STA 642 Homework1

Jae Hyun Lee, jl914

20 January, 2020

HW1 for STA-642

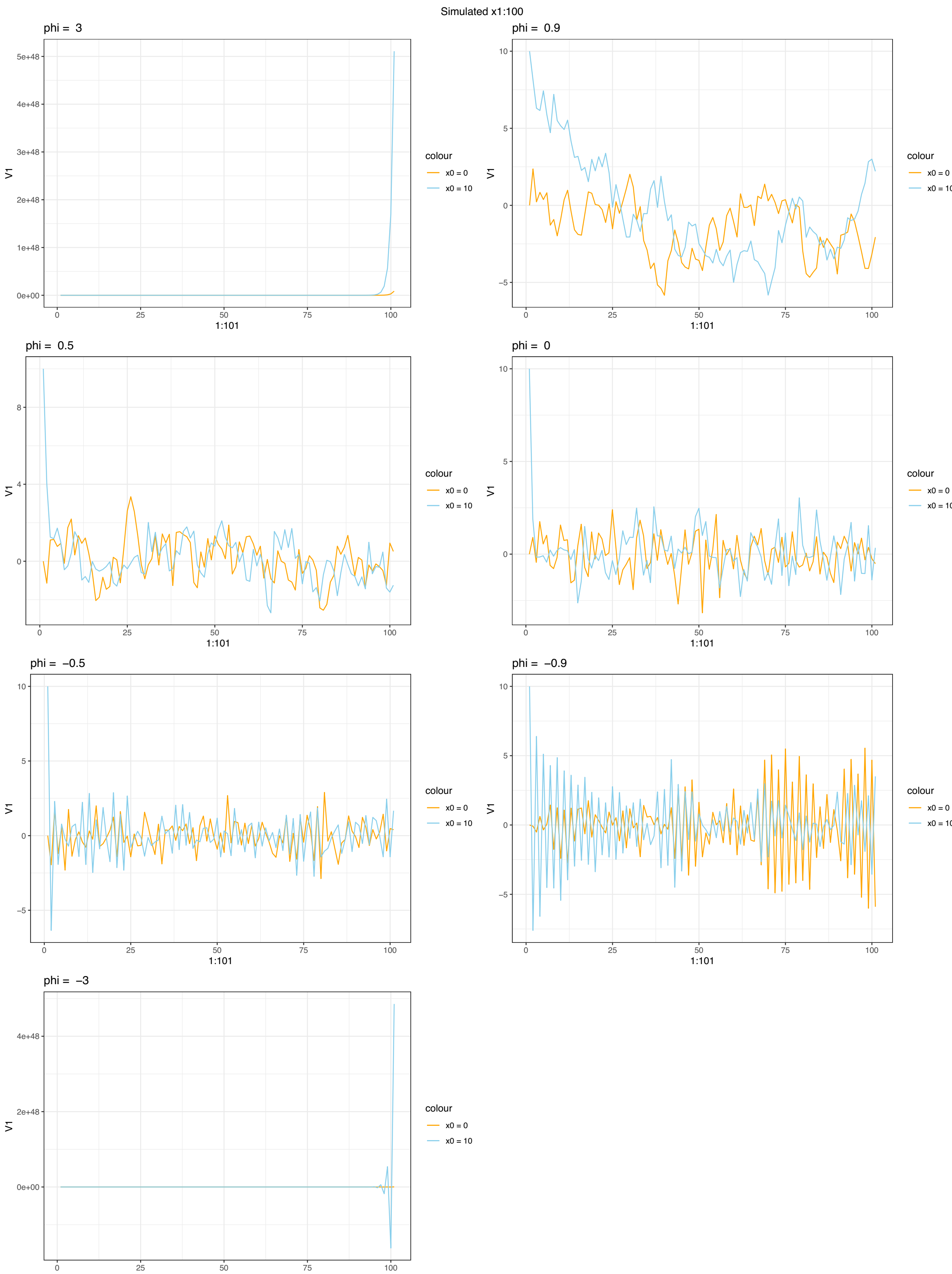
```
library(ggplot2)
library(grid)
library(gridExtra)
```

Exercise 3

I have tested 7 ϕ values 3, 0.9, 0.5, 0 , -0.5, 0, -0.9 and -3 for each simulation of series.

```
pi <- c(3,0.9,0.5,0,-0.5,-0.9,-3)
x0 <- c(0,10)
X <- array(rep(0,101*2*7),dim = c(101,2,7))
X[1,,] <- c(0,10)
set.seed(642)
for (i in 1:7){
  for(j in 2:101){
    X[j,,i] <- c(rnorm(1,X[j-1,1,i]*pi[i],1),rnorm(1,X[j-1,2,i]*pi[i],1))
  }
}
```

```
plots <- list()
for (i in 1:7){
  plots[[i]] <- ggplot(data = as.data.frame(X[,,i])) +
    geom_line(aes(x = 1:101, y = V1, col = "orange")) +
    geom_line(aes(x = 1:101, y = V2, col = "skyblue")) +
    labs(title = paste("phi = ",pi[i])) +
    scale_color_manual(values = c("orange","skyblue"), labels = c("x0 = 0","x0 = 10"))
}
grid.arrange(grobs = plots, ncol = 2, top ="Simulated x1:100", )
```


Simulated x1:100
phi = 3
phi = 0.9
phi = 0.5
phi = 0
phi = −0.5
phi = −0.9
phi = −3
V1
1:101
colour
x0 = 0
x0 = 10
0e+00
1e+48
2e+48
3e+48
4e+48
5e+48
0
25
50
75
100

Effect of absolute size of phi

When $\phi = 3$ or $\phi = -3$ which has large absolute value, the series increases exponentially and becomes explosive and has very big value at last of series. When $\phi = -3$ which is negative value, the series shows oscillation which become explosive and lastly has -1.5e^48. On contrary with previous 2 series, the stationary process which $|\phi| < 1$ show random walks that has most of values between -5 and 5 and difference in initial point seems to be significant. We could find that the decaying rate of the process is much faster and effect of initial point is weak when ϕ has absolute small value. When $\phi < 0$, amplitude of oscillation become small.

Effect of sign of phi

When $0 < \phi < 1$, the process decaying monotonically. On the other hand, when $-1 < \phi < 0$, the process show dumped oscillation.

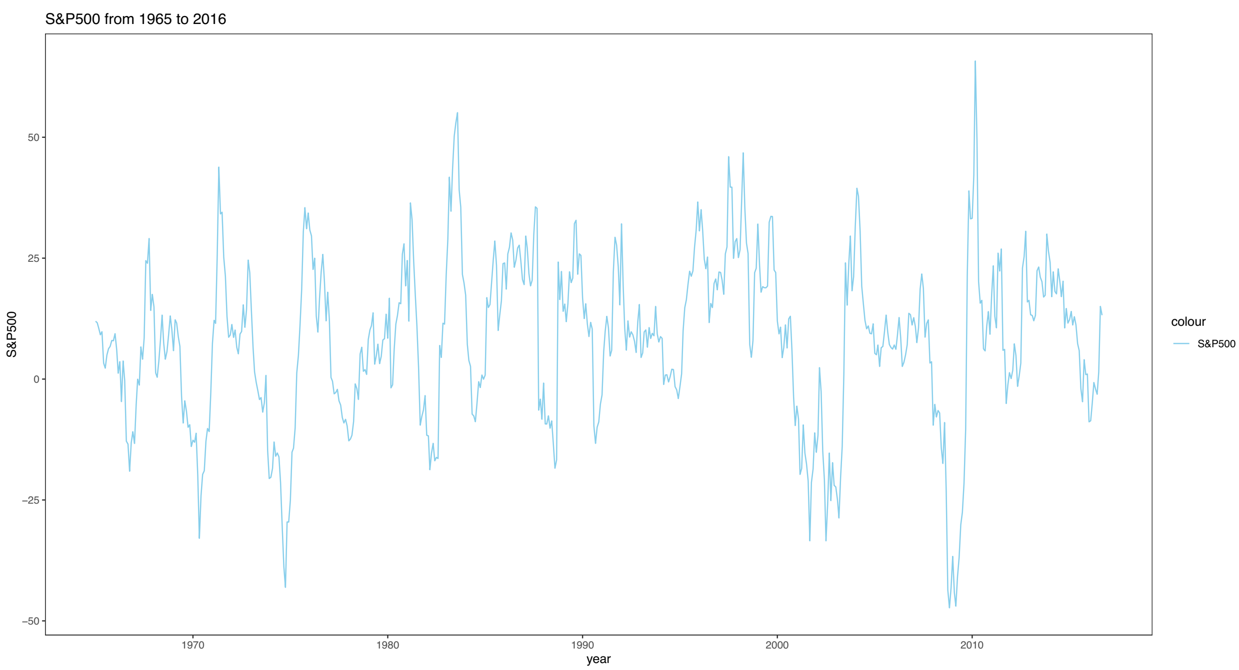
Effect of initial point

When the process is not stationary, initial point does not have effect on the process at all. When the process is stationary, the effect of initial point get decreased as ϕ has small absolute value

Exercise 4

S&P500 process

```r
data <- readxl::read_xlsx("USMacroData1965_2016updated.xlsx")
SP <- data[,length(data)]
ggplot(data = data, aes(x = Date, y = `S&P500`)) +
  geom_line(aes(col = "skyblue")) +
  labs(title = "S&P500 from 1965 to 2016",x = "year") +
  scale_color_manual(values = "skyblue", labels = "S&P500") +
  theme(panel.grid = element_blank())
```



Theoretical posterior distribution of ϕ is as follow:

$$
\begin{aligned}
&\phi \mid y_{1:n}, \mathbf{F} \sim t_{(n-2)}\left(m(y_{1:n}), \frac{C(y_{1:n})}{n-2}\right) \\
&where \quad m(y_{1:n}) = \frac{\sum_{t=2}^{n} y_{t-1}y_t}{\sum_{t=1}^{n-1} y_t^2} \\
&and \quad C(y_{1:n}) = \frac{\sum_{n=2}^{n} y_t^2 \sum_{n=2}^{n} y_{t-1}^2 - (\sum_{t=2}^{n} y_t y_{t-1})^2}{(\sum_{t=1}^{n-1} y_t^2)^2}
\end{aligned}
$$

Functions for subseries and summary statistics

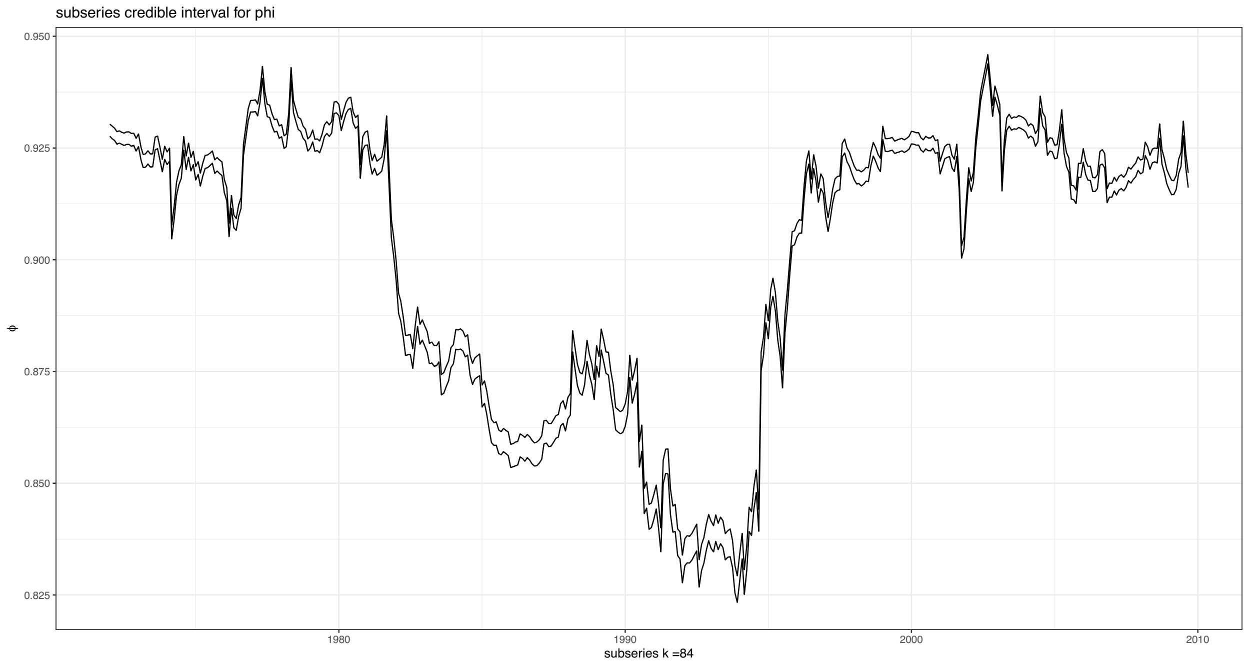
```
m <- function(x){
  x_0 = x[1:(length(x)-1)]
  x_1 = x[2:length(x)]
  result = sum(x_0*x_1)/sum(x_1^2)
  return(result)
}
C <- function(x){
  x_0 = x[1:(length(x)-1)]
  x_1 = x[2:length(x)]
  result = (sum(x_0^2)*sum(x_1^2) - sum(x_0*x_1)^2)/sum(x_1^2)^2
  return(result)
}
df <- function(x){
  return(length(x)-2)
}

subseries <- function(x,k){
  result <- data.frame(matrix(rep(0,(2*k+1)*(nrow(x) - 2*k)), ncol = (nrow(x) - 2*k)))
  num = ncol(result)
  for(i in 1:num){
    result[,i] <- x[i:(i+2*k),]
  }
  return(result)
}

credible <- function(x,m,C){
  values <- rt(5000, df = df(x)) * C / df(x) + m
  return(quantile(values, c(0.05,0.95)))
}
```

```
series_84 <- subseries(SP, 84)
series_84_2 <- apply(series_84, 2, function(x){x - mean(x)})
series_84_date <- as.POSIXct(t(subseries(data[,1],84)[85,]))
intervals <- apply(series_84_2, 2, function(x){credible(x,m(x),C(x))})
```

```
ggplot()+
  geom_line(aes(x = series_84_date, y = intervals[1,])) +
  geom_line(aes(x = series_84_date, y = intervals[2,])) +
  labs(title = "subseries credible interval for phi", y = expression(phi), x = "subseries k =84")
```



(a)

Comment on what you see in the plot and comparison, and what you might conclude in terms of changes over time

By investigating credible plot for ϕ over time, we can conlcude that dependency that x_t has on x_{t-1} had been changed over time, because we could find decreasing trend from 1975 to 1987. After that decreasing, we could also find that there was dramatic increase in dependency at from 1987 to 1990. In addition, we could find that all credible interval are included in boundary that satisfy stationarity $|\phi| < 1$

(b)

Do you believe that short-term changes in S&P have shown real changes in month-month dependencies since 1965?

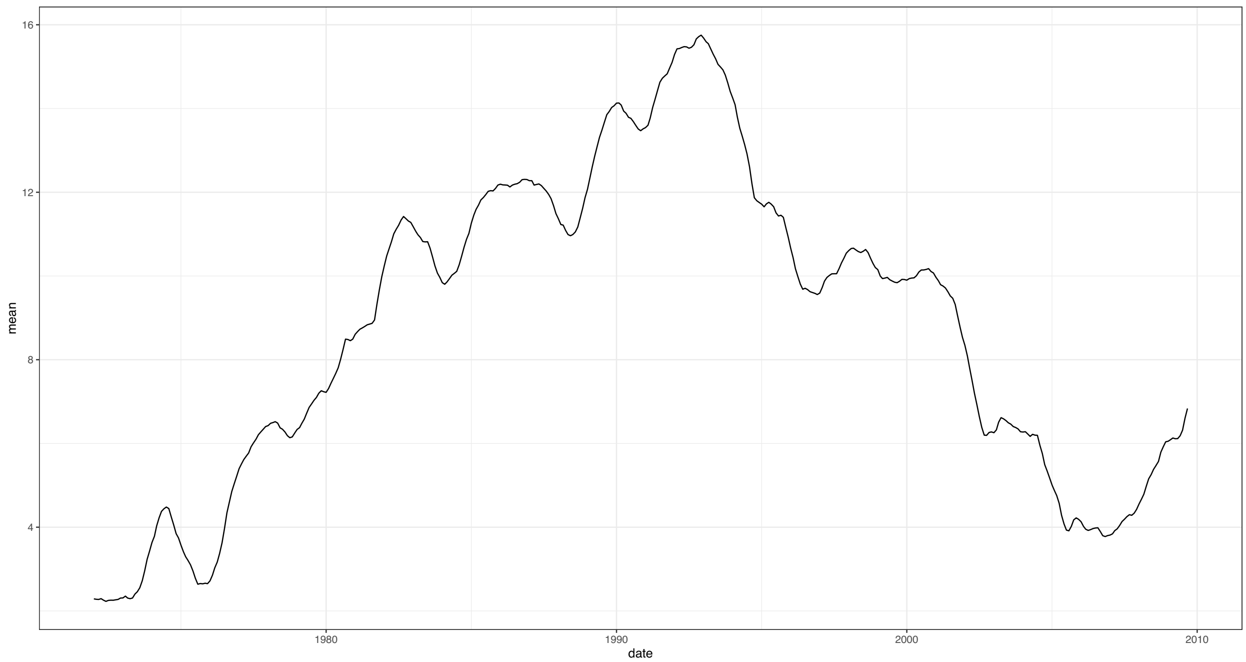
Yes I do, because the short-term is large enough to show stably estimated dependency between x_t and x_{t-1}. We can also find that small amplitude during the terms which show small estimated dependency. i.e from 1990 to 1995 in original S&P500 plot. I think this changes show real change in month-month dependency.

(c)

How would you suggest also addressing the question of whether or not the underlying mean of the series is stable over time?

Weak stationarity is defined that mean, covariance, varianec is stable or constant over all sub-series process. Thus, as we did before, we can show that mean of the series is stable by showing mean of subseries of the process is stable as below. We can find that mean of series is unstable.

```
series_84_mean <- apply(series_84, 2, mean)
ggplot()+
  geom_line(mapping = aes(x = series_84_date, y = series_84_mean)) +
  labs(y = "mean", x = "date")
```

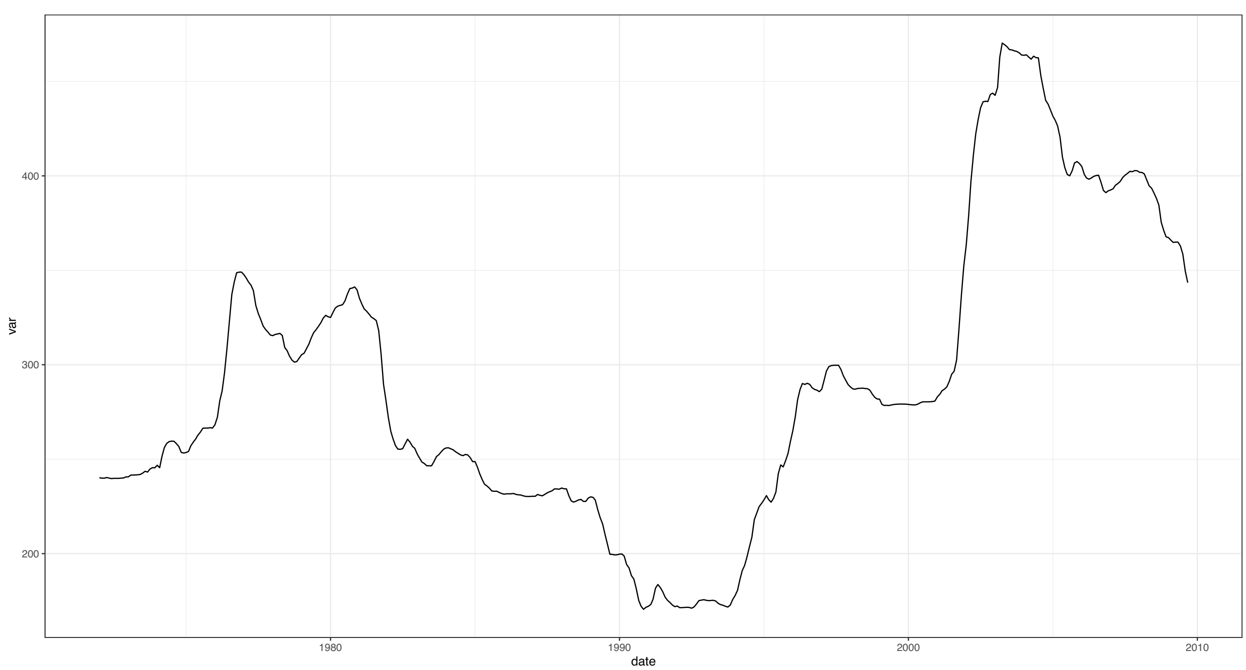


(d)

What about the innovations variance?

We can show innovation variance's stability with same way as (c). We can examine variance of subseries which might show innovation variance over time

```
series_84_var <- apply(series_84, 2, function(x){sd(x)^2})
ggplot()+
  geom_line(mapping = aes(x = series_84_date, y = series_84_var)) +
  labs(y = "var", x = "date")
```



(e)

What does this suggest for more general models that might do a better job of imitating this data?

This subseries check for parameters can confirm whether parameter we want to estimate is changable or constant. By deciding its changability, we can easily choose appropriate model.

Exercise 5

Let $\lambda = \nu^{-1}$. Then

$$\phi \mid \lambda, D_{t-1} \sim N(m_{t-1}, \lambda^{-1}\frac{C_{t-1}}{s_{t-1}}), \quad P(\phi \mid \lambda, D_{t-1}) \propto \lambda^{1/2} exp\{-\frac{\lambda s_{t-1}}{2C_{t-1}}(\phi - m_{t-1})^2\}$$
$$\lambda \mid D_{t-1} \sim G(n_{t-1}/2, n_{t-1}s_{t-1}/2), \quad P(\lambda \mid D_{t-1}) \propto \lambda^{\frac{n_{t-1}}{2}-1} exp\{-\frac{\lambda n_{t-1}s_{t-1}}{2}\}$$
$$\rightarrow P(\phi, \lambda \mid D_{t-1}) \propto \lambda^{\frac{n_{t-1}+1}{2}-1} exp\{-\frac{\lambda s_{t-1}}{2}[n_{t-1} + \frac{(\phi - m_{t-1})^2}{C_{t-1}}]\}$$

(a)

What is the current marginal posterior for ϕ, namely $P(\phi \mid D_{t-1})$?

$$\begin{aligned}
P(\phi \mid D_{t-1}) &= \int P(\phi \mid \lambda, D_{t-1}) P(\lambda \mid D_{t-1}) d\lambda \\
&\propto \lambda^{\frac{n_{t-1}+1}{2}-1} exp\{-\frac{\lambda s_{t-1}}{2}[n_{t-1} + \frac{(\phi - m_{t-1})^2}{C_{t-1}}]\} \\
&= \Gamma\left(\frac{n_{t-1}+1}{2}\right)\left[\frac{s_{t-1}}{2}(n_{t-1} + \frac{(\phi - m_{t-1})^2}{C_{t-1}})\right]^{-\frac{n_{t-1}+1}{2}} \int \Gamma\left(\frac{n_{t-1}+1}{2}\right)^{-1}\left[\frac{s_{t-1}}{2}(n_{t-1} + \frac{(\phi - m_{t-1})^2}{C_{t-1}})\right]^{\frac{n_{t-1}+1}{2}} \lambda \\
&\propto \left[\frac{s_{t-1}}{2}(n_{t-1} + \frac{(\phi - m_{t-1})^2}{C_{t-1}})\right]^{-\frac{n_{t-1}+1}{2}} \\
&\propto \left[1 + \frac{(\phi - m_{t-1})^2}{n_{t-1}C_{t-1}}\right]^{-\frac{n_{t-1}+1}{2}} \quad \rightarrow \phi \sim t(n_{t-1}, m_{t-1}, C_{t-1})
\end{aligned}$$

That is $\frac{(\phi - m_{t-1})^2}{C_{t-1}} \sim t_{n_{t-1}}$ with n_{t-1} degree of freedom

(b)

Let $\frac{s_{t-1}}{C_{t-1}} = k$. Then

$$x_t \mid \phi, \lambda, D_{t-1} \sim N(\phi x_{t-1}, \lambda^{-1})$$
$$\rightarrow P(x_t \mid \phi, \lambda, D_{t-1}) = \lambda^{1/2} exp\{-\frac{\lambda}{2}(x_t - \phi x_{t-1})^2\}$$
$$P(\phi \mid \lambda, D_{t-1}) \propto \lambda^{1/2} exp\{-\frac{\lambda k}{2}(\phi - m_{t-1})^2\}$$

Then,

$$
\begin{aligned}
P(x_t \mid \lambda, D_{t-1}) &= \int P(x_t \mid \phi, \lambda, D_{t-1}) P(\phi \mid \lambda, D_{t-1}) d\phi \\
&\propto \int \lambda exp\{-\frac{\lambda}{2}(\phi^2(x_{t-1}^2 + k) - 2\phi(x_t x_{t-1} + k m_{t-1}) + x_t^2 + k m_{t-1}^2)\} d\phi \\
&\propto \lambda^{1/2} exp\{-\frac{\lambda}{2}(x_t^2 + k m_{t-1}^2 - \frac{(x_t x_{t-1} + k m_{t-1})^2}{x_{t-1}^2 + k})\} \int \lambda^{1/2} exp\{-\frac{\lambda(x_{t-1}^2 + k)}{2}(\phi^2 - 2\phi(\frac{x_t x_{t-1} + k m_{t-1}}{x_{t-1}^2 + k}) + (\\
&\propto \lambda^{1/2} exp\{-\frac{\lambda}{2}(x_t^2 + k m_{t-1}^2 - \frac{(x_t x_{t-1} + k m_{t-1})^2}{x_{t-1}^2 + k})\} \\
&\propto \lambda^{1/2} exp\{-\frac{\lambda}{2}(x_t^2(1 - \frac{x_{t-1}^2}{x_{t-1}^2 + k}) - 2x_t \frac{k x_{t-1} m_{t-1}}{x_{t-1}^2 + k} + k m_{t-1}^2 (1 - \frac{k}{x_{t-1}^2 + k}))\} \\
&\propto \lambda^{1/2} exp\{\frac{\lambda k}{2(x_{t-1}^2 + k)}(x_t^2 - 2 x_t x_{t-1} m_{t-1} + x_{t-1}^2 m_{t-1}^2)\} \\
&\to x_t \sim N(x_{t-1} m_{t-1}, \lambda^{-1}(1 + x_{t-1}/k)), \quad where\ \lambda^{-1}(1 + x_{t-1}/k)) = \nu(1 + \frac{C_{t-1}}{s_{t-1}} x_{t-1}^2) = \frac{\nu}{s_{t-1}} q_t \\
&\to x_t \sim N(f_t, q_t v / s_{t-1})
\end{aligned}
$$

(c)

$P(x_t \mid \lambda, D_{t-1}) \propto \lambda^{1/2} exp\{-\frac{\lambda k}{2(x_{t-1}^2 + k)}(x_t - x_{t-1} m_{t-1})^2\}$ then,

$$
\begin{aligned}
P(x_t \mid D_{t-1}) &\propto \int P(x_t \mid \lambda, D_{t-1}) P(\lambda \mid D_{t-1}) d\lambda \\
&\propto \int \lambda^{\frac{n_{t-1}+1}{2} - 1} exp\{-\frac{\lambda}{2}\left[n_{t-1} s_{t-1} + \frac{k}{x_{t-1}^2 + k}(x_t - x_{t-1} m_{t-1})^2\right]\} \\
&\propto \left[n_{t-1} s_{t-1} + \frac{k}{x_{t-1}^2 + k}(x_t - x_{t-1} m_{t-1})^2\right]^{-\frac{n_{t-1}+1}{2}} \\
&\propto \left[1 + 1/n_{t-1} \times 1/(C_{t-1} x_{t-1}^2 + s_{t-1}) \times (x_t - x_{t-1} m_{t-1})^2\right]^{-\frac{n_{t-1}+1}{2}} \\
&\to x_t \sim t(n_{t-1}, x_{t-1} m_{t-1}, C_{t-1} x_{t-1}^2 + s_{t-1})
\end{aligned}
$$

(d)

$m_t = m_{t-1} + A_t e_t$ $C_t = r_t(C_{t-1} - A_t^2 q_t)$ $n_t = n_{t-1} + 1$ $s_t = r_t s_{t-1}$ where $e_t = x_t - f_t$ and $A_t = C_{t-1} x_{t-1} / q_t$

$$
\begin{aligned}
P(\phi, \lambda \mid D_t) &= P(\phi, \lambda \mid x_t, D_{t-1}) \\
&= \frac{P(\phi, \lambda, x_t \mid D_{t-1})}{P(x_t \mid D_{t-1})} \\
&= \frac{p(x_t \mid \phi, \lambda, D_{t-1}) P(\phi \mid \lambda, D_{t-1}) P(\lambda \mid D_{t-1})}{P(x_t \mid D_{t-1})} \\
&\propto \frac{\lambda^{1/2} exp\{-\frac{\lambda(x_t - \phi x_{t-1})^2}{2}\} (\lambda k)^{1/2} exp\{-\frac{\lambda k(\phi - m_{t-1})^2}{2} \lambda^{n_{t-1}/2}\} exp\{-\frac{\lambda n_{t-1} s_{t-1}}{2}\}}{[1 + \frac{x_t - x_{t-1} m_{t-1}}{n_{t-1} \times (C_{t-1} x_{t-1}^2 + s_{t-1})}]^{-\frac{n_{t-1}+1}{2}}} \\
&\propto \lambda^{1/2} exp\{-\frac{\lambda(x_t - \phi x_{t-1})^2}{2}\} (\lambda k)^{1/2} exp\{-\frac{\lambda k(\phi - m_{t-1})^2}{2} \lambda^{n_{t-1}/2}\} exp\{-\frac{\lambda n_{t-1} s_{t-1}}{2}\} \\
&= (\lambda k)^{1/2} exp\{-\frac{\lambda k}{2}(\phi^2 - 2m_{t-1}\phi + \phi^2 x_{t-1}/k - 2\phi x_t x_{t-1}/k)\} \times \lambda^{n_{t-1}+1/2} exp\{-\frac{\lambda}{2}(n_{t-1}s_{t-1} + km_{t-1}^2 + x_t^2)\} \\
&= (\lambda k)^{1/2} exp\{-\frac{\lambda(k + x_{t-1}^2)}{2}(\phi^2 - 2\phi \frac{km_{t-1} + x_t x_{t-1}}{k + x_{t-1}^2} + (\frac{km_{t-1} + x_t x_{t-1}}{k + x_{t-1}^2})^2)\} \times \lambda^{n_{t-1}+1/2} exp\{-\frac{\lambda}{2}(n_{t-1} + s_{t-1} - \\
&\to P(\phi \mid \lambda, D_t) P(\lambda \mid D_t) \\
P(\phi \mid \lambda, D_t) &\sim N(\frac{km_{t-1} + x_t x_{t-1}}{k + x_{t-1}^2}, \frac{\nu}{k + x_{t-1}^2}), \quad P(\lambda \mid D_t) \sim Gamma(\frac{n_{t-1}+1}{2}, (n_{t-1} + s_{t-1} + km_{t-1}^2 + x_t^2 - \frac{(k
\end{aligned}
$$

we can find that

$$
\begin{aligned}
n_t &= n_{t-1} + 1 \\
n_t s_t &= n_{t-1}s_{t-1} + km_{t-1}^2 + x_t^2 - \frac{k^2 m_{t-1}^2 + 2km_{t-1}x_t x_{t-1} + x_t^2 x_{t-1}^2}{k + x_{t-1}^2} \\
&= n_{t-1}s_{t-1} + \frac{1}{k + x_{t-1}^2}(k^2 m_{t-1}^2 + km_{t-1}^2 x_{t-1}^2 + kx_t^2 + x_t^2 x_{t-1}^2 - k^2 m_{t-1}^2 - 2km_{t-1}x_t x_{t-1} - x_t^2 x_{t-1}^2) \\
&= n_{t-1}s_{t-1} + \frac{k}{k + x_{t-1}^2}(x_t - m_{t-1}x_{t-1})^2 \\
&= n_{t-1}s_{t-1} + \frac{s_{t-1}}{s_{t-1} + C_{t-1}x_{t-1}^2}(x_t - m_{t-1}x_{t-1})^2 \\
&= s_{t-1}(n_{t-1} + \frac{e_t^2}{q_t}) \\
&\to s_t = \frac{s_{t-1}}{n_t}(n_{t-1} + \frac{e_t^2}{q_t})
\end{aligned}
$$

$$
C_t = s_t/(k + x_{t-1}^2) = r_t s_{t-1}/(k + x_{t-1}^2) = r_t C_{t-1} s_{t-1}/q_t = r_t(C_{t-1} - A_t^2 q_t)
$$

$$
\begin{aligned}
m_t &= \frac{km_{t-1} + x_t x_{t-1}}{k + x_{t-1}^2} \\
&= \frac{s_{t-1}m_{t-1} + C_{t-1}x_t x_{t-1}}{s_{t-1} + C_{t-1} + x_{t-1}^2} \\
&= m_{t-1} + \frac{C_{t-1}x_t x_{t-1} - m_{t-1}C_{t-1}x_{t-1}^2}{s_{t-1} + C_{t-1} + x_{t-1}^2} \\
&= m_{t-1} + \frac{C_{t-1}x_{t-1}(x_t - m_{t-1}x_{t-1})}{s_{t-1} + C_{t-1} + x_{t-1}^2} = m_{t-1} + A_t e_t
\end{aligned}
$$

from above factorized pdf

(e)

1&2. Above expression, $m_t = m_{t-1} + A_t e_t = A_t x_t + (1 - A_t) m_{t-1}$ shows that m_t depends on x_t as much as adaptive coefficent A_t relative to prior mean. It means that A_t is weight given to new observation. In my opinion, r_t is indicator of how much uncertainty of process is changed by new observation by introducing realized error. With r_t, C_t depends on C_{t-1} and we can confirm that C_t was reduced by adaptive coefficient square A_t^2 and q_t which can only positive values. That means , in the case of forecast error is not that large i.e. r_t is unchanged, C_t is reduced by new observations. Reduced amount would be proportional to A_t^2

3. We can find that n_t is increased as much as the number of new observations x_t. For s_t, it is affected indirectly by x_t by forecast error $e_t = x_t - f_t$. If forecast error is large which means realized x_t is largely different from what expected, s_t is increased. On the contrary, if x_t is similar with our expectation, s_t would be reduced.

(f)

In the case that the forecast error is so large that e_t^2/q_t is larger than 1, s_t which is point estimator of innovation variance would be increased. Moreover, since C_t depends on s_t, it would be also increased. As result, the uncertainty of posterior distribution of (ϕ, ν) would increase.