-8,8))
```

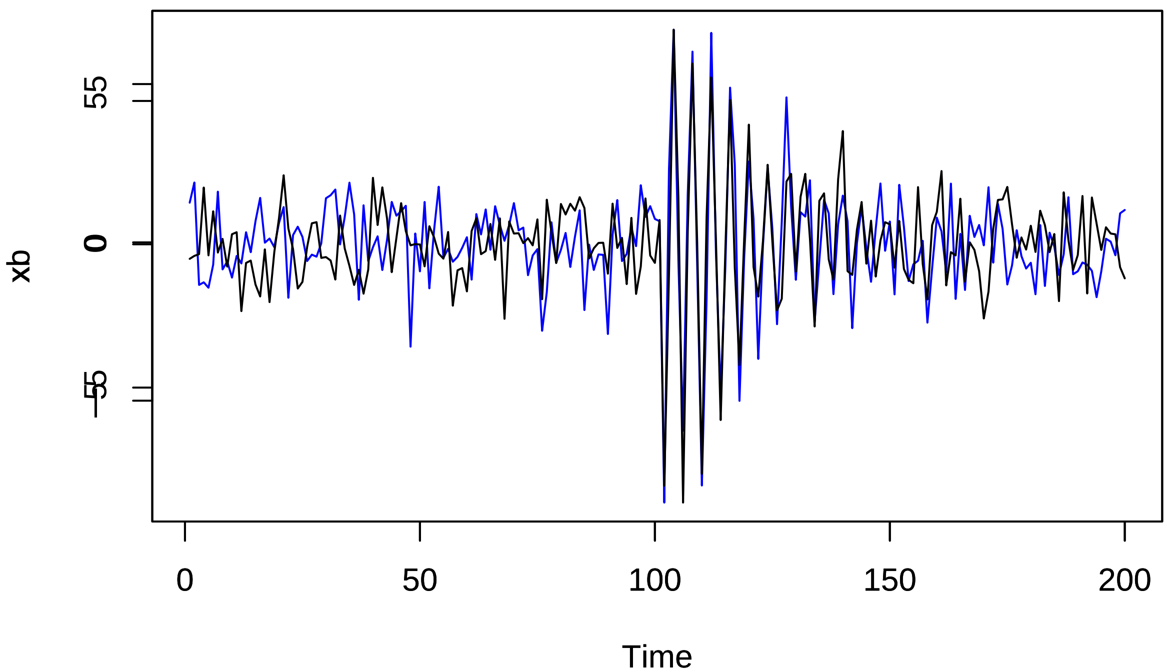



```
par(new = F)
```

Após o início da parcela exclusivamente aleatória, a alteração do modulador não alterou tanto a amplitude do sinal.

c)

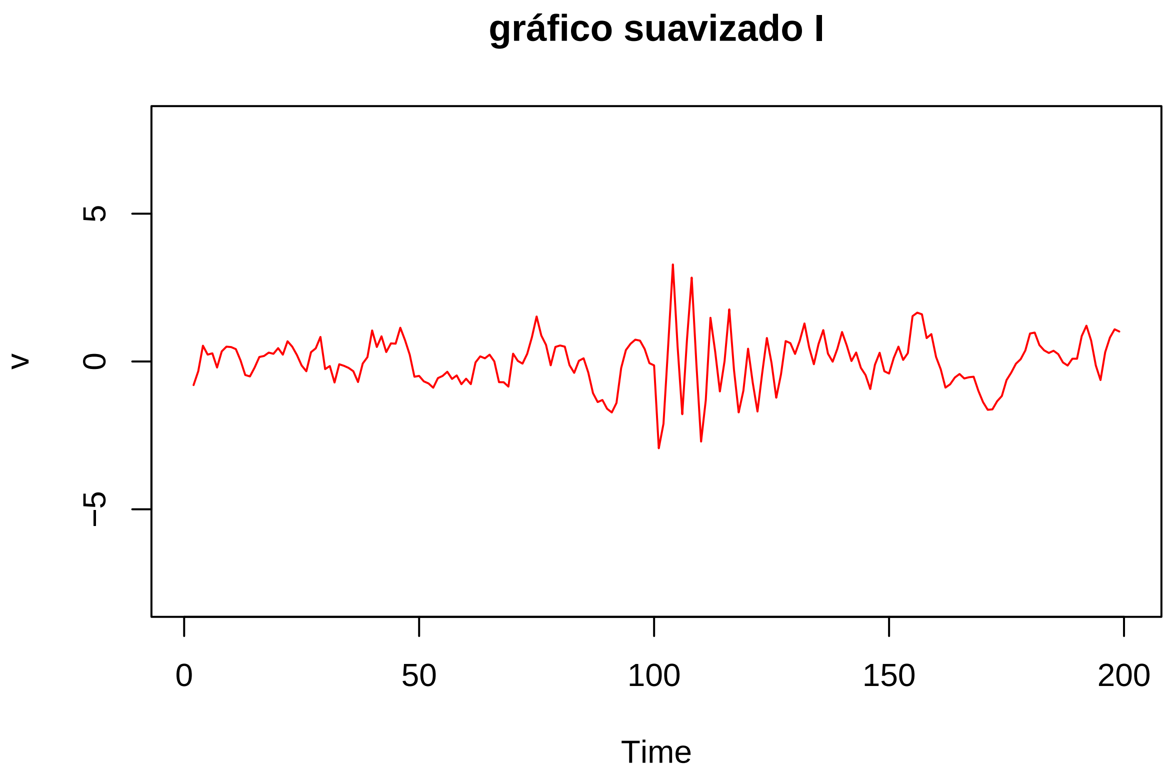
```
sc = c(rep(0,100), 10*exp((1:100)/200)*cos(2*pi*(101:200)/4))
xc = s + rnorm(200)
plot.ts(xc, col = "blue")
par(new = TRUE)
plot.ts(xb)
```



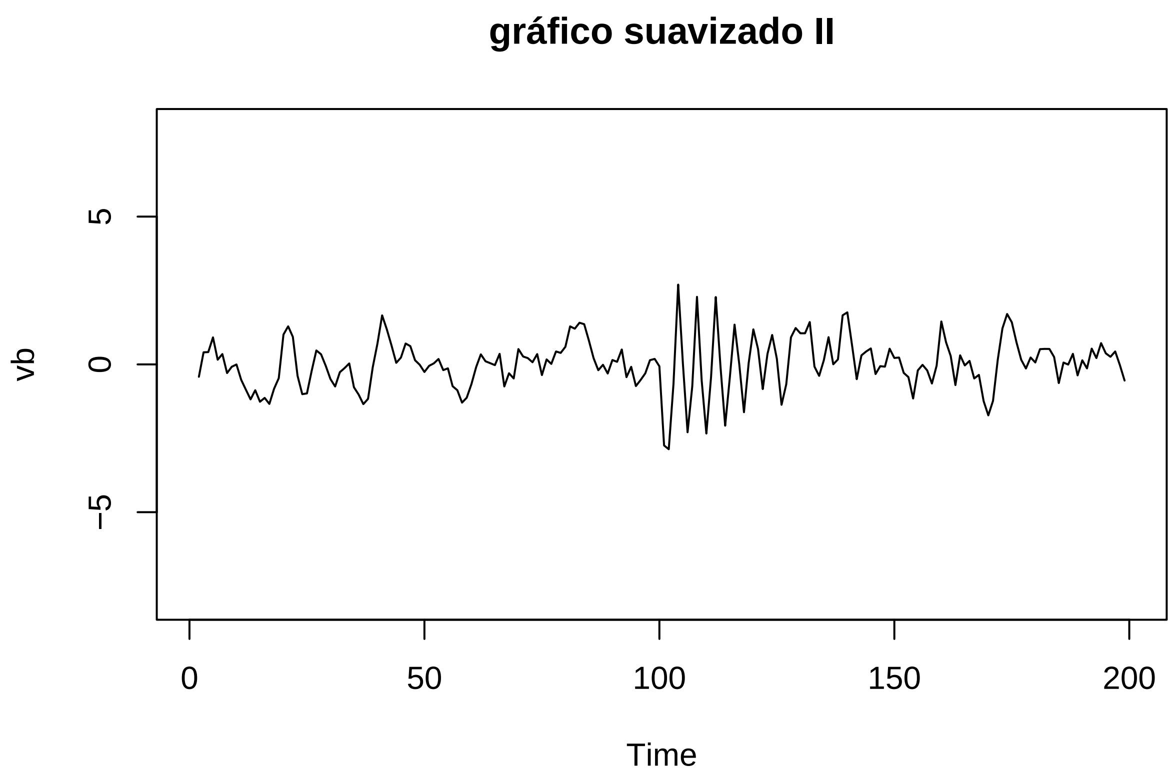
A alteração do sinal não alterou a informação da série, algo que cogitei que poderia acontecer seria a inversão da série, porém acredito que isso só aconteceria caso opuvesse mudança na função coseno para a seno

A comparação direta com a figura disponibilizada no material se parece com a série temporal relacionada a explosões, dado ao aumento repentino da variancia.

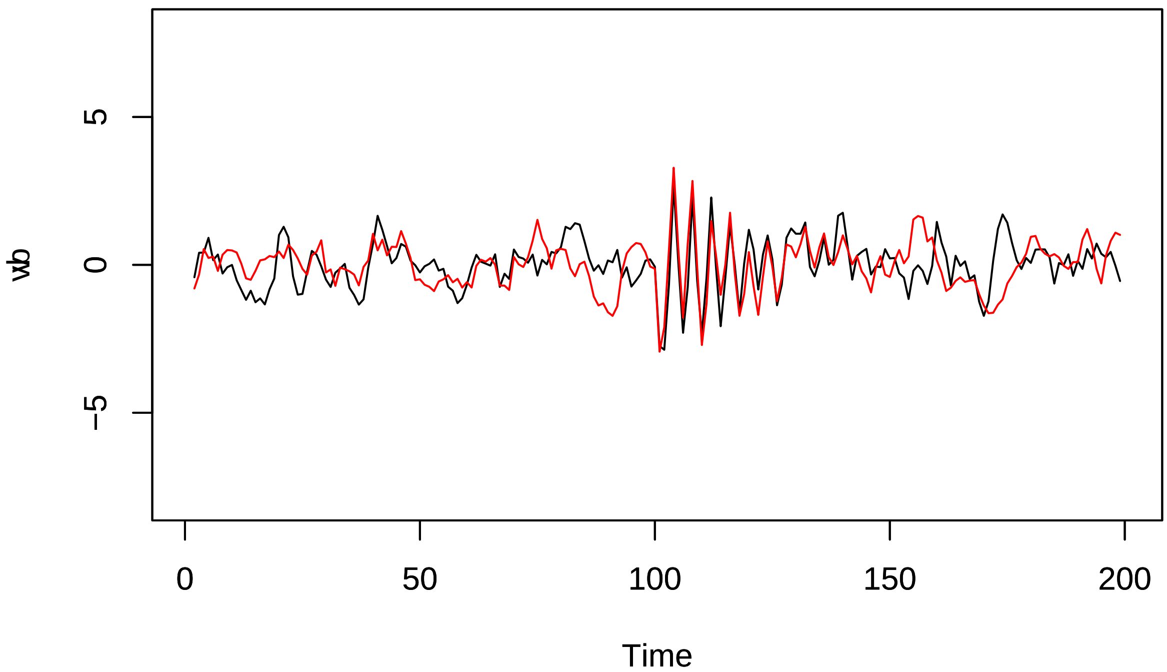
```
v = filter(x, sides=2, rep(1/3,3)) # médias móveis
plot.ts(v, col = "red", ylim = c(-8,8), main = "gráfico suavizado I")
```




```r
vb = filter(xb, sides=2, rep(1/3,3)) # médias móveis
plot.ts(vb, col = "black", ylim = c(-8,8), main = "gráfico suavizado II")
```



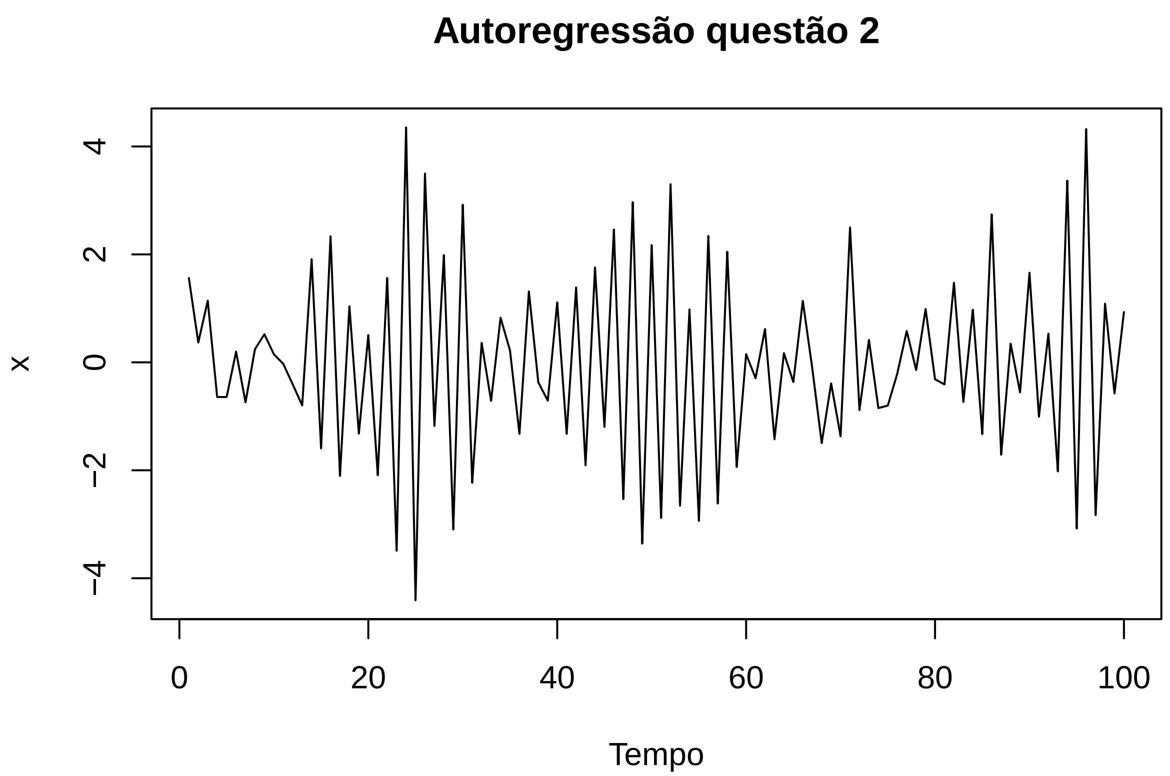
```r
v1 <- plot.ts(vb, col = "black", ylim = c(-8,8))
par(new = TRUE)
v2 <- plot.ts(v, col = "red", ylim = c(-8,8))
```



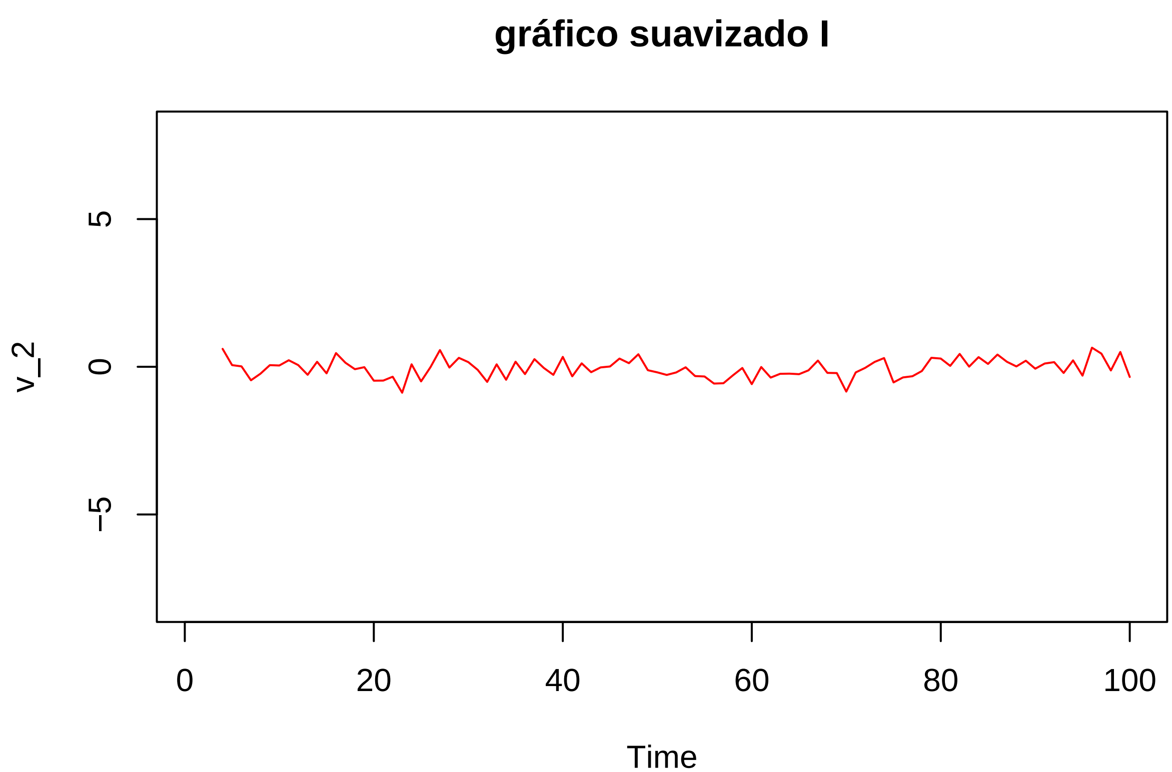

```r
par(new = F)
```

Questão 2

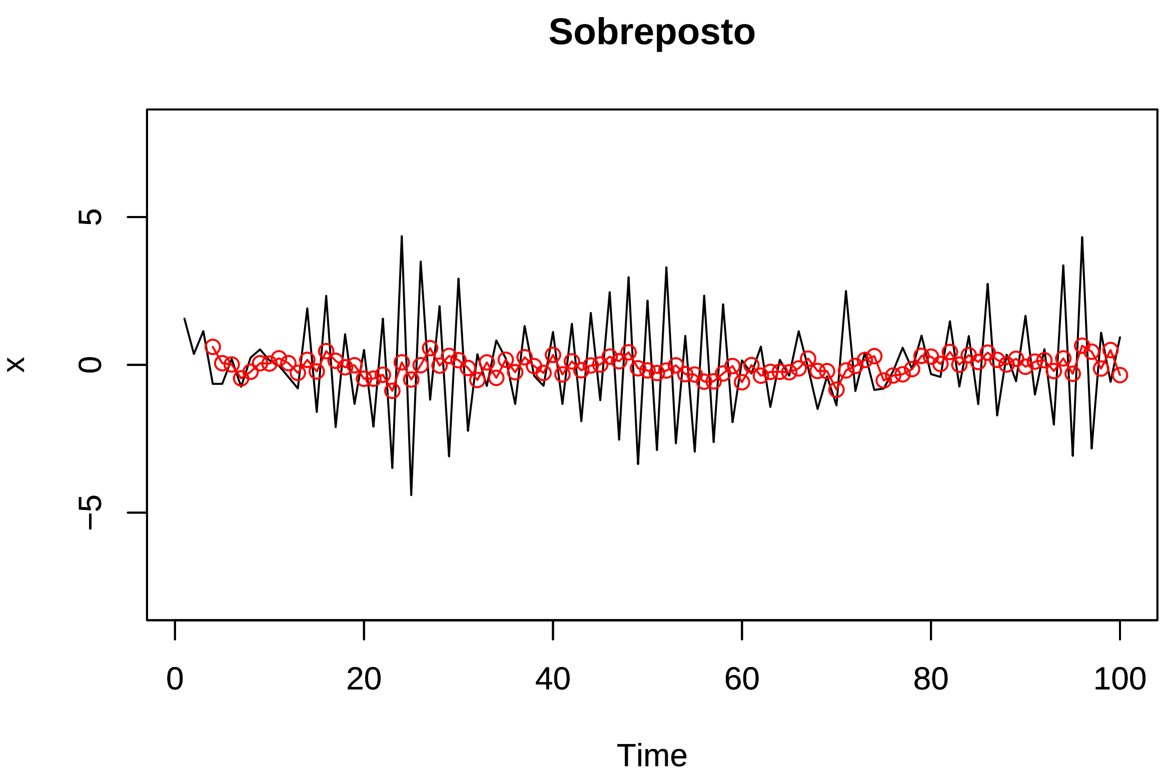
```r
w = rnorm(150,0,1) # 50 extras para evitar problemas de inicialização
x = filter(w, filter=c(-.9), method="recursive")[-(1:50)] # removendo os primeiros 50
x_2plot <- plot.ts(x, xlab="Tempo", main="Autoregressão questão 2")
```




```
v_2 <- filter(x, sides = 1, rep(1/4,4))
V_2plot<- plot.ts(v_2, col = "red", ylim = c(-8,8), main = "gráfico suavizado I")
```



```r
x_2plot <- plot.ts(x, ylim = c(-8,8))
par(new = T)
V_2plot<- plot.ts(v_2,ylab = "", col = "red", ylim = c(-8,8), main = "Sobreposto", type = "o")
```

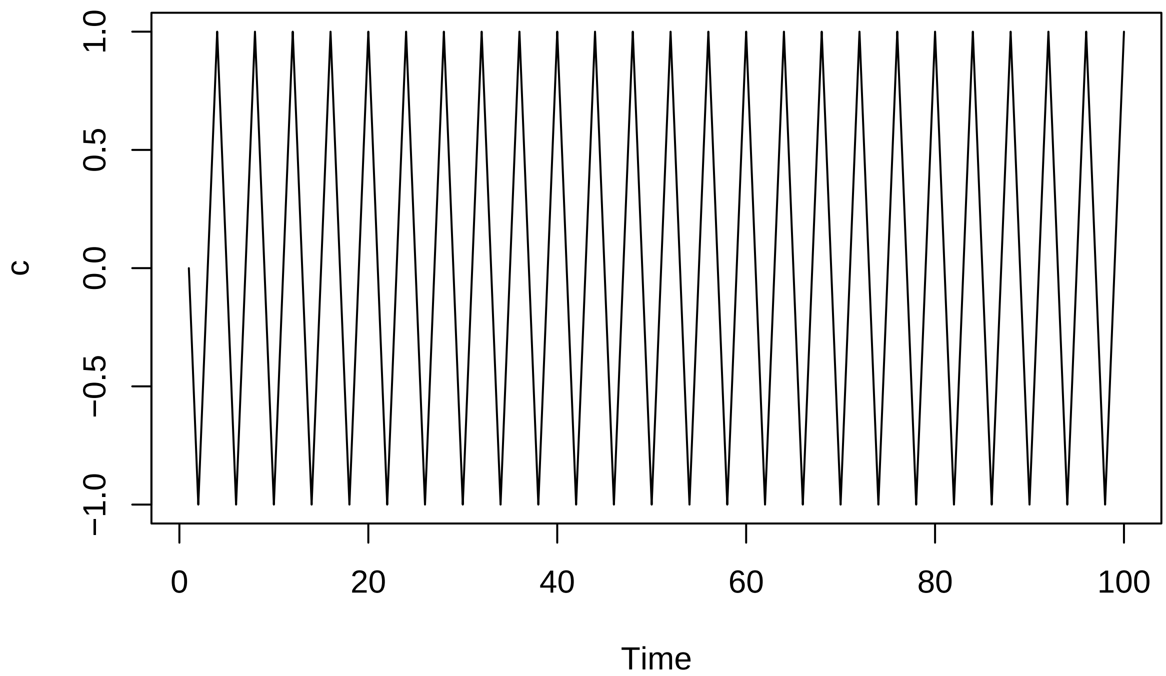



```r
par(new = F)
```

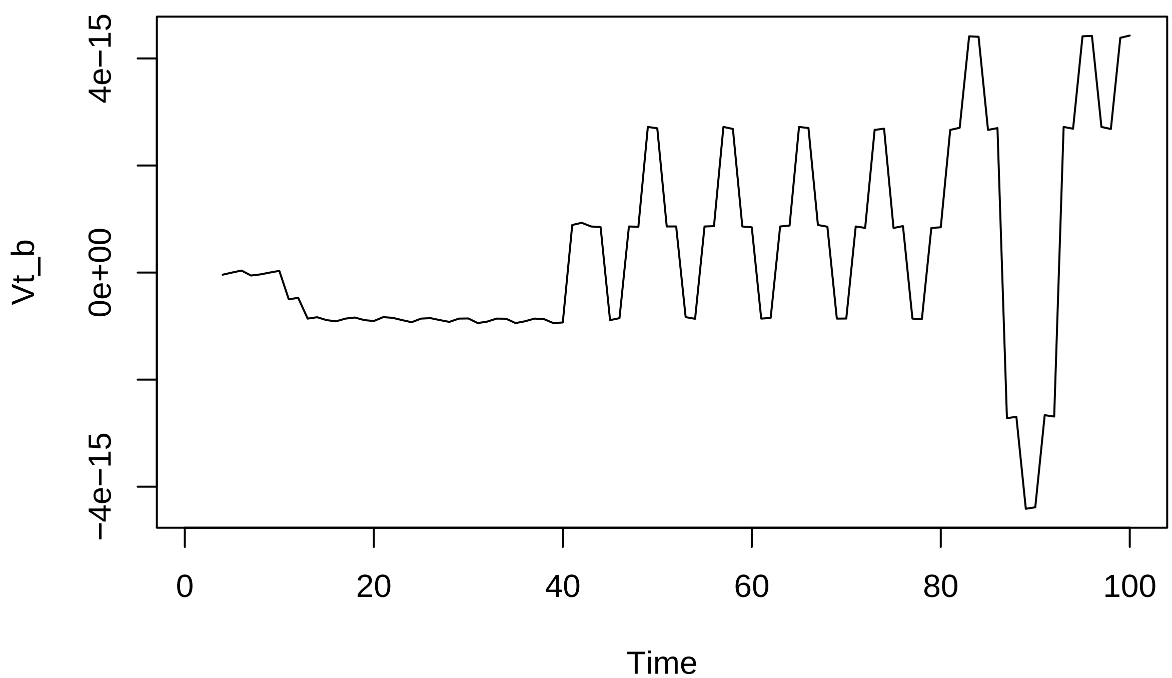
Suavisou a função de mais, neste ponto não é possível se obter informações sobre o comportamento da série.

b)

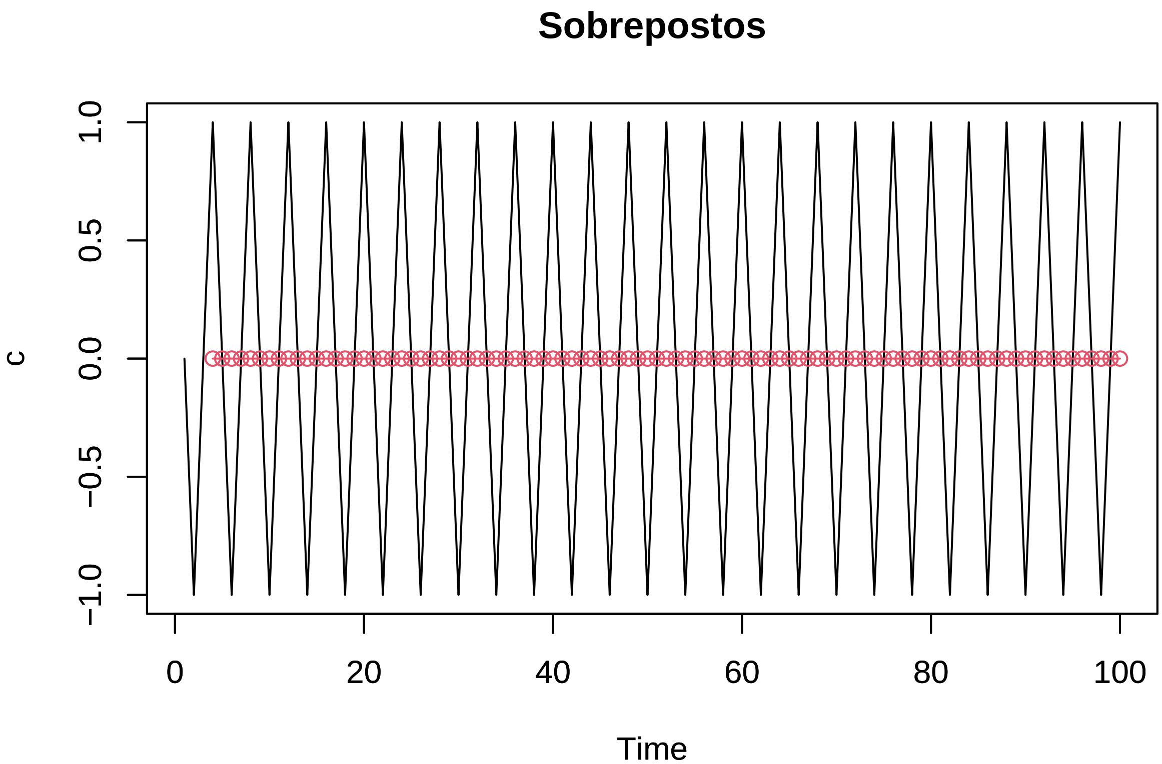
```r
c <- c(cos((2*pi*1:100)/4))
plot.ts(c)
```




```
Vt_b <- filter(c, sides = 1, rep(1/4,4)) # Criação de média móvel
plot.ts(Vt_b)
```




```
plot.ts(c, ylim = c(-1,1))
par(new = T)
plot.ts(Vt_b, ylim = c(-1,1),
        type = "o",
        ylab = "",
        main = "Sobrepostos",
        col = 2)
```

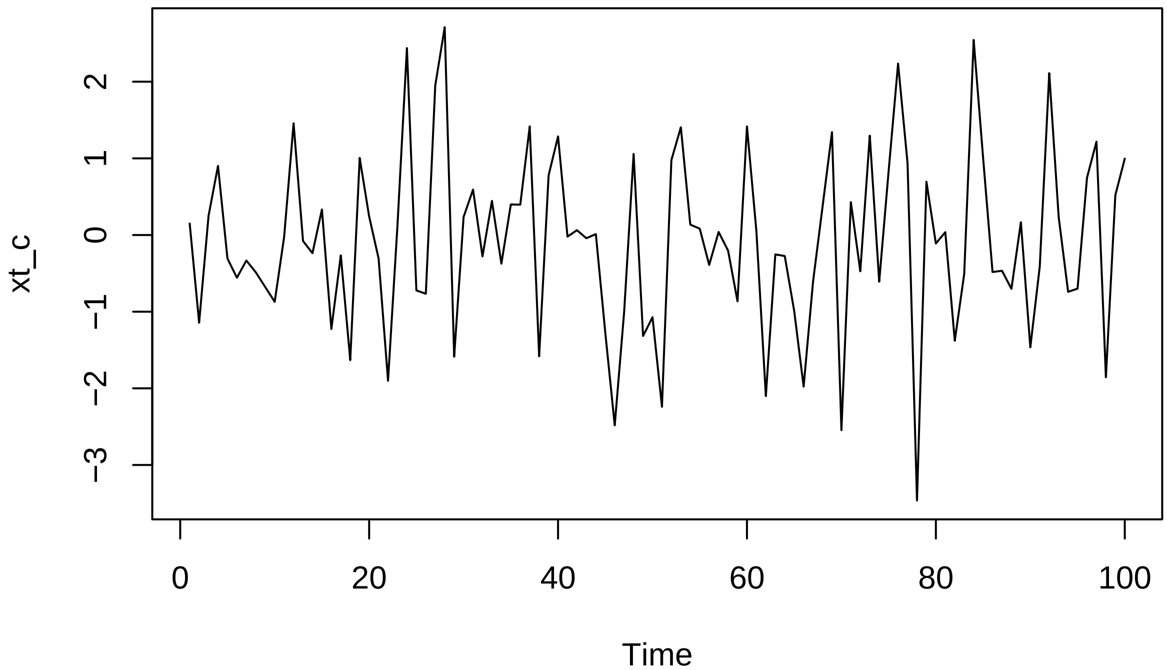


A suavisação foi tão forte que se perdeu completamente o padrão apresentado pela função

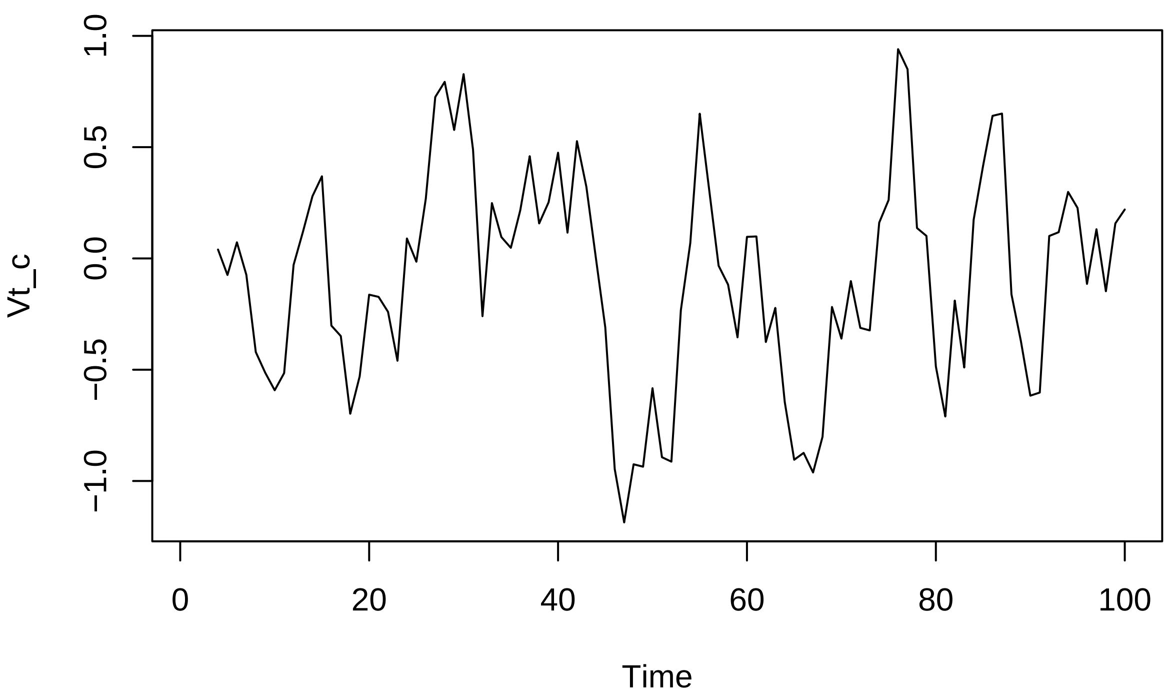
$$X_t$$

c)

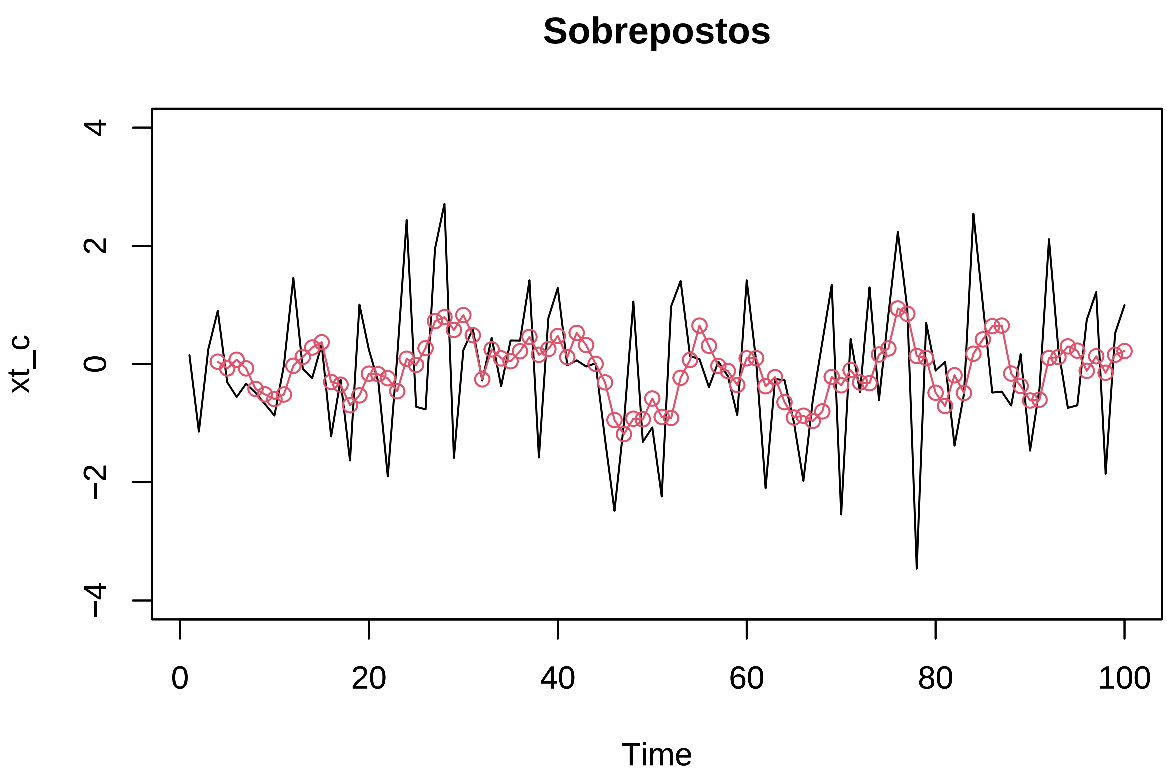
```
c <- c(cos((2*pi*1:100)/4))
x <- rnorm(100,0,1)
xt_c <- c + x
plot.ts(xt_c)
```




```
Vt_c <- filter(xt_c, sides = 1, rep(1/4,4)) # Criação de média móvel
plot.ts(Vt_c)
```




```
plot.ts(xt_c, ylim = c(-4,4))
par(new = T)
plot.ts(Vt_c, ylim = c(-4,4),
        type = "o",
        ylab = "",
        main = "Sobrepostos",
        col = 2)
```

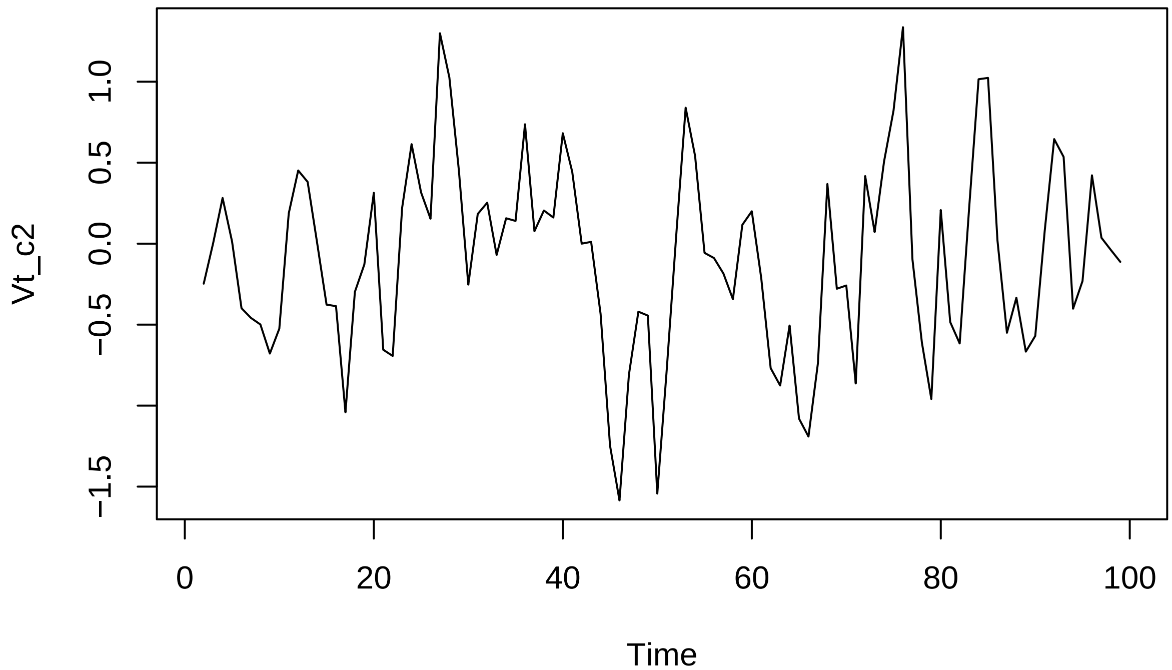


A série num geral ficou bem errática, não encontrei nenhuma comparação no material que se compare, porém a suavisação aqui não pareceu diminuir suficientemente o padrão da série.

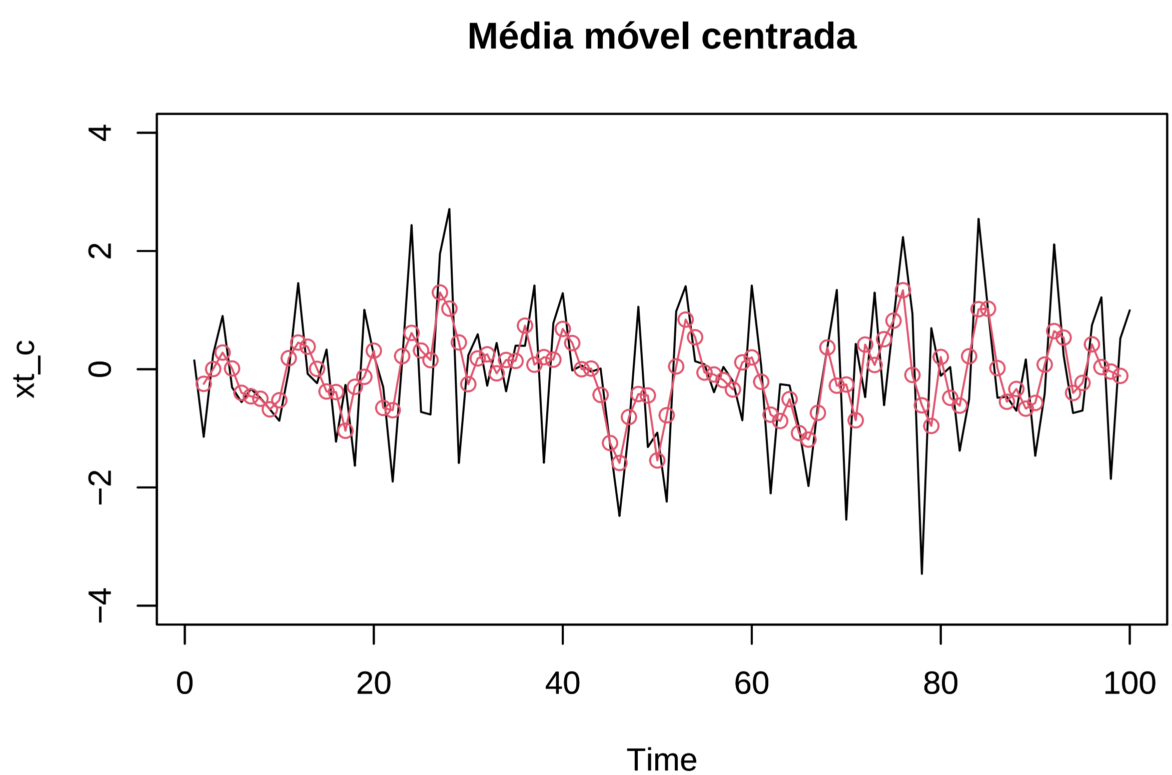
d)

Tanto no caso a) quanto no caso c) tivemos situações em que houve um aumento repentino na variância da série, podendo indicar um evento porém a média móvel não conseguiu captar essa mudança.

```
Vt_c2 <- filter(xt_c, sides = 2, rep(1/3,3)) # Criação de média móvel
plot.ts(Vt_c2)
```




```
plot.ts(xt_c, ylim = c(-4,4))
par(new = T)
plot.ts(Vt_c2, ylim = c(-4,4),
        type = "o",
        ylab = "",
        main = "Média móvel centrada",
        col = 2)
```



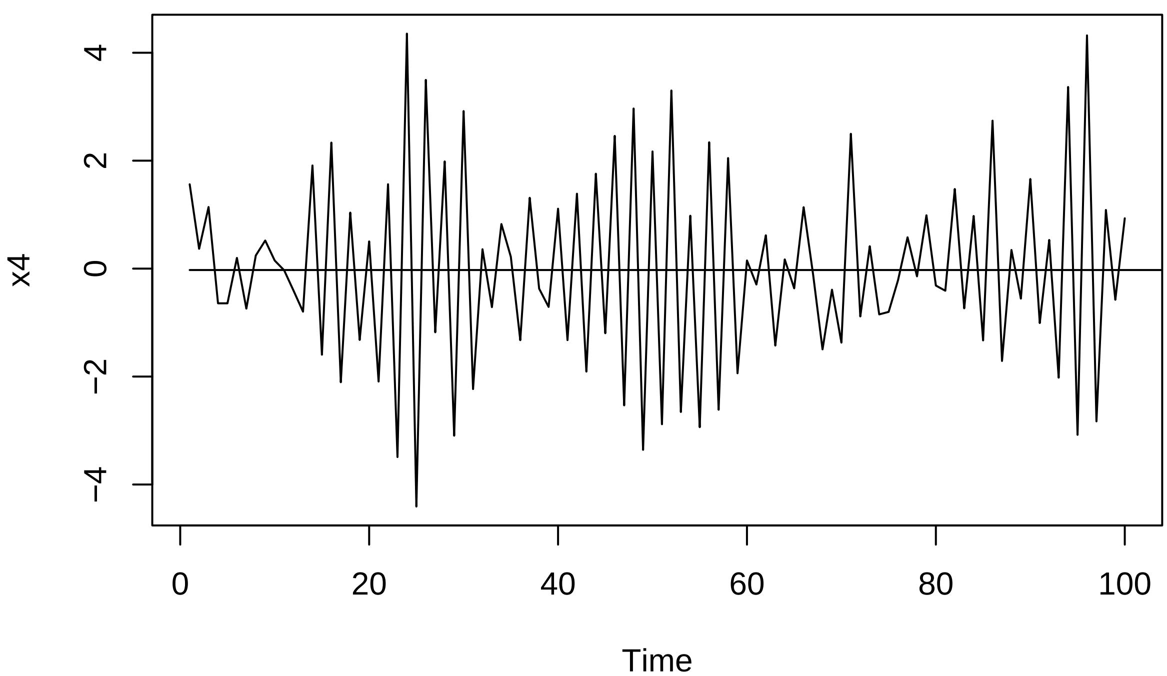
Enquanto a criação da média móvel centrada em zero como no exemplo I.13 vêmos uma série melhor descrita utilizando a média móvel centrada em 0.

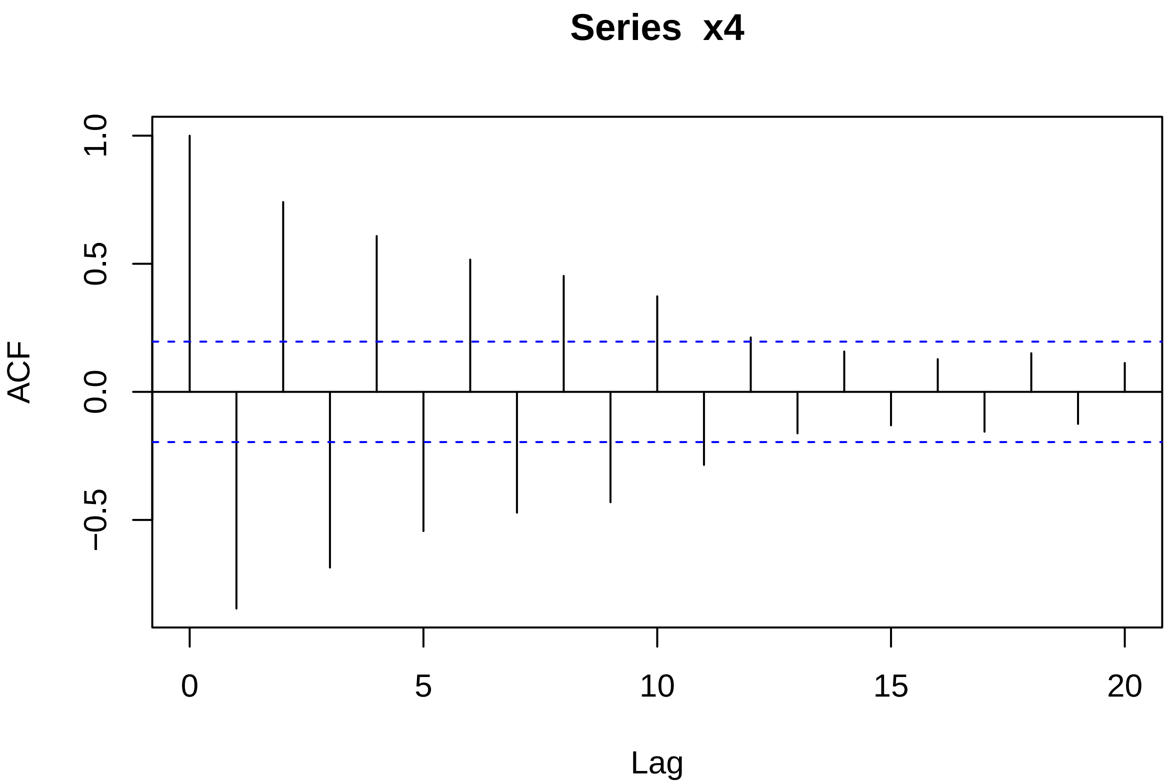
#Questão 3

$$
\begin{aligned}
E[X_t] &= E[X_s X_t - X_s\mu_t - X_t\mu_s + \mu_s\mu_t] \\
&= E[X_s X_t] - E[X_s\mu_t] - E[X_t\mu_s] + E[\mu_s\mu_t] \\
&= E[X_s X_t] - \mu_t\mu_s - \mu_s\mu_t + \mu_s\mu_t \\
&= E[X_s X_t] - \mu_t\mu_s
\end{aligned}
$$

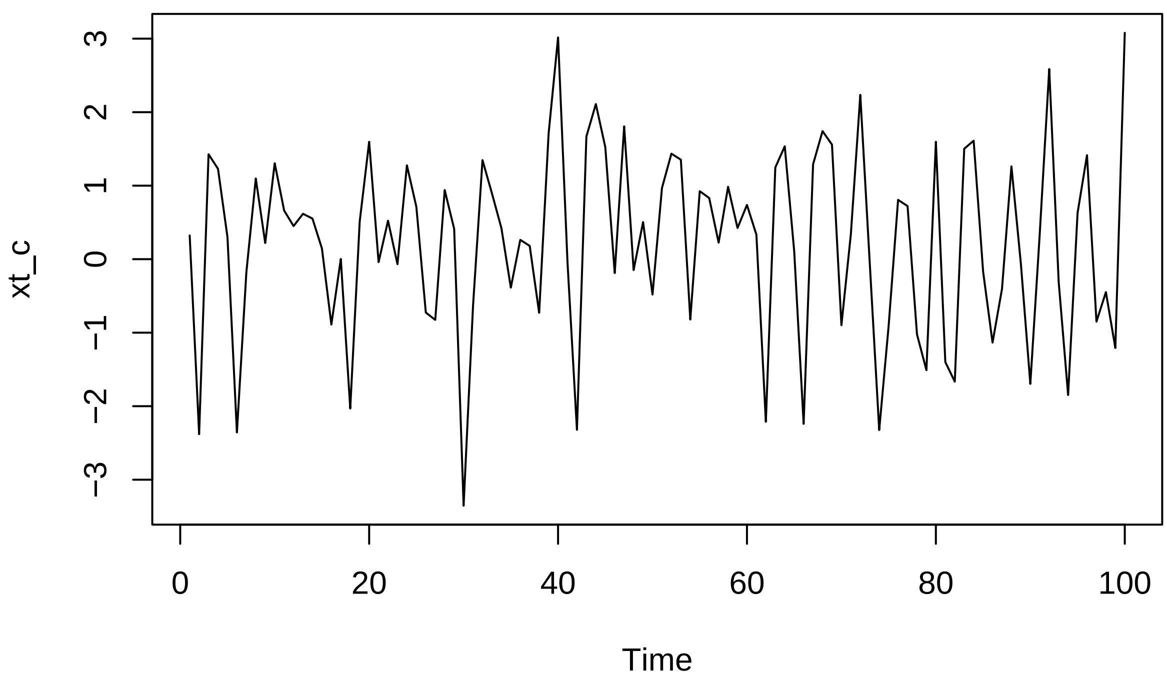
#Questão 4

```
w4 = rnorm(250,0,1) # 50 extras para evitar problemas de inicialização
x4 = filter(w, filter=c(-.9), method="recursive")[-(1:50)] # removendo os primeiros 50


xbar_4 <- sum(x4)/ length(x4) #Função de médias
x <- 1:200

df <- data.frame(x = x, y = xbar_4)
plot.ts(x4)
with(df, lines(y ~ x), col = 'red')
```

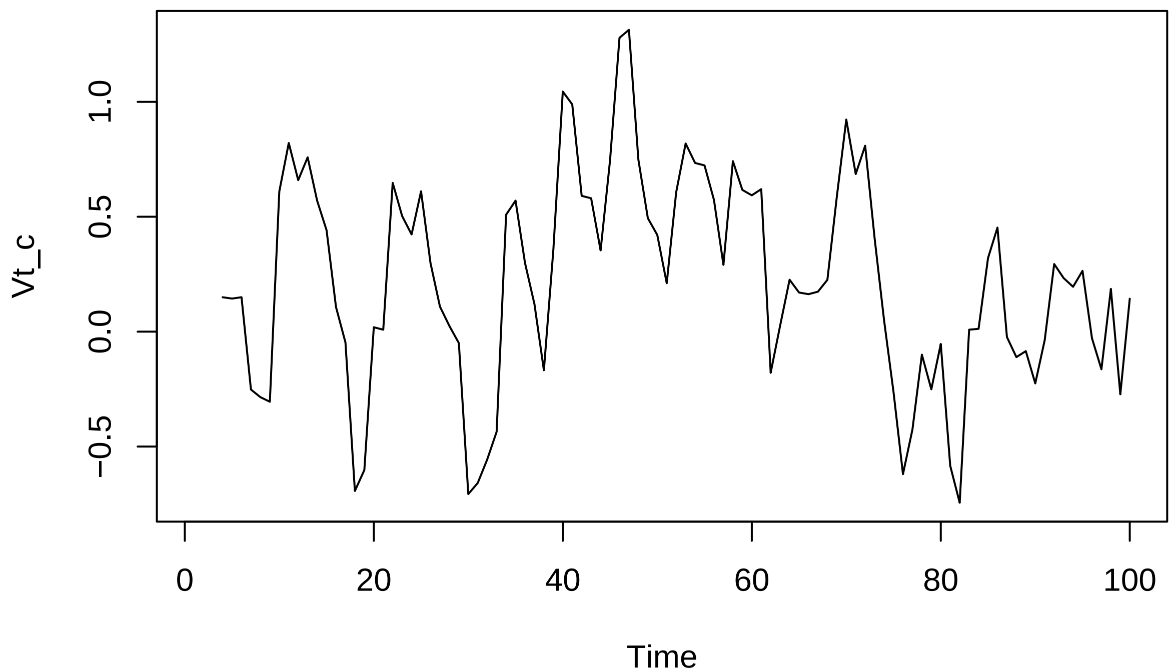




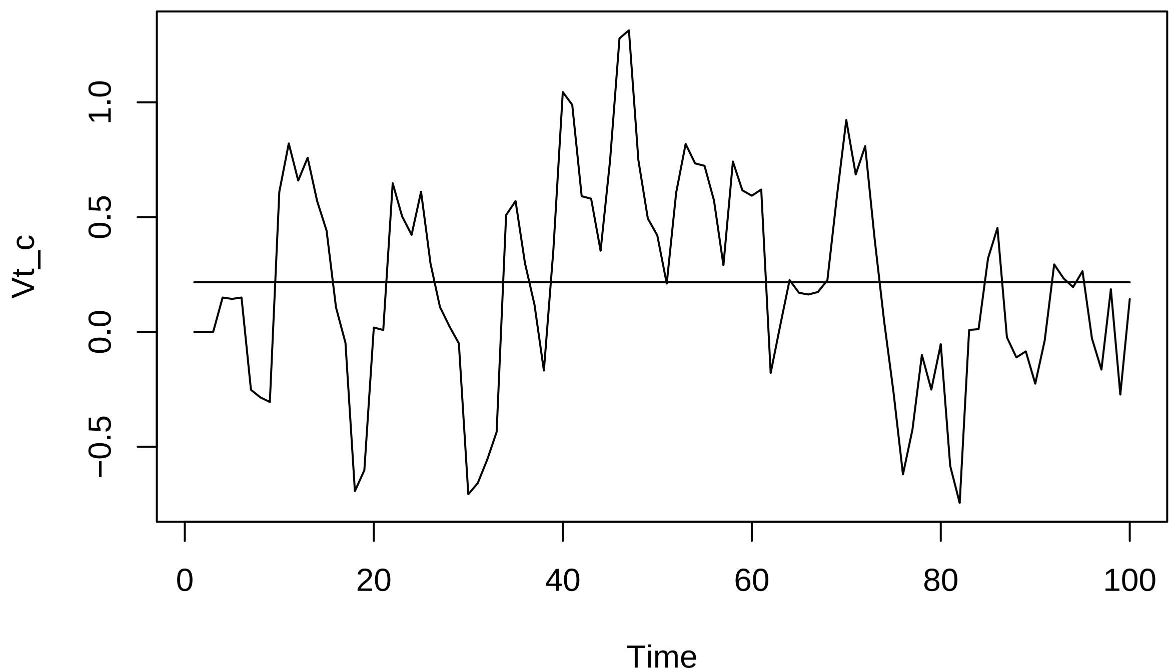

```
c <- c(cos((2*pi*1:100)/4))
x <- rnorm(100,0,1)
xt_c <- c + x
plot.ts(xt_c)
```

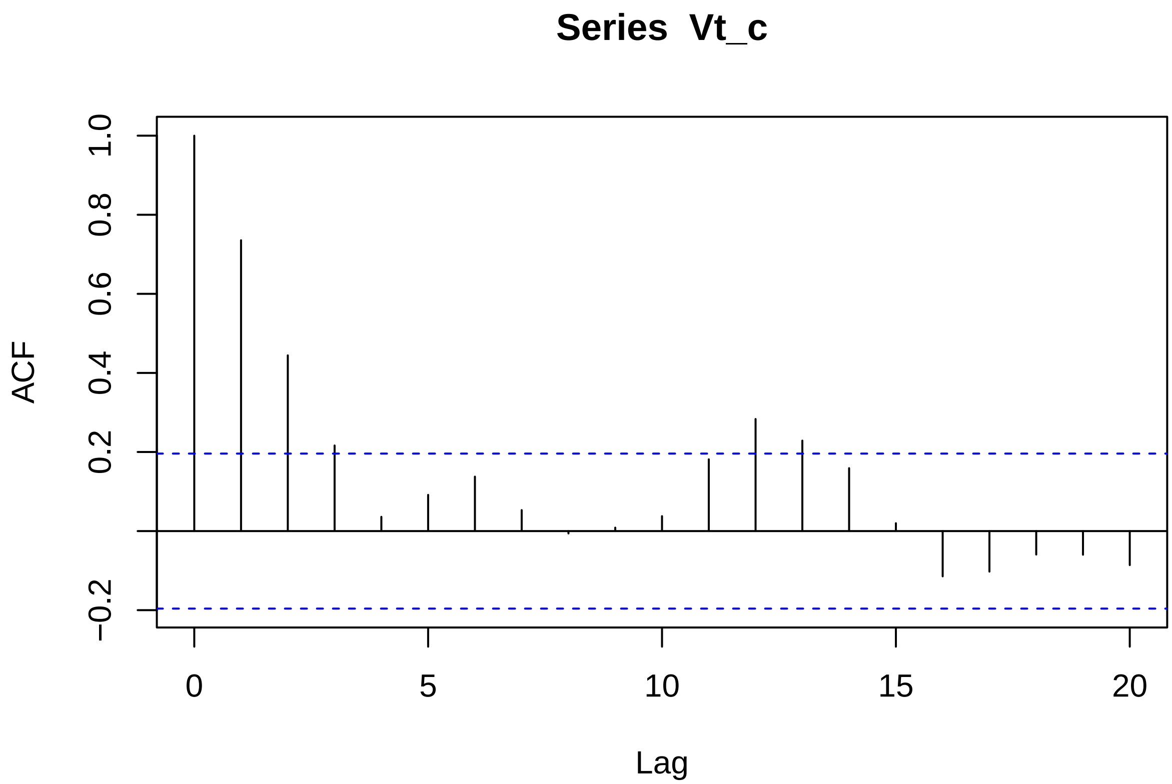


```
Vt_c <- filter(xt_c, sides = 1, rep(1/4,4)) # Criação de média móvel
plot.ts(Vt_c)
```




```
x <- 1:100
Vt_c<- tidyr::replace_na(Vt_c, 0)
vt_c_bar <- sum(Vt_c)/length(Vt_c)
df <- data.frame(x = x, y = vt_c_bar)
plot.ts(Vt_c)
with(df, lines(y ~ x), col = 'red')
```





Questão 5

$$X_t = \beta_0 + \beta_1 t + W_t$$
$$E[X_t] = E[\beta_0] + E[\beta_1 t] + E[W_t]$$

$$E[X_t] = \beta_0 + \beta_1 t + 0$$

Como a esperança depende do instante \$ t \$ a série não pode ser estacionária

#b

$$\beta_0 + \beta_1 t + W_t - (\beta_0 + \beta_1(t-1) + W_(t-1))$$
$$\beta_0 + \beta_1 t + W_t - (\beta_0 + \beta_1 t - \beta_1 + W_{t-1})$$

$$Y_t = \beta_1 + W_t + W_{t-1}$$

$$E[Y_t] = \beta_1 + 0 + 0$$

Como a esperança não depende de \$ t \$ a série poderá ser considerada estacionária se a covariância depender de $|s-t|$}

$$Cov[s,t] = E[(\beta_0 + W_s - W_{s-1} - \beta_0) * (\beta_0 + W_t - W_{t-1})]$$
$$E[(W_s - W_{s-1}) * (W_t - W_{t-1})]$$

Questão 6 lista 2

Questão a e b

Os calculos de comprovação são iguais ao da questão anterior. a) Não é estacionário b) É estacionário

c)

A substituição de uma constante por outra não irá influenciar no resultado para a primeira necessidade de estacionáriedade, estaremos substituindo $E[W_t] = 0$ por $E[Y_t] = \mu_y$ que continua não dependendo de t e a função de covariância será de:

$$\gamma_y(s) * \gamma_y(t)$$

podendo ser escrito em função da posição de s e t, portanto a série continua sendo estacionário

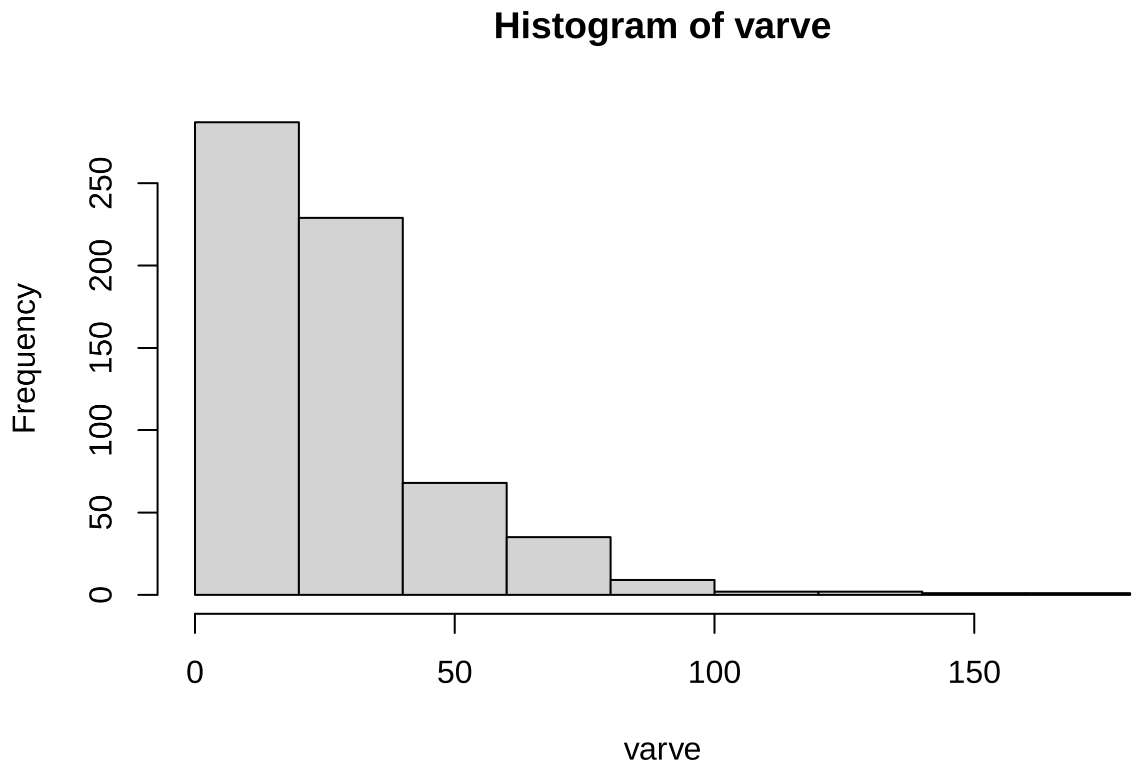
Questão 7

a)

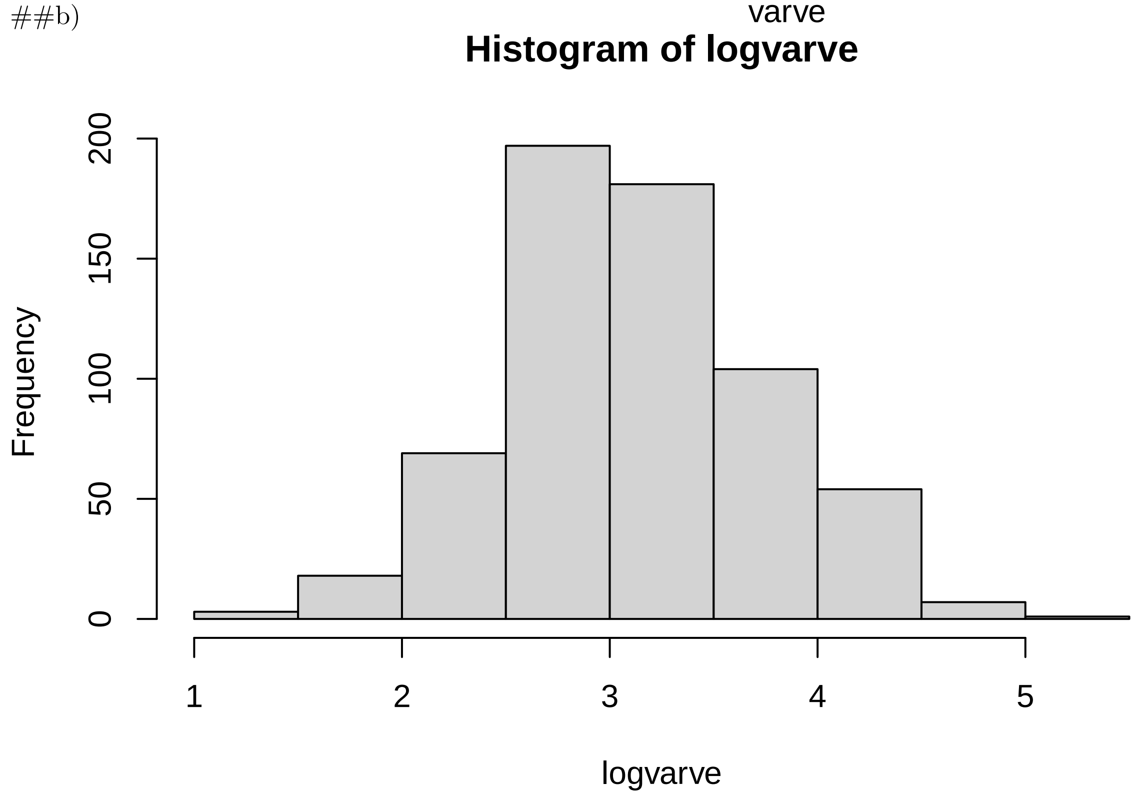
```r
library(astsa)
n<-length(varve)
var0 <- var(varve)

#Variancia primeira metade
var1 <- var(varve[1:(n/2)])

#Variancia segunda metade
var2 <- var(varve[((n+2)/2):n])

logvarve <- log(varve)

logvar1 <- var(logvarve[1:(n/2)])
logvar2 <- var(logvarve[((n+2)/2):n])
data.frame("Total" = var0,
           "Primeira metade" = var1,
           "Segunda metade" = var2,
           "Primeira metade log" = logvar1,
           "Segunda metade log" = logvar2)
```

```
##      Total Primeira.metade Segunda.metade Primeira.metade.log
## 1 412.6488        133.4574       594.4904           0.2707217
##   Segunda.metade.log
## 1           0.451371
```

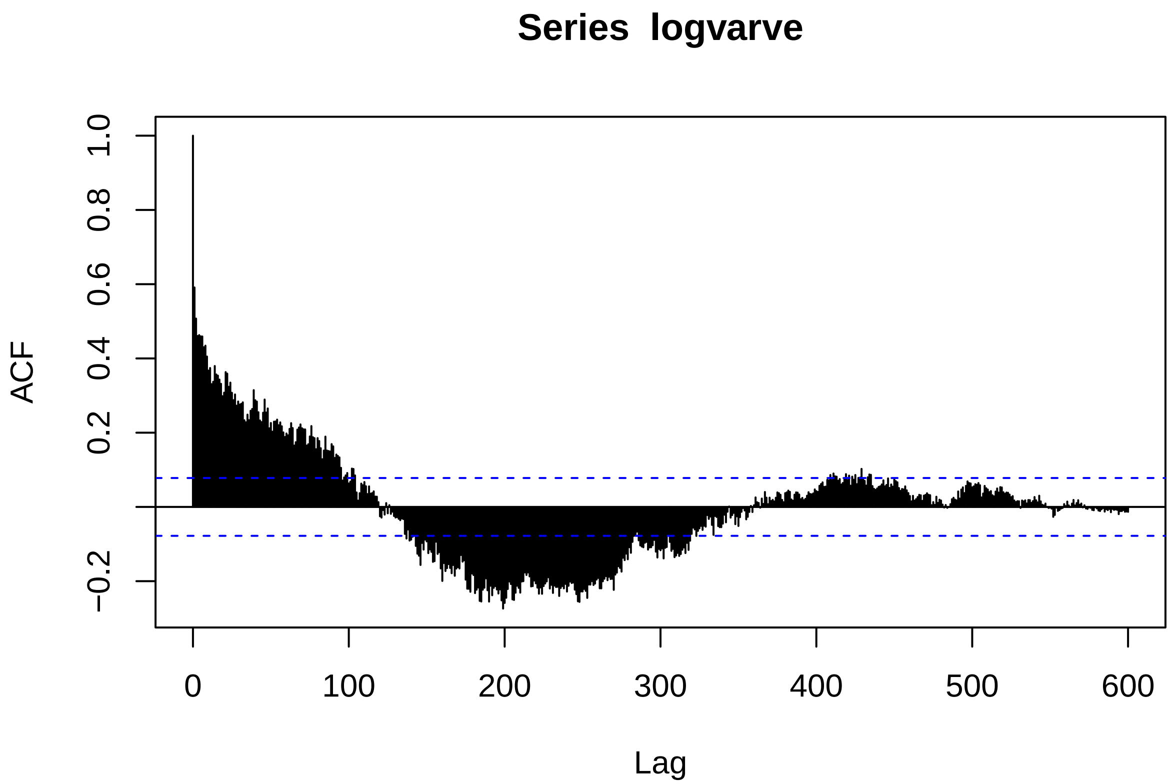
Passando um log nos dados nós conseguimos ver uma aproximação da variância entre a primeira e a segunda metade



##b)



O ideal era demonstrar os qqplots porém não consegui rodar o código para demonstrar eles, porém fica evidente a aproximação da normalidade dos dados após passarmos um logarítmo neles

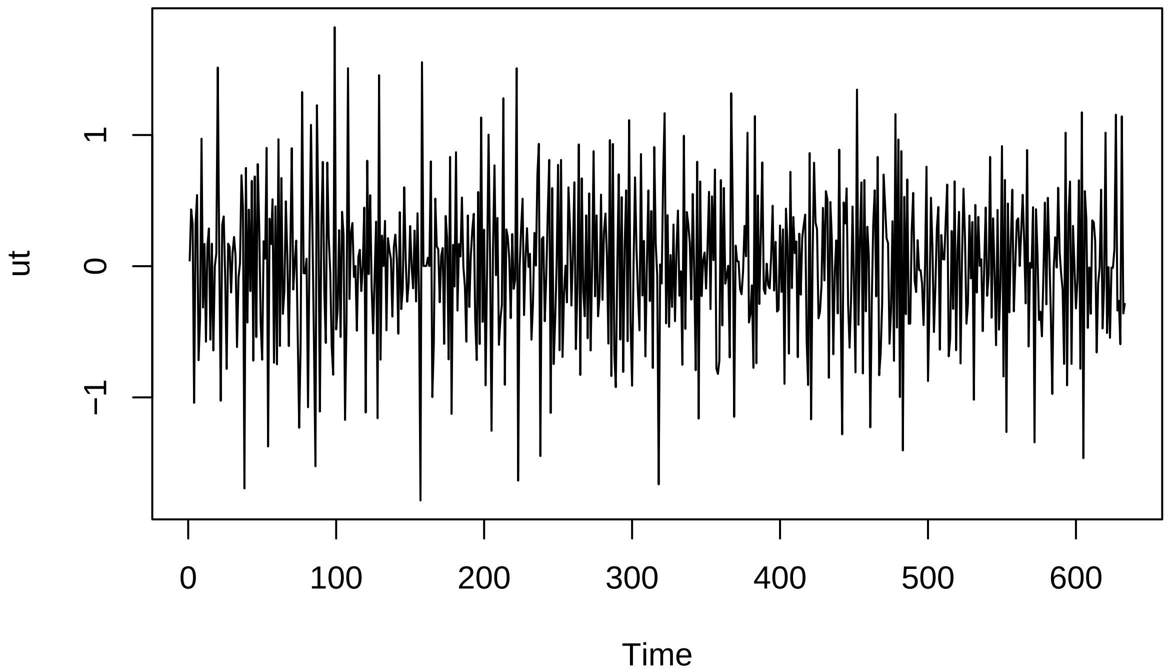


##c)

Olhando com um lag de 600 vemos mudanças de inf+luencia para negativa em torno de 100 a 350 anos atrás, enquanto no resto da série é demonstrado uma ifluencia positiva, bem alta nos ultimos 100 anos.

##d)

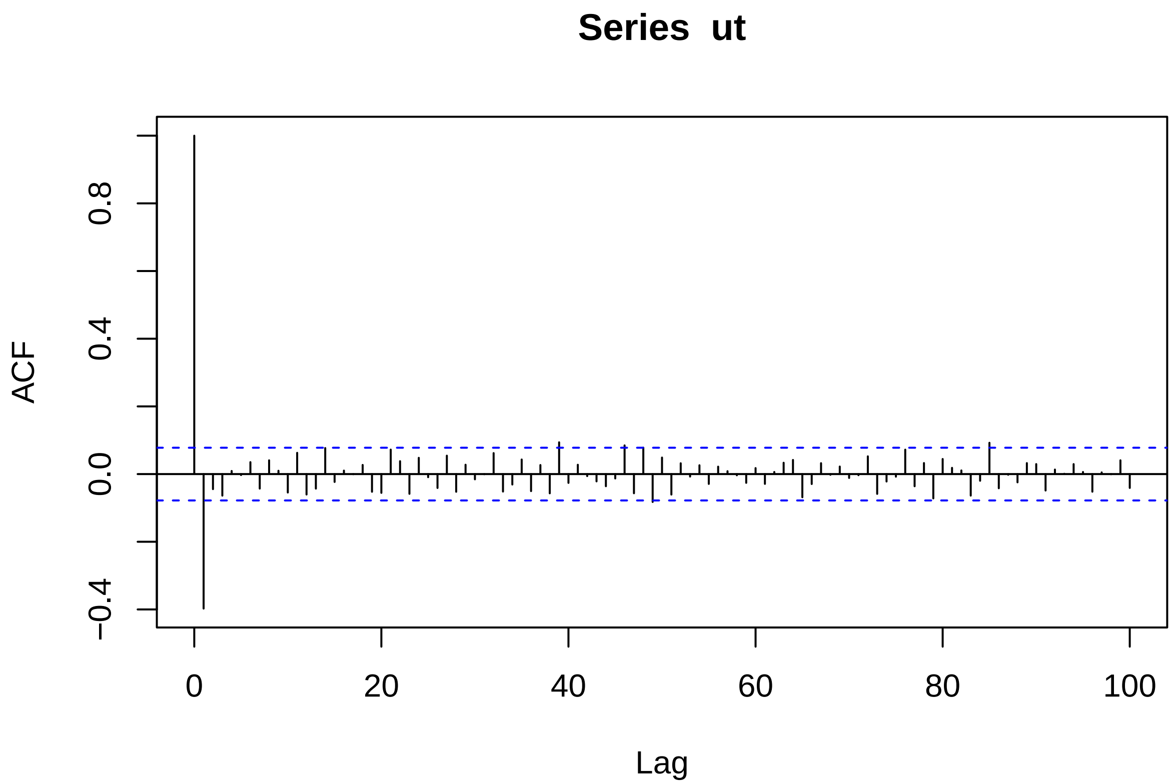
```
ut <- numeric(length(logvarve))
for (i in 2:length(ut)) {
  ut[i] <- logvarve[i] - logvarve[i-1]
}

ut <- ut[-1] #Remover o zero inicial

plot.ts(ut)
```




```
data.frame(
"media" = mean(ut),
"Media primeira metade" = mean(ut[1:(length(ut)/2)]),
"Media segunda metade" = mean(ut[((length(ut)+2)/2):length(ut)]) )
```

```
##          media Media.primeira.metade Media.segunda.metade
## 1 -0.001125366           0.001732166           -0.0030773
```

Dá para notar, pelo gráfico e pela tabela que a média da série samba em torno do zero, parecendo uma certa estacionáriedade.



O ACF amostral demonstra que a série só depende dos ultimos 2 instantes e pouco do resto da série, um compartamenento esperado dada a transformação feita em Ut. Como a série varve já demonstra uma variação de um tempo em relação ao anterior da espessura da areia recolhida, a utilização do Log transforma a variável resposta para uma resposta a nível de escala log, e comparar a diferença de dois tempos numa escala diferente da resposta não me deixou muitas pistas de qual seria a interpretação da série Ut