Problem 2 (10 credits)

HW2

Danny Moncada (monca016)

February 24, 2020

```r
suppressWarnings(suppressPackageStartupMessages({
  library(TSA)
  library(forecast)
  library(ggplot2)
  library(dplyr)
}))
```

Characteristic Polynomials

Question 1

Assume Y_t is the following stochastic process such as

$$Y_t = 2.2 \cdot Y_{t-1} - 1.57 \cdot Y_{t-2} + 0.36 \cdot Y_{t-3} + e_t$$

where $e_i \sim N(0,1)$ i.i.d

a) (1 credit)

First, let's determine whether the process is stationary or not by computing the roots of the characteristic polynomial.

Hints:

- use `polyroot()` function

Please pick the smallest root as x_1 and the larger root as x_3:

```r
"Based on the stochastic process equation, we rewrite it as 1 - 2.2 + 1.57 - 0.36."
```

```
## [1] "Based on the stochastic process equation, we rewrite it as 1 - 2.2 + 1.57 - 0.36."
```

```r
"Does this mean that the largest root is 2.2?  That's the assumption I'm making."
```

```
## [1] "Does this mean that the largest root is 2.2?  That's the assumption I'm making."
```

```r
polyroot(c(1, -2.2, 1.57, -0.36))
```

```
## [1] 1.111111-0i 1.250000+0i 2.000000-0i
```

```r
x_1 <- 2.0
x_2 <- 1.25
x_3 <- 1.11
```

(please include only the numerical answer above not the computation. An acceptable format for your answer is such as `x_1 <- 5`)

b) (1 credit)

Based on your answer above conclude whether the process is stationary or not:

```
stationary <- TRUE # type a boolean: TRUE or FALSE
```

c) (2 credits)

Please generate $N = 100$ sample paths of length $T = 100$ for this stochastic process.

- Please save the results into a data.frame `df2c` where:
  - column `df2c$Y` has the values of the process
  - column `df2c$id` has the id of the sample path
  - column `df2c$t` has the time

```
set.seed(42) # Please do not change the seed

N <- 100L
T <- 100L

df2c <- data.frame(Y = arima.sim(model=list(ar=c(2.2, -1.57, 0.36), sd = 1), n = N*T),
                   id = rep(1:N, each = T),
                   t= rep(1:T, N))
head(df2c)
```

```
##            Y id t
## 1 -3.0438630  1 1
## 2 -4.5287222  1 2
## 3 -5.0575687  1 3
## 4 -3.7126110  1 4
## 5 -2.5849935  1 5
## 6 -0.3763685  1 6
```

d) (1 credit)

Please plot the sample paths that you generated in the previous question

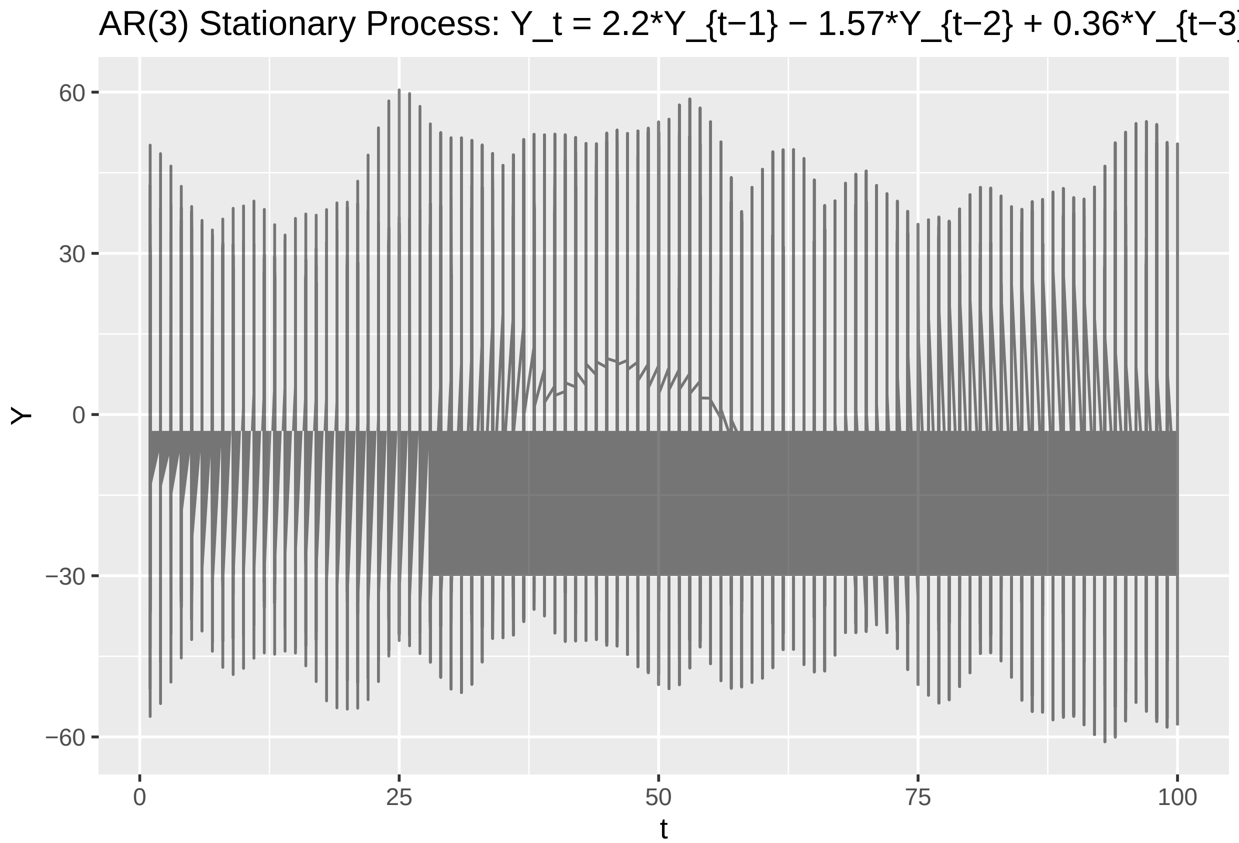
- Please save your plot into variable `p1d`

Hints:

- use `ggplot` and take advantage of the long format of the data
- please don't change the color (keep the lines black) but do put `alpha=0.05` into your `geom_line` to make sample paths somewhat transparent.
- do not use `geom_points` just `geom_line` is fine
- As you will see from your plot:
  - the fainter the line the less likely the stochastic process would reach this spot

```
p1d <- ggplot(data = df2c, aes(x = t, y = Y)) + geom_line(alpha = 0.5) +
         ggtitle("AR(3) Stationary Process: Y_t = 2.2*Y_{t-1} - 1.57*Y_{t-2} + 0.36*Y_{t-3} + e_t")
p1d
```

```
## Don't know how to automatically pick scale for object of type ts. Defaulting to continuous.
```



Question 2 (5 credits)

Repeat a) - d) in Question 1 for the following stochastic process Y_t:

$$Y_t = 2.4 \cdot Y_{t-1} - 1.55 \cdot Y_{t-2} + 0.3 \cdot Y_{t-3} + e_t$$

where $e_i \sim N(0, 1)$ i.i.d

Compared with Question 1, we expect to see significant difference in the stationarity from the plot, although the coefficients are very close.

a)

First, let's determine whether the process is stationary or not by computing the roots of the characteristic polynomial.

Hints:

- use `polyroot()` function

Please pick the smallest root as x_1 and the larger root as x_3:

```
"Based on the stochastic process equation, we rewrite it as 1 - 2.4 + 1.55 - 0.3."
```

```
## [1] "Based on the stochastic process equation, we rewrite it as 1 - 2.4 + 1.55 - 0.3."
```

```
polyroot(c(1, -2.4, 1.55, -0.3))
```

```
## [1] 0.6666667+0i 2.0000000-0i 2.5000000+0i
```

```r
x_1 <- 2.5
x_2 <- 2.0
x_3 <- 0.66
```

(please include only the numerical answer above not the computation. An acceptable format for your answer is such as `x_1 <- 5`)

b)

Based on your answer above conclude whether the process is stationary or not:

```r
stationary <- FALSE # type a boolean: TRUE or FALSE
```

c)

Please generate $N = 100$ sample paths of length $T = 20$ for this stochastic process.

- Please save the results into a data.frame `df2c` where:
  - column `df2c$Y` has the values of the process
  - column `df2c$id` has the id of the sample path
  - column `df2c$t` has the time

```r
set.seed(42) # Please do not change the seed

N <- 100L
T <- 20L

Y = c()
for(i in 1:T){

    x <- numeric(100)
    x[1] = x[2] = x[3] = rnorm(1)
    for(i in 4: length(x)) {
      x[i] <- 2.4*x[i-1] - 1.55*x[i-2] + 0.3*x[i-3] + rnorm(1)
    }
    Y = append(Y, x)
}

df2c <- data.frame(Y = Y, id = rep(1:N, each = T), t= rep(1:T, N))
head(df2c)
```

```
##          Y id t
## 1 1.370958  1 1
## 2 1.370958  1 2
## 3 1.370958  1 3
## 4 1.011904  1 4
## 5 1.078000  1 5
## 6 2.062899  1 6
```

d)

Please plot the sample paths that you generated in the previous question. You should see the effect of the roots of the polynomial on the sample paths of the process.

- Please save your plot into variable `p1d`

Hints:

- use `ggplot` and take advantage of the long format of the data
- do not use `geom_points` just `geom_line` is fine

```
p1d <- ggplot(data = df2c, aes(x = t, y = Y)) + geom_line(alpha = 0.5) +
          ggtitle("Non Stationary Process: Y_t = 2.4*Y_{t-1} - 1.55*Y_{t-2} + 0.3*Y_{t-3} + e_t")
p1d
```