

# Series de Tiempo 2018

Maestría en Estadística Aplicada, UNR  
Unidad 4

---

Luis Damiano

[damiano.luis@gmail.com](mailto:damiano.luis@gmail.com)

2018-05-04

- Procesos no estacionarios en la media
  - Tendencia lineal
  - Tendencia estocástica
  - Ejercicio: Energías alternativas
- Procesos no estacionarios en la varianza
  - Transformaciones.
  - Ejercicio: Ventas de supermercados

# Procesos no estacionarios

---

Un proceso **débilmente** estacionario de segundo orden:

- La media y la varianza son constantes.
- La covariancia y la correlación sólo son función de la diferencias de tiempo ( $k$ ) entre las dos variables.

## Discusión en clases

¿De qué forma podrían fallar cada uno de los requisitos de la definición?

# Tipos de no estacionariedad (continuación)

- Procesos no estacionarios en la media
  - Tendencia determinística
  - Tendencia estocástica
  - Tendencia que varía en función del tiempo
- Procesos no estacionarios en la varianza
  - Transformaciones para estabilizar la varianza

# ¿Cómo encarar una serie?

1. Primero y principal, ¿la teoría subyacente indica algo sobre la estacionariedad?
2. Inspección visual (serie original, ACF, PACF).
3. Ajustar ARMA si la serie:
  - No muestra una violación aparente a los supuestos de estacionariedad, y
  - Tiene una función de autocovariancia que decae rápidamente.
4. Si el valor medio muestra variación por tendencia o estacionalidad, capturar el componente (ej. aplicar descomposición, diferenciar).
5. Si la varianza muestra variación, buscar una transformación para estabilizarla.

# Procesos no estacionarios en la media

---

$$Z_t = \alpha_0 + \alpha_1 t + a_t, \quad t = 0, \pm 1, \dots, \quad a_t \sim \mathcal{N}(0, \sigma^2)$$

```
simTL <- function(alpha0, alpha1, sigma, T) {  
  # Ruido  
  at <- rnorm(T, 0, sigma)  
  
  # Observaciones  
  t  <- 1:T  
  Zt <- alpha0 + alpha1 * t + at  
  
  # Devolver observaciones simuladas  
  Zt  
}
```

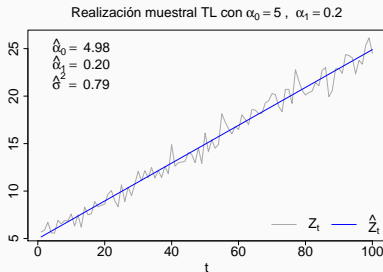


# Ejemplo

$$Z_t = 5 + 0.2t + a_t, \quad t = 0, \pm 1, \dots, \quad a_t \sim \mathcal{N}(0, 1)$$

```
set.seed(9000)
z <- simTL(5, 0.2, 1, 100)
t <- 1:100

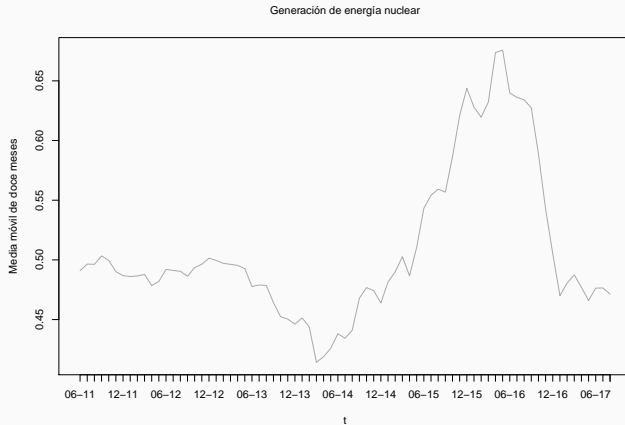
lmfit <- lm(z ~ t)
```



```
##
## Call:
## lm(formula = z ~ t)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.36261 -0.57223  0.00738  0.71255  2.46057
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  4.975067   0.179551   27.71  <2e-16 ***
## t            0.199233   0.003087   64.54  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.891 on 98 degrees of freedom
## Multiple R-squared:  0.977, Adjusted R-squared:  0.9768
## F-statistic:  4166 on 1 and 98 DF, p-value: < 2.2e-16
```

- Proceso no estacionario: raíces dentro del círculo unitario.
- Proceso integrado  $I(d)$ : raíz unitaria en el polinomio autorregresivo.
  - Asociado a una ACF que decae lentamente y toma valores cercanos a la unidad para los primeros rezagos.
  - Prueba de Dickey-Fuller.
  - Se torna estacionario al tomar la  $d$ -ésima diferenciación, donde  $d$  es la cantidad de raíces ubicadas en el círculo unitario.

# Ejemplo

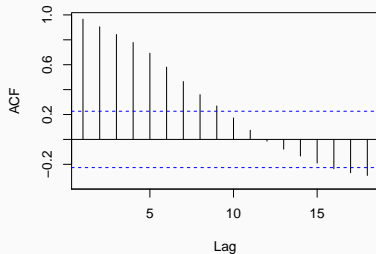


## Discusión en clases

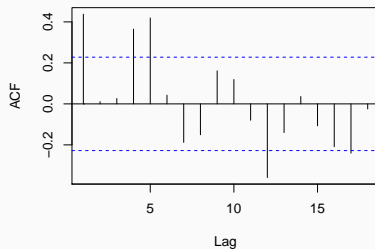
¿Qué tipo de tendencia tiene esta serie? ¿Cómo avanzarían en el análisis?

# Ejemplo (continuación)

**Serie original**



**Primera diferencia**



```
##  
## Augmented Dickey-Fuller Test  
##  
## data: z  
## Dickey-Fuller = -1.6567, Lag order = 1, p-value = 0.7155  
## alternative hypothesis: stationary
```

# Ejercicio: Energías alternativas

## Ejercicio en clases

Analizar la serie de tiempo de ejemplo.

Además de poner práctica nuestras habilidades para analizar datos reales, el ejercicio nos permite aprender algunas funciones de R que son específicas para la manipulación de datos secuenciales. Solución disponible en el anexo.

Algunos pasos:

- Descargar los datos de <https://bit.ly/2IdUjsP>, sección Energías Alternativas.
- Tomando solamente aquellos registros de frecuencia *mensual* y alcance *país*, leer la columna *generacion\_nuclear*.<sup>1</sup>
- Calcular la generación de energía promedio de doce meses centrados.<sup>2</sup>
- Graficar y describir la serie original.
- Probar si tiene raíz unitaria. En tal caso, diferenciar, graficar, y describir la serie.
- Calcular los estadísticos descriptivos.
- Proponer un modelo.
- Evaluar el ajuste.

---

<sup>1</sup> Hay una copia local en `data/nuclear.txt` en caso de que el sitio esté fuera de línea.

<sup>2</sup> La serie original tiene estacionalidad mensual. Al tomar el promedio, eliminamos el componente estacional pero perdemos información valiosa. Más adelante, aprendemos métodos más específicos.

# Procesos no estacionarios en la varianza

---

# Procesos no estacionarios en la varianza

- Un proceso estacionario en la media puede o no ser estacionario en la varianza.
  - Aplicar transformación para estabilizar la varianza.
- Todo proceso no estacionario en la media es no estacionario en la varianza.
  - Aplicar primera diferencia.

Asumamos que la varianza varía junto a la media.

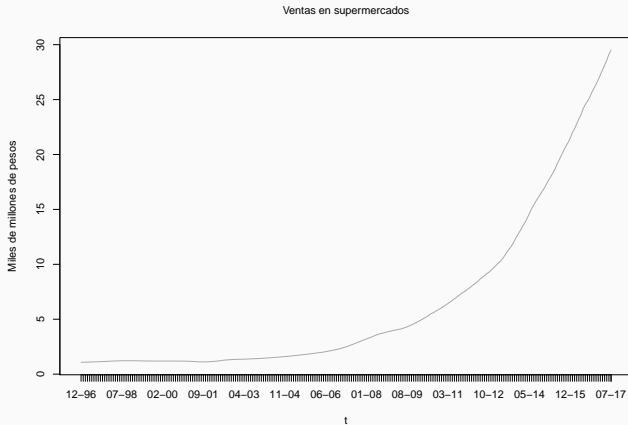
$$V \langle Z_t \rangle = cf(\mu_t)$$

Para lograr una varianza constante, entonces:

- Desvío estándar de una serie crece linealmente con la media: aplicar logaritmo (cualquier base).
- varianza de una serie es proporcional al nivel: aplicar raíz cuadrada.
- Desvío estándar de una serie es proporcional al cuadrado del nivel: aplicar recíproca.



# Ejemplo



## Discusión en clases

A juzgar sólo por el gráfico, ¿qué transformaciones esperan aplicar?

# Ejercicio: Ventas de supermercados

## Ejercicio en clases

Analizar la serie de tiempo de ejemplo.

El Anexo incluye algunos gráficos útiles.

Algunos pasos:

- Descargar los datos desde <https://bit.ly/2GXzXoa>.
- De la Sección A 1.11, leer los datos mensuales para la columna *Ventas totales*.<sup>3</sup>
- Calcular las ventas promedio de doce meses centrados.<sup>4</sup>
- Graficar y describir la serie original.
- Probar si tiene raíz unitaria. En tal caso, diferenciar, graficar, y describir la serie.
- Una vez obtenida una serie estacionaria, proponer una transformación y aplicarla sobre los datos originales.
- Calcular los estadísticos descriptivos.
- Proponer un modelo.
- Evaluar el ajuste.

---

<sup>3</sup> Hay una copia local en `data/INDECSuper.txt` en caso de que el sitio esté fuera de línea.

<sup>4</sup> La serie original tiene estacionalidad mensual. Al tomar la media, eliminamos el componente estacional pero perdemos información valiosa. Más adelante, aprendemos métodos más específicos.

## Anexo: Generación nuclear

---

**Publicado por:** Ministerio de Hacienda. Secretaría de Política Económica. Subsecretaría de Programación Microeconómica.

**Sección:** Indicadores Sectoriales de Energías alternativas.

**Subsección:** Indicadores de Energías alternativas en valores anuales, trimestrales y mensuales.

**Campo:** Generación de energía fuente nuclear.

**Frecuencia:** mensual.

**Agrupado por:** agregado a nivel país.

**Unidad:** MWh (megavatio hora).<sup>5</sup>

---

<sup>5</sup>Transformado en GWh a los fines de la resolución.

# Lectura & procesamiento

```
# https://bit.ly/2IdUjsP Energías Alternativas
```

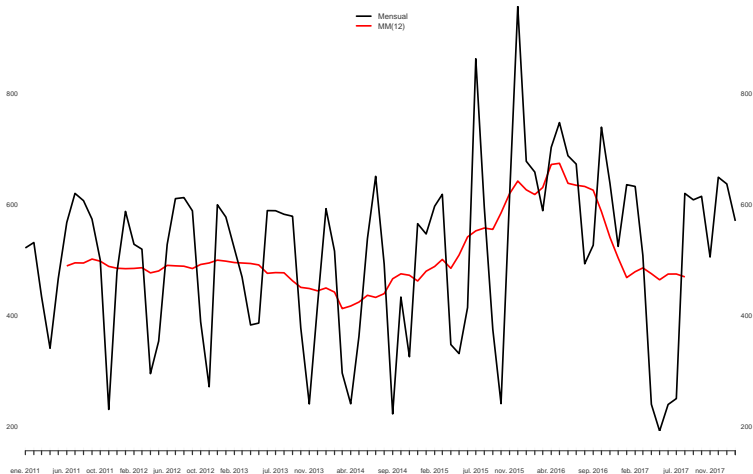
```
df <- read.table(  
  file = "data/nuclear.txt",  
  header = TRUE,  
  sep = "\t"  
)  
  
df[, 1] <- as.POSIXct(df[, 1], format = "%Y-%m-%d")  
df[, 2] <- df[, 2] / 1000  
  
x <- xts(x = df[, 2], order.by = df[, 1])  
xsum <- na.omit(rollmean(x, 12, fill = NA))  
  
t(head(cbind(mes = x, ult12 = xsum), 9))
```

```
##      2011-01-01 2011-02-01 2011-03-01 2011-04-01 2011-05-01 2011-06-01  
## mes      523.425   533.124   435.933   342.776   468.288   569.1960  
## ult12      NA      NA      NA      NA      NA      491.0303  
##      2011-07-01 2011-08-01 2011-09-01  
## mes      621.6130  608.5010  575.2080  
## ult12      496.5097  496.2487  503.3488
```

```
plot(  
  cbind(x, xsum),  
  main = "Generación de energía nuclear en Argentina (GWh)",  
  col = 1:2,  
  grid.col = "white"  
)  
  
addLegend("top", c("Mensual", "MM(12)"), col = 1:2, lwd = 2)
```

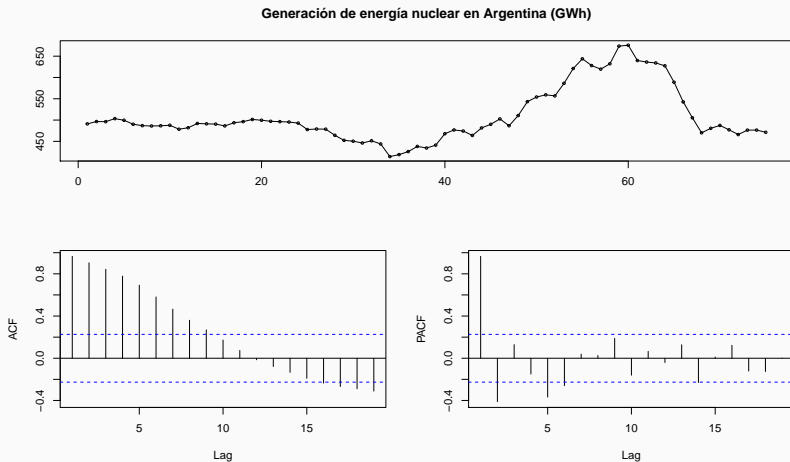
Generación de energía nuclear en Argentina (GWh)

2011-01-01 / 2018-02-01



# Análisis de la serie original original

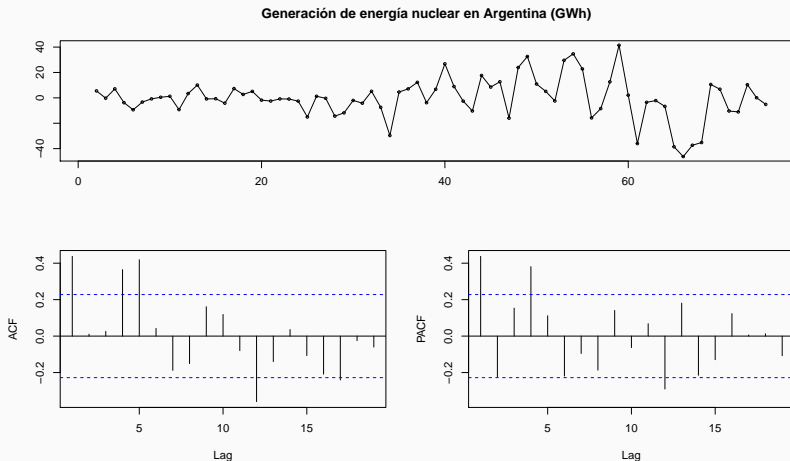
```
tsdisplay(  
  xsum,  
  main = "Generación de energía nuclear en Argentina (GWh)",
```





# Análisis de la serie original diferenciada

```
tsdisplay(  
  diff(xaum),  
  main = "Generación de energía nuclear en Argentina (GWh)",
```



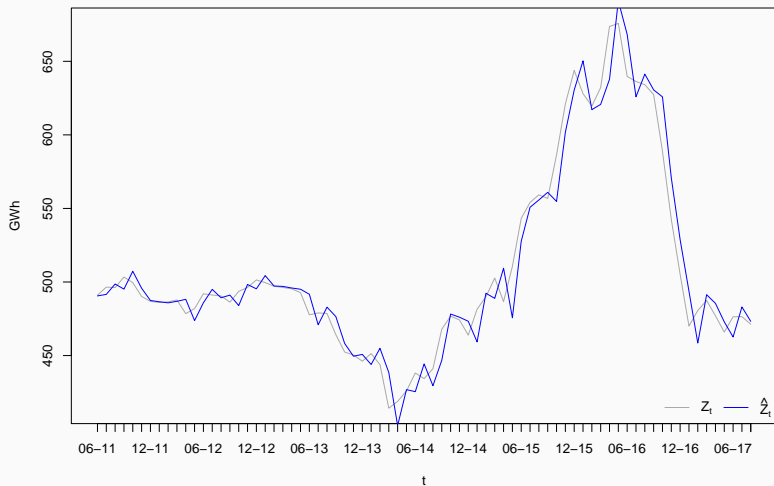
```
fit <- Arima(xsum, order = c(0, 1, 1))
```

```
print(fit)
```

```
## Series: xsum
## ARIMA(0,1,1)
##
## Coefficients:
##          ma1
##          0.4870
## s.e.  0.0938
##
## sigma^2 estimated as 200:  log likelihood=-300.67
## AIC=605.33   AICc=605.5   BIC=609.94
```

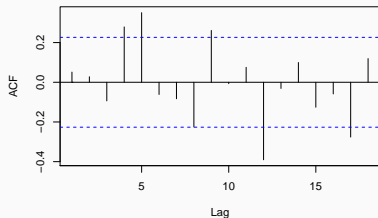
# Ajuste versus observado

Generación de energía nuclear en Argentina

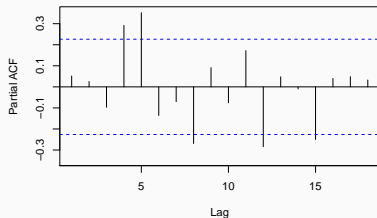


# Diagnóstico de residuos

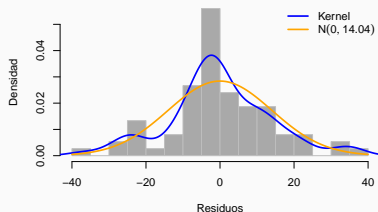
Residuos



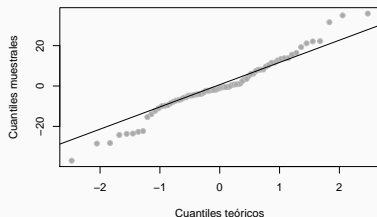
Residuos



Histograma de residuos



QQ Residuos



## **Anexo: Ventas en supermercados**

---

# Lectura & procesamiento

```
# https://bit.ly/2GXzXoa
df <- read.table(
  file = "data//INDECSuper.txt",
  header = TRUE,
  sep = "\t"
)

df[, 1] <- as.POSIXct(df[, 1], format = "%Y-%m-%d")

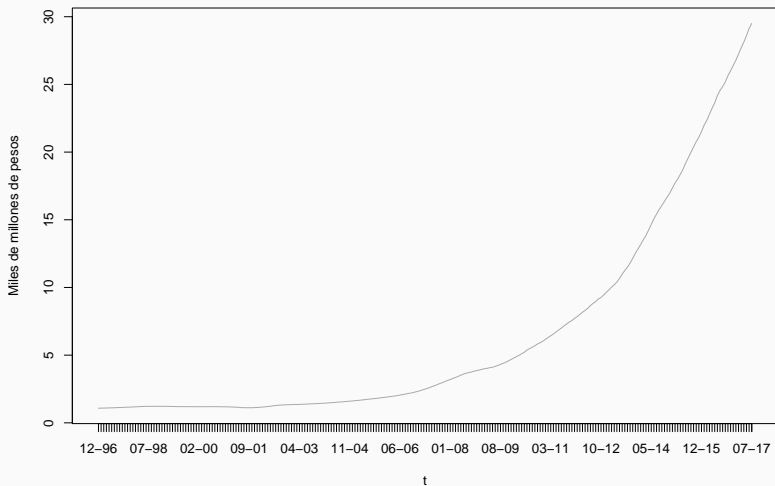
x <- xts(x = df[, 2] / 1000, order.by = df[, 1])
z <- na.omit(rollmean(x, 12, fill = NA))

t(head(cbind(x, z), 9))
```

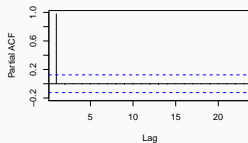
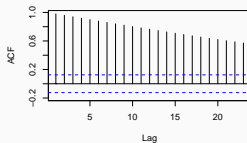
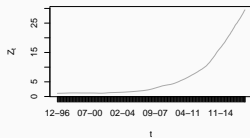
```
##      1996-07-01 1996-08-01 1996-09-01 1996-10-01 1996-11-01 1996-12-01
## ..1      1.036      1.064      0.975      1.025      1.073      1.371000
## ..2      NA      NA      NA      NA      NA      1.079833
##      1997-01-01 1997-02-01 1997-03-01
## ..1      1.013000 1.000000 1.172000
## ..2      1.084417 1.090167 1.095083
```

Ventas en supermercados

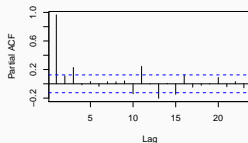
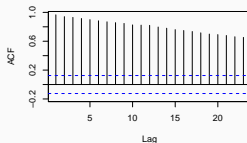
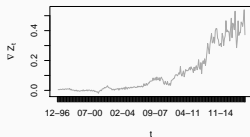


# Serie original

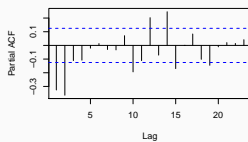
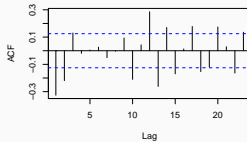
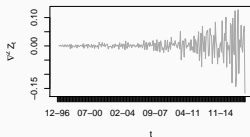
Ventas en supermercados (ARS MM)



Primera diferencia



Segunda diferencia





# Serie transformada (ln)

