Taller dos de Estadística III

Simón Cuartas Rendón

Marzo de 2022

```
rm(list = ls())

# Disponibilización de librerías importantes
library(forecast)
```

```
## Registered S3 method overwritten by 'quantmod':
##   method            from
##   as.zoo.data.frame zoo
```

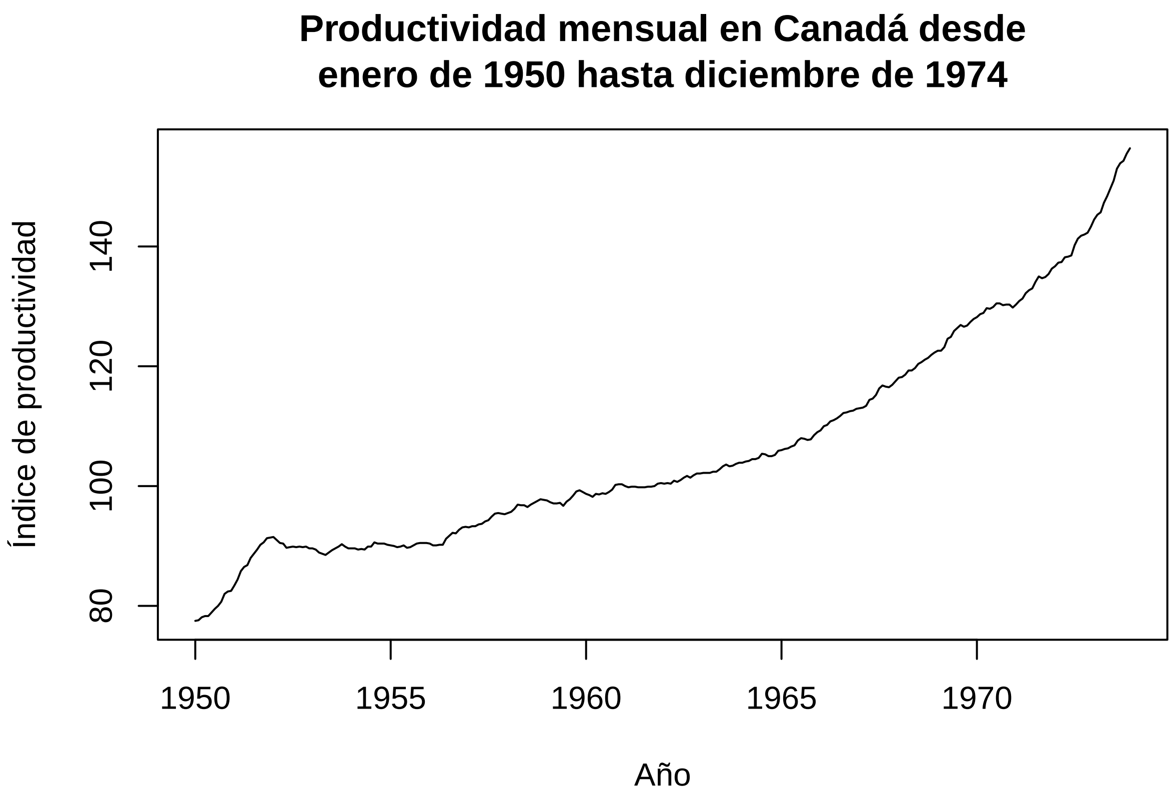
```
# Disponibilización de funciones de usuario
source("https://raw.githubusercontent.com/NelfiGonzalez/Funciones-de-Usuario-Estadistica-III/main/Funci
source("https://raw.githubusercontent.com/NelfiGonzalez/Funciones-de-Usuario-Estadistica-III/main/Funci
```

```
# Lectura de los datos
produccion = ts(scan(file.choose(), dec = "."),
                frequency=12,
                start = c(1950, 1))
```

Punto uno. ¿Qué tipo de tendencia tiene la serie?

Se tiene una base de datos sobre el índice de productividad mensual de Canadá entre los años 1950 y 1974, por lo que se tiene un total de 288 datos, y cuya serie luce como se observa a continuación:

```
plot(produccion,
     main = "Productividad mensual en Canadá desde\nenero de 1950 hasta diciembre de 1974",
     xlab = "Año",
     ylab = "Índice de productividad")
```



Como se puede observar, esta tendencia tiene una tendencia creciente, ya que, en general, el índice de productividad está aumentando mes a mes salvo por algunos pocos meses en los que esto no ocurre. Asimismo, es importante advertir que no existe una tendencia lineal y que esta parece más bien ser polinomial de grado tres o superior, exponencial o logpolinomial.

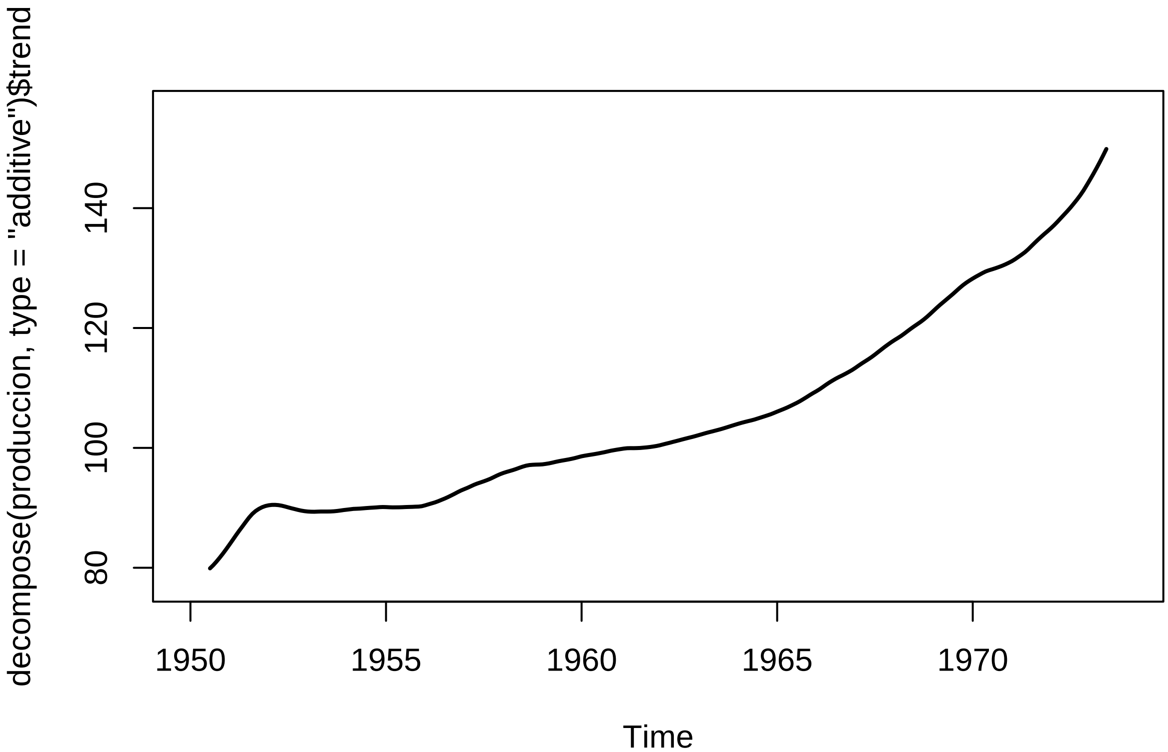
Por otro lado, es importante mencionar que en la serie se puede observar una intervención hacia el año de 1952, pues durante los primeros dos años de la serie se puede ver un crecimiento constante hasta que ocurre dicha perturbación, a partir de la cual el crecimiento del índice de productividad se estanca con un índice que crece lentamente hasta inicios de la década de los sesenta, cuando comienza a acelerarse nuevamente el aumento de este índice para Canadá. Con lo anterior, se puede decir que la tendencia de esta serie es local.

Por otro lado, se puede decir que la tendencia es determínistica, ya que los datos pueden ser representados con ayuda de una función suave del tiempo de forma sencilla.

Punto dos. Estacionalidad.

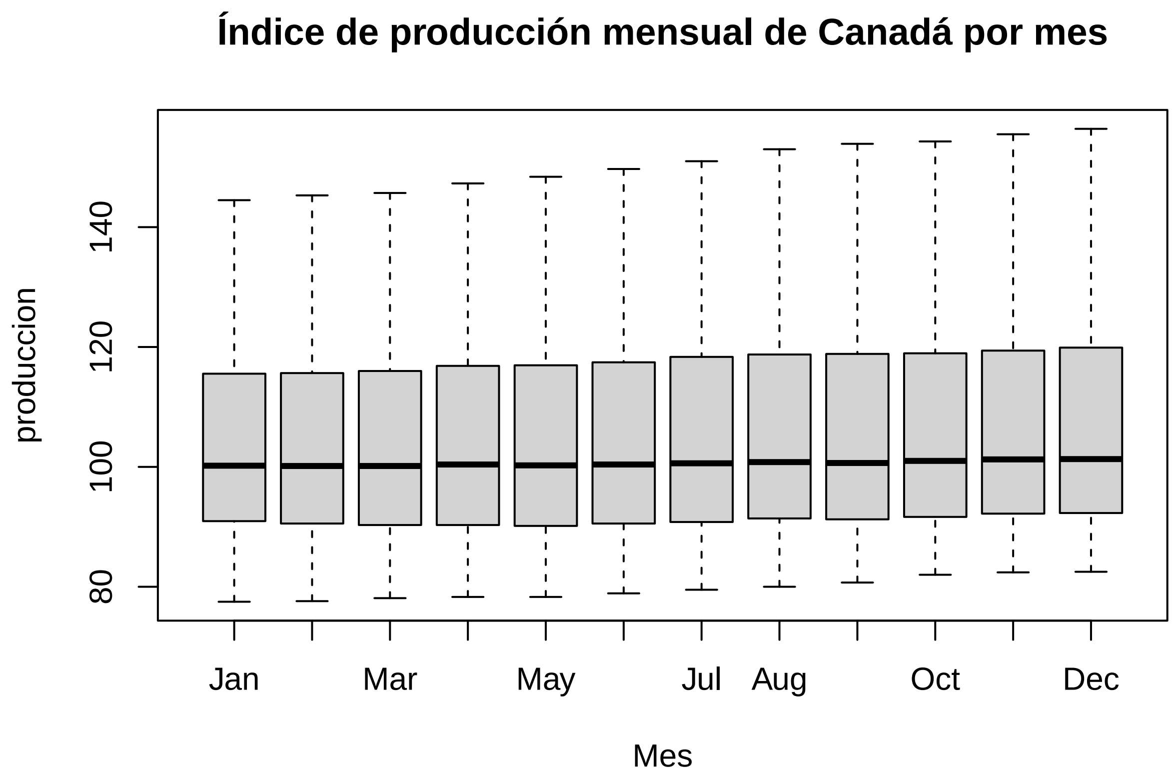
A partir de lo que se ve en la gráfica se puede juzgar que no existe una componente estacional, ya que no se observa en esta un cambio regular en cada año. En todo caso es interesante ver la componente de la tendencia en la descomposición de esta serie en cada una de sus componentes:

```
plot(decompose(produccion,
               type = "additive")$trend,
               ylim = c(min(produccion), max(produccion)),lwd = 2)
```



Y se puede visualizar, al comparar esta figura con la primera, que conserva muy bien sus propiedades. En todo caso, para estar seguros de cómo se está comportando la estacionalidad, si la hay, es útil chequear los gráficos de cajas y bigotes del índice de producción de Canada por mes.

```
boxplot(produccion~cycle(produccion),names=month.abb,
        main = "Índice de producción mensual de Canadá por mes",
        xlab = "Mes",
        y = "Índice de producción")
```

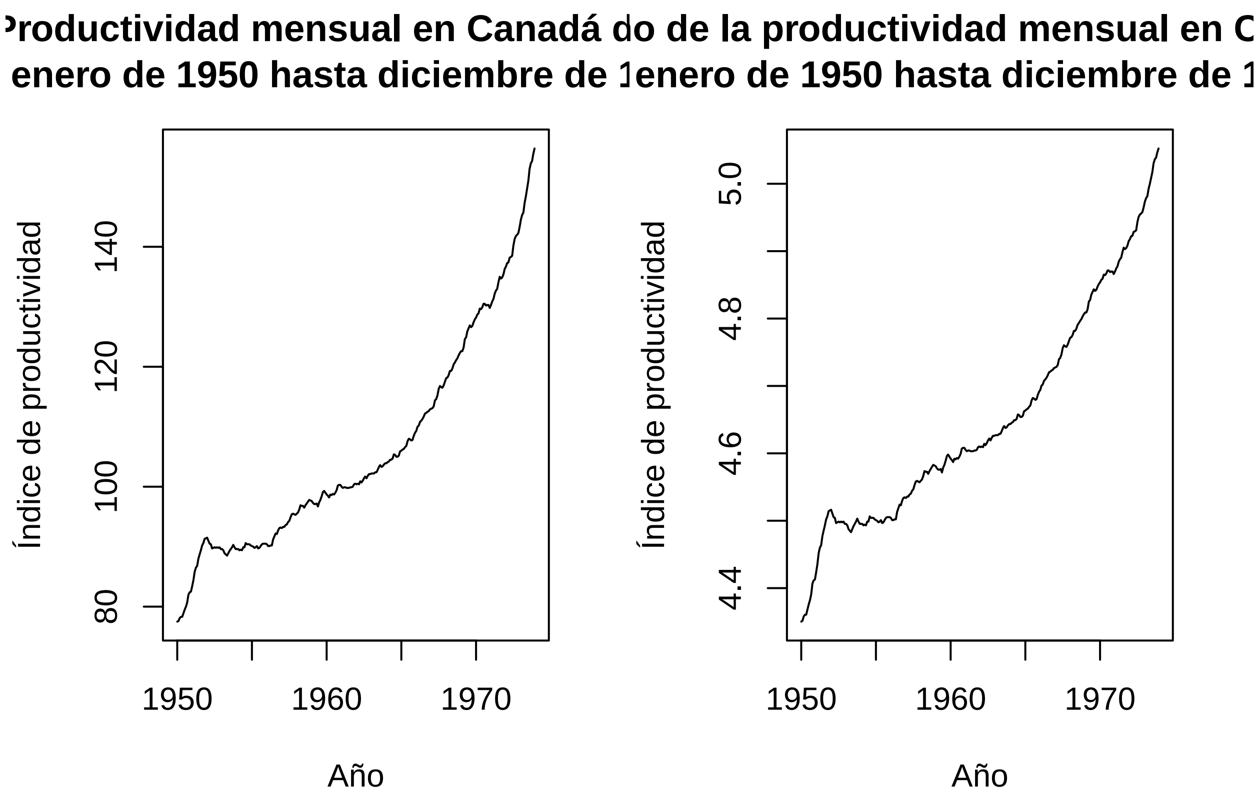


Y se puede detallar que la mediana es prácticamente la misma para cada uno de los meses, por lo que se puede afirmar que esta serie de tiempo carece de estacionalidad.

Punto tres. ¿Es la serie de componentes aditivas o multiplicativas?

Para definir esto es necesario observar la serie de tiempo con sus unidades originales y luego transformada por un logaritmo natural, de tal forma que pueda ser verificada el comportamiento de la varianza en cada una de ellas, lo cual se observa a continuación:

```
par(mfrow = c(1, 2))
plot(produccion,
     main = "Productividad mensual en Canadá desde\nenero de 1950 hasta diciembre de 1974",
     xlab = "Año",
     ylab = "Índice de productividad")
plot(log(produccion),
     main = "Logaritmo de la productividad mensual en Canadá desde\nenero de 1950 hasta diciembre de 19
     xlab = "Año",
     ylab = "Índice de productividad")
```



Esta serie se destaca por no tener especialmente una gran varianza, y en el gráfico anterior se observa que luce de forma muy semejante en la escala regular como en la logarítmica, por lo que es razonable afirmar que la tendencia es *aditiva.*

Punto cuatro. Planteamiento del modelo.

Para esta serie de tiempo se van a proponer cinco modelos distintos, a saber:

- **Modelo uno.** Model cúbico. $Y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + E_i, \quad E_i \overset{i.i.d.}{\sim} Normal(0,\ \sigma^2)$.
- **Modelo dos** Modelo polinomial de grado cuatro. $Y_t = \beta_0 + (\sum_{j=1}^{4} \beta_j t^j) + E_i, \quad E_i \overset{i.i.d.}{\sim} Normal(0,\ \sigma^2)$.
- **Modelo tres.** Modelo polinomial de grado seis.$Y_t = \beta_0 + (\sum_{j=1}^{6} \beta_j t^j) + E_i, \quad E_i \overset{i.i.d.}{\sim} Normal(0,\ \sigma^2)$.
- **Modelo cuatro.** Modelo exponencial polinomial de grado seis. $Y_t = exp(\beta_0 + (\sum_{j=1}^{6} \beta_j t^j)) + E_i, \quad E_i \overset{i.i.d.}{\sim} Normal(0,\ \sigma^2)$.
- **Modelo cinco.** Modelo logpolinomial de grado seis. $log(Y_t) = \beta_0 + (\sum_{j=1}^{6} \beta_j t^j) + E_i, \quad E_i \overset{i.i.d.}{\sim} Normal(0,\ \sigma^2)$.