

UNIVERSIDAD NACIONAL DE COLOMBIA

PREGRADO EN ESTADISTICA

DEPARTAMENTO DE ESTADÍSTICA

FACULTAD DE CIENCIAS

— SERIES DE TIEMPO UNIVARIADAS —

Profesor:

Norman Diego Giraldo Gómez

Trabajo 1

*Procesos estacionarios, Gaussianos
autocovarianza, pruebas de hipótesis
densidad espectral*

Integrantes:

Ricardo William Salazar Espinal C.C. 1017219472

Alejandro Velásquez Rendón C.C. 1000913570

Medellín, Colombia

Medellín, Enero de 2025

Índice

Índice de Figuras	1
Índice de Tablas	1
1 Enunciado general	2
2 Punto a)	2
2.1 Solución	2
3 Punto b)	9
3.1 Solución	9
4 Punto c)	14
4.1 Solución	14
5 Apéndice-Código usado	18
5.1 Llamado de librerías	18
5.2 Punto a)	18
5.3 Punto b)	20
5.4 Punto c)	21
Referencias	22

Índice de figuras

1	Gráfica de las series-Dafuturo	3
2	Gráfica FAC-Dafuturo	4
3	Gráfica de las series-Popular	6
4	Gráfica FAC-Popular	7
5	Gráfica para elegir orden q-BIC	10
6	Residuos del modelo MA(q)	12
7	Fac de residuos	13
8	Periodograma acumulado de los residuos	14
9	Periodogramas de las series	15
10	Histograma	16
11	Gráfica de Periodogramas acumulados	17

Índice de cuadros

1	Coeficientes del modelo MA(7)	11
---	---	----

1 Enunciado general

Se utilizan datos del valor de la unidad de fondos de pensiones voluntarias administrados por sociedades fiduciarias. Son fondos para ahorro complementario a las pensiones obligatorias. Se consideran los siguientes casos o fondos:

- a) Fiduciaria Davivienda. Fondo VANGUARD US 500 STOCK FUND F.P.V DAFUTURO
- b) Fiduciaria Davivienda. MULTIFONDO BASICO F.V.P DAFUTURO
- c) Fiduciaria Davivienda. FONDO DE PENSIONES VOLUNTARIO DAFUTURO
- d) POPULAR. FONDO DE PENSIONES VOLUNTARIAS MULTIOPCION
- e) ALIANZA FIDUCIARIA S.A. FONDO DE PENSIONES ABIERTO VISION

En nuestro caso se trabajarán con los fondos de los literales “b” y “d”.

Se utilizan las variaciones porcentuales del valor de la unidad, de la pareja de fondos mencionada. $X_{j,n} = \log(U_{j,n}/U_{j,n-1})$, $j = 1, 2$, donde $U_{j,n}$ es el valor de la unidad en el día n .

2 Punto a)

Reporte las gráficas de $X_{j,n}$, $j = 1, 2$, junto con las funciones fac. Aplique la prueba KPSS a $X_{j,n}$. Se acepta la nula de estacionario?

2.1 Solución

Como se había dicho, se trabajarán con los fondos “Fiduciaria Davivienda. MULTIFONDO BASICO F.V.P. DAFUTURO”, que se nombrará de forma simplificada como “Dafuturo”, y con el fondo “POPULAR. FONDO DE PENSIONES VOLUNTARIAS MULTIOPCION”, llamado a partir de ahora como “Popular” para que sea más simple.

Se tomará el valor de la unidad del fondo, denotados por $U_{1,n}$ y $U_{2,n}$, para el fondo Dafuturo y Popular respectivamente, estos datos fueron obtenidos en el año 2015 de forma diaria, por lo que se obtuvieron $n = 1, 2, \dots, 365$ observaciones.

Sin embargo, no se usarán estos valores de forma directa, pues estos valores suelen tender a la alza, en su lugar, para tener valores quizá más centrados a un valor, y para intentar ajustar un modelo, se usarán las variaciones porcentuales diaras del valor de la unidad, esto es, para cada fondo se tiene la siguiente serie de tiempo:

$$X_{j,n} = \log(U_{j,n}/U_{j,n-1}), j = 1, 2$$

Lógicamente, $X_{1,n}$ corresponde a los datos de Dafuturo y $X_{2,n}$ corresponde a los datos de Popular, en este caso, $n = 1, \dots, 364$, pues al ser diferencias de los datos anteriores, se pierde una observación, pues, los valores se desplazan en un día.

2.1.1 Fondo Dafuturo

En primera instancia se hace el análisis con el fondo Dafuturo, en la gráfica 1, a la izquierda se ve la gráfica de $U_{1,n}$, o sea, del valor de la unidad diario, presenta una clara tendencia a la alza, o sea, el fondo tiene rendimientos positivos, hay pequeñas variaciones en la tendencia, es decir, momentos en que la curva se aplana o se empina más según el momento, aunque hay una tendencia general clara.

A la derecha se ve la gráfica de la serie que nos interesa, la gráfica de $X_{1,n}$, que son las variaciones porcentuales, la serie parece fluctuar alrededor de un valor cercano a 0, (es como un valor ligeramente superior a cero), es decir, que las fluctuaciones diarias no suelen ser muy extremas, se ve que oscilan aproximadamente en el intervalo $(-0.0002, 0.0002)$, sin embargo, hay picos bastantes pronunciados que reflejan días de caídas o, en general, movimientos del fondo poco comunes.

De entrada, la serie $X_{1,n}$ podría ser estacionaria alrededor de un valor cercano a cero, tiene una variabilidad constante, pero no se pueden sacar conclusiones determinantes en este punto.

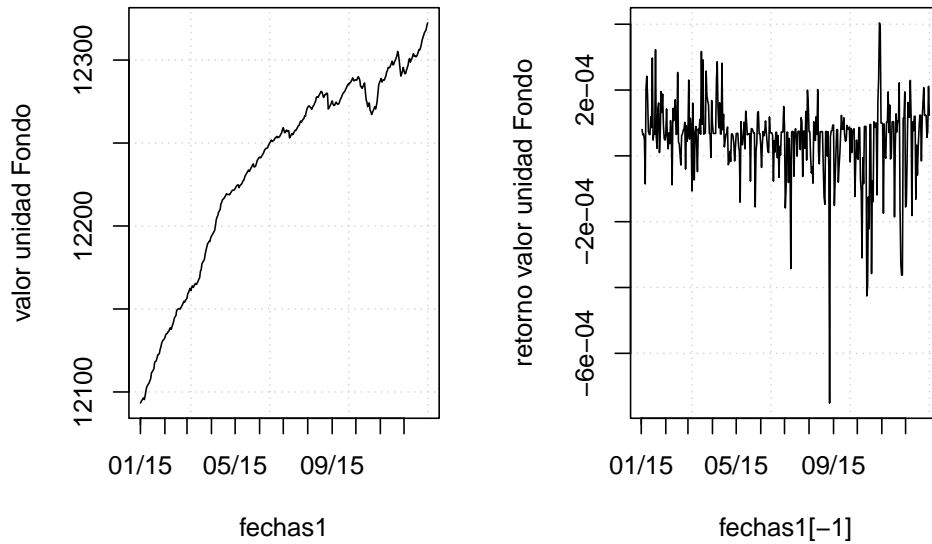


Figura 1: Gráfica de las series-Dafuturo

Ahora, se mostrará la función FAC (función a autocorrelación) de la serie $X_{1,n}$, para esta función, primero definimos la función de autocovarianza entre $X_{1,n}$ y $X_{1,n+k}$ como:

$$R(k) = Cov(X_{1,n}, X_{1,n+k})$$

Con esto, la función de autocorrelación es:

$$\rho(k) = \text{Corr}(X_{1,n}, X_{1,n+k}) = R(k)/R(0)$$

Sin embargo, no se tiene un modelo estocástico, tenemos una serie de tiempo, por tanto se deben estimar las correlaciones. Con los datos, se pueden estimar las correlaciones como sigue:

$$\hat{\rho}(k) = \frac{\sum_{j=1}^{N-k} (X_{1,j} - \bar{X}_1)(X_{1,j+k} - \bar{X}_1)}{\sum_{j=1}^{N-k} (X_{1,j} - \bar{X}_1)^2}$$

Donde N es la longitud de la serie y $\bar{X}_1 = \frac{1}{N} \sum_{j=1}^N X_{1,j}$.

Se decide presentar la gráfica de la función FAC con 100 rezagos, pues a partir de allí la tendencia es muy monótona y muestra solo correlaciones que no son significativas.

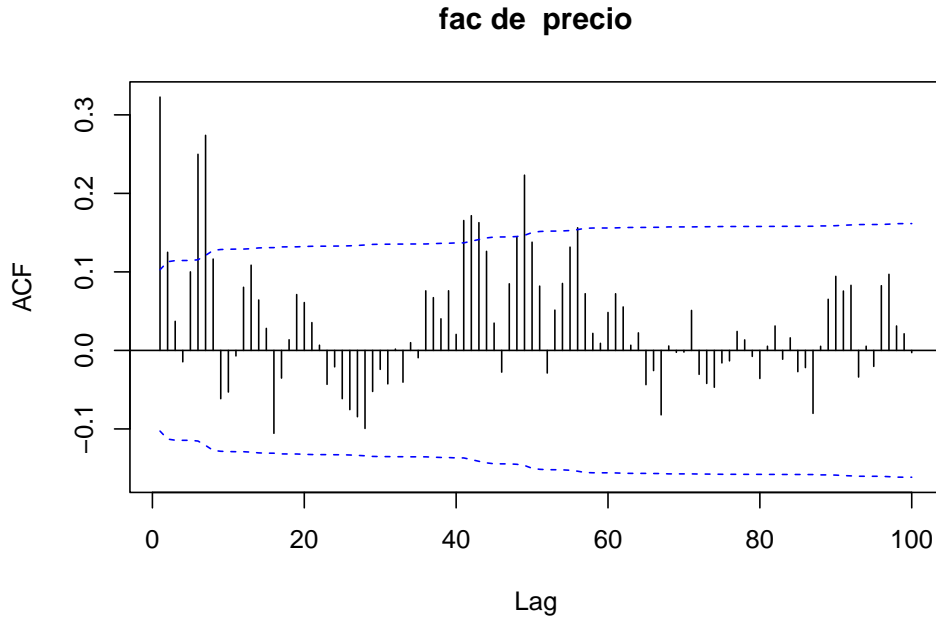


Figura 2: Gráfica FAC-Dafuturo

En todo caso, en la gráfica 2 se puede ver la FAC para $X_{1,n}$, los primeros rezagos muestran autocorrelaciones positivas y significativas, lo que indica que hay patrones de dependencia en el corto plazo, a medida que aumentan los rezagos, las autocorrelaciones empiezan a oscilar en 0 y tienden a dejar de ser significativas. Los picos en lags mayores, y la forma en general de la FAC puede indicar una estacionalidad débil.

Para confirmar si la serie es o no estacionaria en covarianza, se usará la función `kpss.test` de la librería `aTSA`, esta se basa en escribir a nuestra serie como:

$$X_{1,n} = a + bn + Y_n + Z_n$$

Donde a , b son constantes, Y_n es estacionario en covarianza, y $Z_n = Z_{n-1} + \varepsilon_n$, con $\varepsilon \stackrel{iid}{\sim} N(0, \sigma^2)$.

Con esto se propone el siguiente juego de hipótesis:

$$\begin{cases} H_0 : \sigma^2 = 0 \\ H_1 : \sigma^2 \neq 0 \end{cases}$$

No rechazar H_0 es lo mismo que decir que $X_{1,n}$ es estacionario con tendencia. La función `kpss.test` arroja tres valores p que se hacen en base a los valores de a y b en la definición que se hizo de $X_{1,n}$, la salida en cuestión se presenta a continuación:

```
KPSS Unit Root Test
alternative: nonstationary
```

```
Type 1: no drift no trend
```

```
lag stat p.value
  4 10.2    0.01
```

```
-----
```

```
Type 2: with drift no trend
```

```
lag stat p.value
  4 0.852    0.01
```

```
-----
```

```
Type 1: with drift and trend
```

```
lag stat p.value
  4 0.146 0.0496
```

```
-----
```

```
Note: p.value = 0.01 means p.value <= 0.01
      : p.value = 0.10 means p.value >= 0.10
```

De forma estándar, usamos un nivel de significancia para las pruebas de hipótesis de $\alpha = 0.05$, con esto, se rechaza la hipótesis nula, es decir, que la serie no es estacionaria en covarianza, pues el valor p es de 0.01.

2.1.2 Fondo Popular

Igual que se hizo para el fondo Dafuturo, para el fondo popular se mostrarán, en primer lugar, las gráficas de las series $U_{2,n}$ y $X_{2,n}$, estas se presentan en la gráfica 3.

La diferencia más notable entre estas gráficas y las vistas en la figura 1 es que, para el fondo Popular, no hay datos del 2016, solo del 2015, como se había dicho.

Ahora, con respecto a las series como tal, igualmente, el valor de la unidad $U_{2,n}$ tiende a la alza, con periodos de estancamiento, pero la tendencia general es creciente, respecto de la serie de interés, $X_{2,n}$ esta suele fluctuar alrededor de un valor cercano a 0, y oscila entre aproximadamente -0.001 y 0.001, aunque no se observa un patrón temporal claro, podría ser

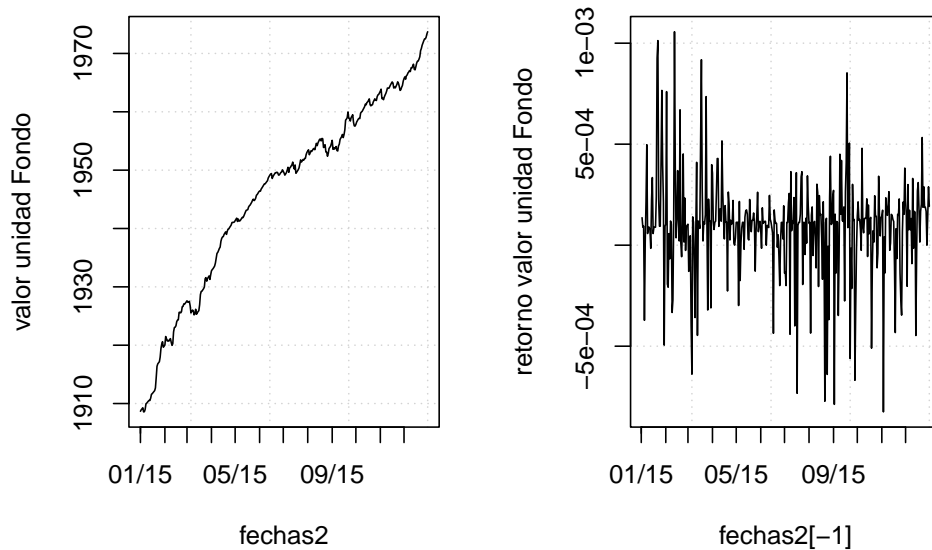


Figura 3: Gráfica de las series-Popular

estacionaria en covarianza, pues se mantiene dentro de un intervalo y oscila alrededor de un valor, pero las caídas y picos elevados hacen dudar de este supuesto.

Por otra parte, en la gráfica 4, se ve la FAC de la serie $X_{2,n}$, no se comentará mucho de ella por ahora, su utilidad vendrá un poco más adelante.

Y, finalmente, al igual que se hizo con la serie $X_{2,n}$, se usará la función `kpss.test` del R para aplicar la prueba de estacionariedad en covarianza, los resultados se ven a continuación:

```

KPSS Unit Root Test
alternative: nonstationary

```

```

Type 1: no drift no trend

```

```

lag stat p.value
 4 9.88    0.01

```

```

-----

```

```

Type 2: with drift no trend

```

```

lag stat p.value
 4 0.433   0.063

```

```

-----

```

```

Type 1: with drift and trend

```

```

lag stat p.value
 4 0.128   0.0833

```

```

-----

```

```

Note: p.value = 0.01 means p.value <= 0.01

```

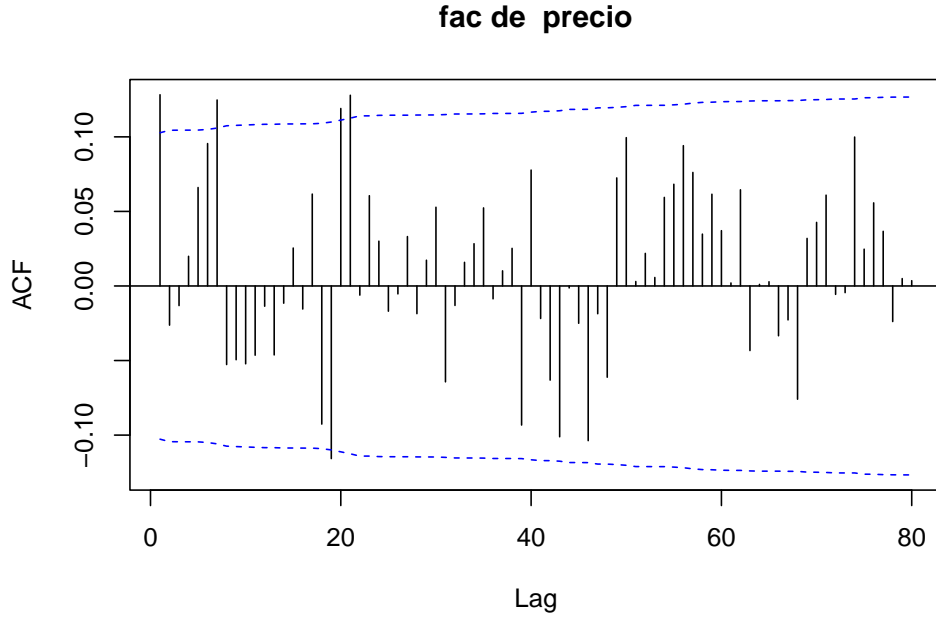



Figura 4: Gráfica FAC-Popular

: p.value = 0.10 means p.value \geq 0.10

Las conclusiones son similares a lo que se dijo sobre la serie de las variaciones porcentuales del fondo Dafuturo, solo que en este caso el segundo valor p parece aceptar que la serie $X_{2,n}$ es estacionaria con tendencia, de forma un poco más fuerte, pues su valor p es mayor.

2.1.3 Filtro LOESS

Como ejercicio extra, dados los valores p que arrojaron los test KPSS, se decide quitar la tendencia de las series $X_{j,n}$, $j = 1, 2$, con un concepto llamado el filtro Loess, no se entrarán en todos los detalles teóricos, sin embargo, se hará una explicación de la aplicación de este método con la serie $X_{1,n}$, y para la serie $X_{2,n}$ será similar.

Como ya se mostró, se puede asumir que $X_{1,n}$ es estacionario con tendencia, entonces se descompone como sigue:

$$X_{1,n} = Y_{1,n} + f_1(n)$$

Donde $f_1(n)$ es la tendencia de la serie, el objetivo es hallar a $\hat{f}_1(n)$, un estimador de $f_1(n)$, una vez obtenido este estimador, se tiene que:

$$Y_{1,n} = X_{1,n} - \hat{f}_1(n)$$

Este nuevo proceso, $Y_{1,n}$ posiblemente sea estacionario en covarianza, a este nuevo proceso se le aplica la prueba KPSS, sus salidas se ven a continuación:

```
KPSS Unit Root Test
alternative: nonstationary
```

```
Type 1: no drift no trend
lag    stat p.value
  4 0.0267    0.1
-----
```

```
Type 2: with drift no trend
lag    stat p.value
  4 0.0279    0.1
-----
```

```
Type 1: with drift and trend
lag    stat p.value
  4 0.0238    0.1
-----
```

```
Note: p.value = 0.01 means p.value <= 0.01
      : p.value = 0.10 means p.value >= 0.10
```

Y, haciendo el mismo proceso con $X_{2,n}$, y aplicando la prueba KPSS se obtiene:

```
KPSS Unit Root Test
alternative: nonstationary
```

```
Type 1: no drift no trend
lag    stat p.value
  4 0.0206    0.1
-----
```

```
Type 2: with drift no trend
lag    stat p.value
  4 0.0208    0.1
-----
```

```
Type 1: with drift and trend
lag    stat p.value
  4 0.0199    0.1
-----
```

```
Note: p.value = 0.01 means p.value <= 0.01
      : p.value = 0.10 means p.value >= 0.10
```

En ambos casos, mejora el valor p, y se concluye que las nuevas series, $Y_{1,n}$ y $Y_{2,n}$ son estacionarias en covarianza.

3 Punto b)

Sólo para uno de los fondos, identifique un posible modelo $MA(q)$ con la función `armasubsets`. Reporte el orden q . Estime el modelo con la función `arima`. Reporte los coeficientes estimados y sus valores p para la prueba de significancia. Reporte la gráfica de los residuos del $MA(q)$, su función `fac`, el periodograma acumulado con la función `cpgram`. Concluya sobre si el modelo cumple con sus especificaciones.

3.1 Solución

Se trabajará con el fondo: POPULAR. FONDO DE PENSIONES VOLUNTARIAS MULTIOPCION (Popular). Para estimar el orden q , usamos dos herramientas, en primer lugar, la gráfica de la función FAC (con las bandas de Bartlett), y la otra está basada en el valor del BIC.

3.1.1 Bandas de Bartlett

Aquí se usa la gráfica 4, se observa que en los rezagos (lags) cercanos a 7, las autocorrelaciones son significativas. Esto podría indicar que un modelo $MA(q)$ con un orden de $q = 7$ puede ser adecuado.

La autocorrelación parece decrecer después del lag 7, pero se siguen presentando algunos rezagos más alejados (por ejemplo, lag 20) con valores significativos. Esto puede ser indicativo de una dependencia periódica o estacionaria.

A manera de conclusión, basados en el análisis de la función de autocorrelación (FAC) y las bandas de Bartlett, un modelo $MA(q=7)$ podría capturar adecuadamente la estructura de los datos, ya que los rezagos significativos se concentran en torno al lag 7, y la mayoría de las autocorrelaciones posteriores caen dentro de las bandas, indicando menor relevancia estadística. Sin embargo, el lag 20 también muestra cierta significancia, lo que podría sugerir un modelo más complejo, por esto, y para una selección más adecuada, se realizará el análisis utilizando la función `armasubsets`, esto para concluir en base al valor BIC.

3.1.2 Análisis del Índice BIC

El criterio de información Bayesiano, conocido como BIC, por sus siglas en inglés, es útil en este tipo de contextos, o sea, en la elección modelos de Medias móviles de orden q ($MA(q)$), es útil porque, en líneas generales, maximiza el ajuste del modelo teórico a los datos que se tienen, y a la vez penaliza la complejidad, para evitar sobreajustes.

En la gráfica 5 se ven las propuestas para el orden q para el modelo de medias móviles, esta basada en el BIC, muestra un valor mínimo del BIC en $q = 7$, lo que sugiere que un modelo $MA(q)$ con este orden es el más adecuado. Sin embargo, también se observa un valor del BIC competitivo en $q=20$, pero como se indicó previamente, y para ser concordantes, el orden seleccionado es el de $q = 7$.

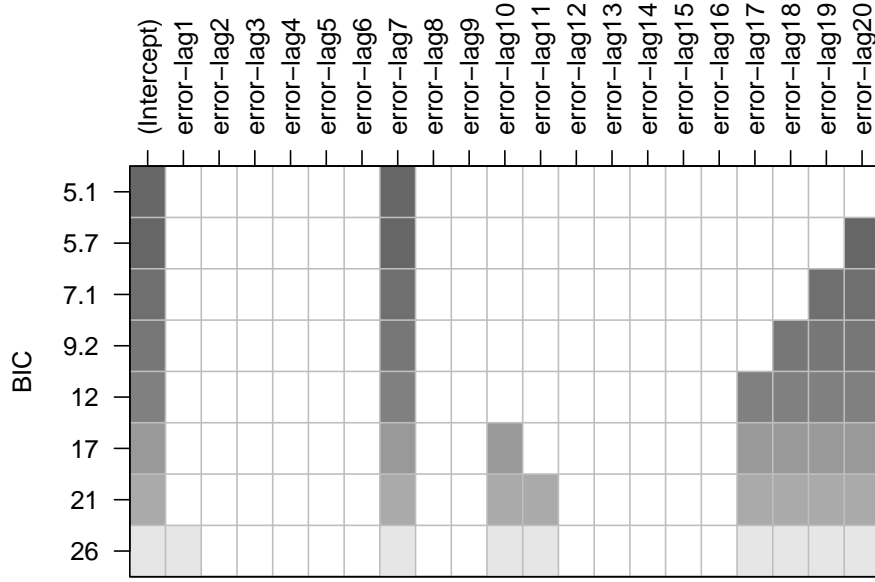


Figura 5: Gráfica para elegir orden q-BIC

3.1.3 Coeficientes y valores p

Queremos entonces, encontrar un modelo de medias móviles para $X_{2,n}$, esto es, pretender que $X_{2,n} \sim MA(7)$, por tanto, se puede expresar al proceso como:

$$X_{2,n} = \mu + \theta_1 Z_{n-1} + \theta_2 Z_{n-2} + \dots + \theta_7 Z_{n-7}$$

Donde $Z_n \sim RBG(0, \sigma^2)$, en este apartado se pretenden calcular los coeficientes θ_i , y ver si son significativos para el modelo.

Para tal fin, se usa la función de R `arma`, al aplicarle `summary`, se obtienen los valores de los coeficientes pero no nos dice si son significativos o no (no arroja los valores p asociados a cada uno), por tanto se deben calcular estos valores como se explica a continuación.

Los valores p se calculan utilizando la distribución t de Student y comparando la estadística t del coeficiente con la distribución teórica. Un valor p pequeño sugiere que el coeficiente estimado es significativamente diferente de cero, lo que implica que tiene un impacto importante en el modelo. Los cálculos se basan en los coeficientes estimados y sus errores estándar, y la hipótesis nula se establece en que los coeficientes son cero (sin efecto).

Para dicha prueba se puede también usar la distribución Z (normal de medio 0 y varianza 1) si la muestra es grande, como es el caso (la función `arma` utiliza por defecto esta última), por tanto se propone el siguiente estadístico:

$$Z_j = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)}$$

Con esto, el valor p se define como:

$$VP = 2 \cdot P(|Z_j| \geq |Z_{observado}|)$$

Si el valor p es pequeño, se podrá rechazar la hipótesis que se había mencionado (que el coeficiente es 0). y concluir que el coeficiente es estadísticamente significativo. Si el valor p es grande, no se puede rechazar la hipótesis nula y concluimos que el coeficiente no tiene un impacto significativo.

En todo caso, los coeficientes y sus respectivos valores p se pueden ver en la tabla 1

Tabla 1: Coeficientes del modelo MA(7)

Coeficiente	Valor	Error.Estándar	Valor.p
ma1	0.1320	0.0522	0.011366
ma2	-0.0205	0.0517	0.692235
ma3	0.0088	0.0523	0.866313
ma4	0.0448	0.0547	0.412511
ma5	0.0679	0.0558	0.223819
ma6	0.0959	0.0488	0.049495
ma7	0.1651	0.0587	0.004905
intercept	0.0001	0.0001	0.098890

Con esto, se propone el siguiente polinomio para el modelo de medias móviles que se ajustará a la serie de tiempo $X_{2,n}$

$$X_{2,n} = \mu + Z_n + \theta_1 Z_{n-1} + \theta_2 Z_{n-2} + \theta_3 Z_{n-3} + \theta_4 Z_{n-4} + \theta_5 Z_{n-5} + \theta_6 Z_{n-6} + \theta_7 Z_{n-7}$$

$$Z_n \sim RBG(0, \sigma^2)$$

Donde $\underline{\theta} = (\theta_1, \theta_2, \dots, \theta_7)'$ son los coeficientes que se presentan en la tabla 1.

Para los coeficientes ma1, ma6 y ma7, valor p < 0.05, lo que implica que estos coeficientes son estadísticamente significativos (para $\alpha = 0.05$). Es decir, esos coeficientes son importantes para el modelo.

Coeficientes no significativos (como ma2, ma3, ma4, ma5): Aunque estos coeficientes no son estadísticamente significativos, el BIC sigue favoreciendo el modelo completo MA(7). Estos coeficientes pueden estar contribuyendo a capturar variabilidad en los residuos o patrones a largo plazo que no se reflejan directamente en la significancia individual.

3.1.4 Análisis de Residuales

Como se dijo, para establecer un modelo de medias móviles, se supone en principio que $Z_n \sim RBG(0, \sigma^2)$, en la figura 7, se presenta la gráfica de los residuales del modelo que se ha ajustado.

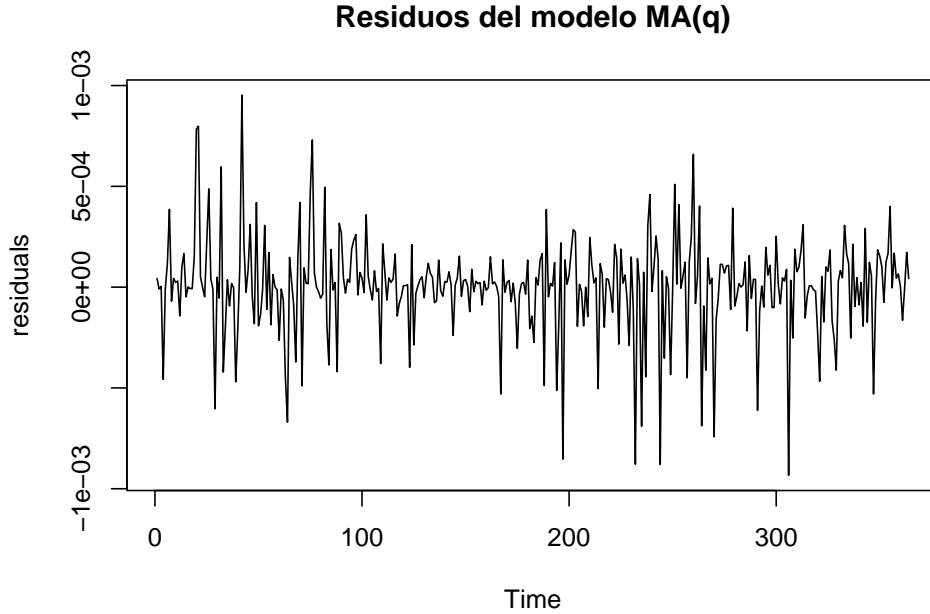


Figura 6: Residuos del modelo MA(q)

Los residuales parecen oscilar alrededor de cero, lo cual es deseable en un modelo bien ajustado. Esto indica que no hay un sesgo sistemático que el modelo no haya capturado, la dispersión de los puntos parece relativamente uniforme a lo largo del tiempo, lo que sugiere que no hay heterocedasticidad (es decir, varianza no constante).

Aunque no podemos determinar la independencia de los residuos solo con esta gráfica, la falta de patrones claros es una señal positiva.

Por otra parte, en la gráfica 7, se ve la FAC de los residuales del modelo

La apariencia general de la FAC es consistente con la de ruido blanco, lo que indica que los residuos son independientes. Esto es una buena señal de que el modelo ha capturado correctamente la estructura de la serie temporal.

Con base en esta gráfica, el modelo MA(7) parece ser adecuado, ya que los residuos no muestran un patrón de autocorrelación significativo.

En la gráfica 8 podemos notar que el periodograma acumulado está entre las líneas de referencia esto sugiere que los residuos tienen un comportamiento similar al de un ruido blanco, es decir, no hay frecuencias dominantes significativas en los datos. Indica que el modelo ha capturado bien la estructura de la serie y que no quedan patrones cíclicos significativos sin

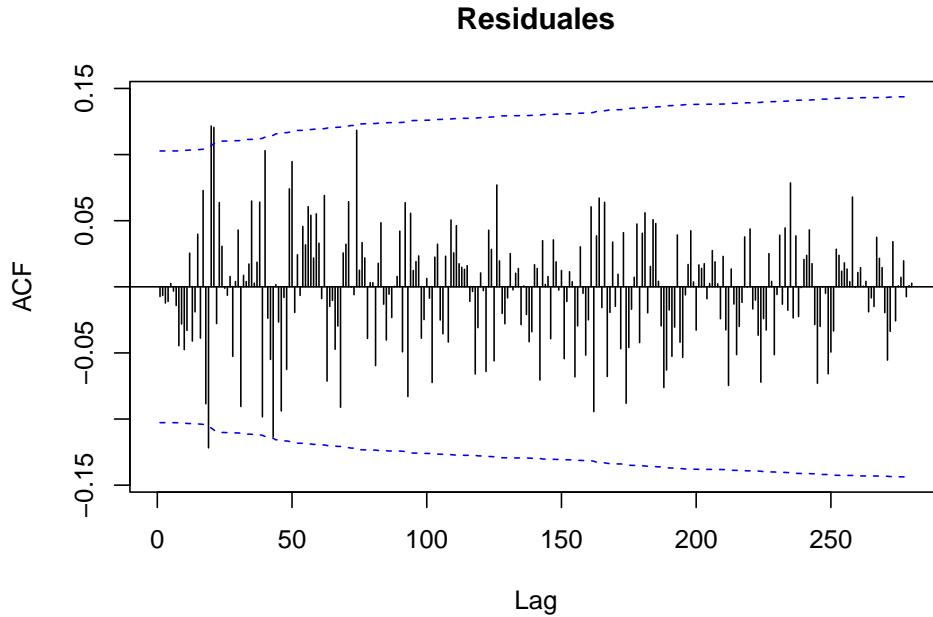


Figura 7: Fac de residuos

modelar, por lo tanto, se puede concluir que el modelo MA(7) es apropiado y que no hay evidencia de patrones cíclicos o frecuencias dominantes en los residuos.

Ya se ha comentado sobre los residuales del modelo, aparentemente se comportan de la forma deseado, como un ruido blanco, sin embargo, se hará una prueba de hipótesis para confirmar esto.

La prueba en cuestión es la de Ljung-Box, se aplicará sobre el modelo Ma(7), así se podrá ver si sus residuos son ruido blanco o no. Para ello se propone el siguiente juego de hipótesis:

$$\begin{cases} H_0 : Z_n \sim RB(0, \sigma^2) \\ H_1 : no(H_0) \end{cases}$$

La salida de R que permite concluir sobre esta prueba de hipótesis se muestra a continuación:

Box-Ljung test

```
data: residuals
X-squared = 0.21682, df = 7, p-value = 1
```

Se obtiene un valor p de 1, es decir, con una significancia de $\alpha = 0.05$, no hay evidencia suficiente como para rechazar la hipótesis nula, lo que implica que para el modelo, se cumple la condición de que sus residuos son ruido blanco.

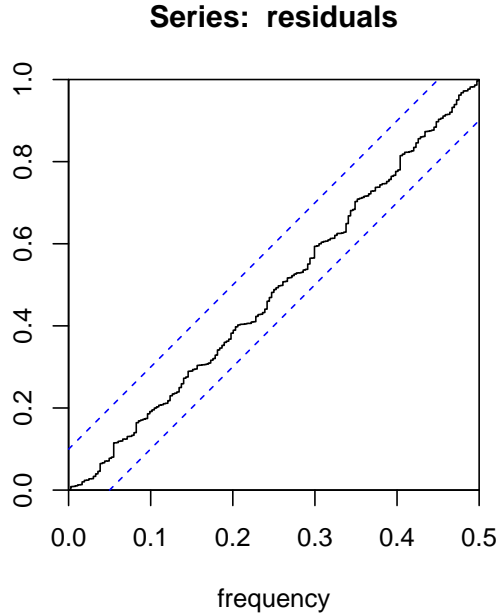


Figura 8: Periodograma acumulado de los residuos

3.1.5 Conclusión

Ajustar un modelo de medias móviles de orden $q = 7$, es decir, un $MA(7)$, para la serie de las variaciones porcentuales del fondo Popular, es decir, para la serie $X_{2,n}$ es adecuado, esto basado en la evidencia gráfica que se presentó, y a las pruebas formales que también se hicieron. El modelo cumple las especificaciones requeridas.

4 Punto c)

Efectúe una prueba de homogeneidad aleatorizada para la densidad espectral suavizada para el estadístico K-S, para probar la nula de que las densidades espectrales de $X_{1,n}$ y $X_{2,n}$ son idénticas. Reporte el histograma y el valor p muestral. Concluya. Qué implicación tendría su conclusión para un usuario del ahorro voluntario?

4.1 Solución

Para la realización de la prueba, es importante que las series sean estacionarias en covarianza, en el literal a), se vió que ambas series eran apenas estacionarias con tendencia, pues para $X_{1,n}$ y $X_{2,n}$, los valores p de la prueba KPSS eran cercanos a 0.05, por tanto, se hará la prueba de comparación de las densidades espectrales aplicando el filtro Loess, es decir, que se van a comparar las series que se denominaron como $Y_{1,n}$ y $Y_{2,n}$, evitando de esta forma cualquier tipo de tendencia que pueda afectar la prueba, y mejorando la condición de ser estacionarias.

Se propone el siguiente juego de hipótesis:

$$\begin{cases} H_0 : I_{Y_{1,n}}(w_k) = I_{Y_{2,n}}(w_k), k = 1, \dots, \lfloor \frac{n}{2} \rfloor \\ H_1 : no(H_0) \end{cases}$$

Para tomar una decisión, se utiliza el estadístico de prueba de Kolgomorov-Smirnoff (K-S) para igualdad de distribuciones, donde $D_n : \max_{k=1, \dots, \lfloor \frac{n}{2} \rfloor} |I_{Y_{1,n}}(w_k) - I_{Y_{2,n}}(w_k)|$. Nótese que en el estadístico original se evalúa en las funciones de distribución acumuladas empíricas, mientras que en este caso son reemplazadas por los periodogramas acumulados que se esbozan en la figura 11.

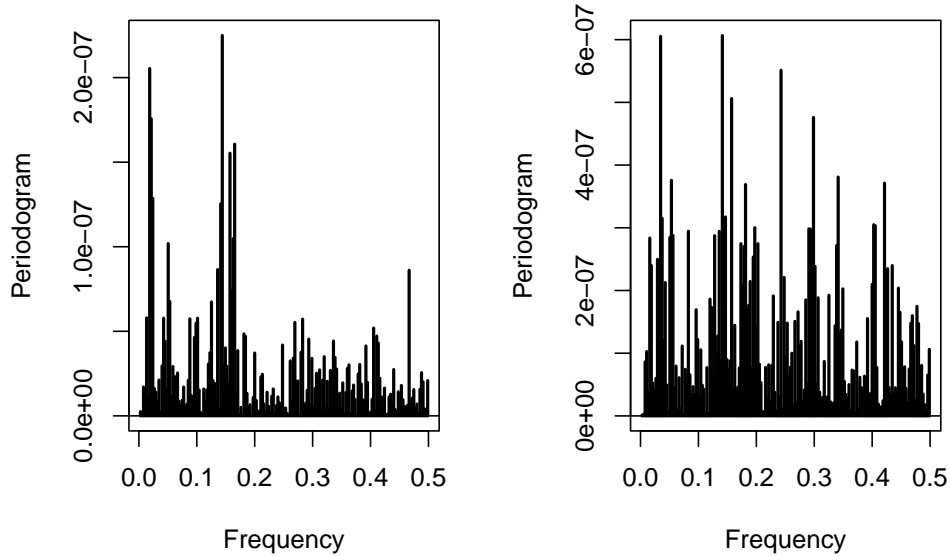


Figura 9: Periodogramas de las series

El valor p de la prueba se imprime a continuación:

[1] 0.02223

El valor p obtenido en la prueba fue de 0.02, lo que indica que se puede rechazar la hipótesis nula a un nivel de significancia del 0.05.

El histograma de las series también fue examinado y presentado en la figura 10, y los resultados sugieren que las densidades espectrales de ambos fondos son diferentes, esto pues se observa que las curvas no se superponen completamente, lo cual indica diferencias en la dinámica de las series. lo que respalda la conclusión de que existen diferencias significativas entre ellos en términos de sus características espectrales.

Dado que se encontraron diferencias significativas en las densidades espectrales de los rendimientos de los dos fondos, podemos concluir que, en términos de sus rendimientos, ambos

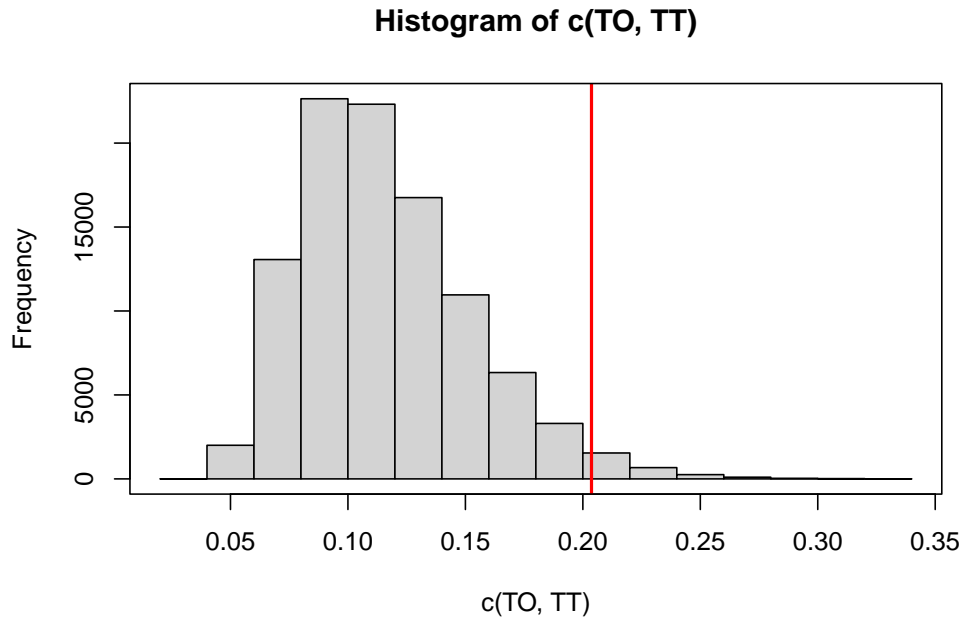


Figura 10: Histograma

fondos presentan comportamientos estadísticamente diferentes. Esto sugiere que, para un usuario interesado en optimizar su estrategia de ahorro voluntario, la elección de uno u otro fondo podría tener implicaciones importantes, ya que los rendimientos de cada fondo varían de manera distinta.

Este hallazgo, basado en una prueba formal y el análisis de su periodograma acumulado, indica que no se puede asumir que los dos fondos sean intercambiables desde el punto de vista de los rendimientos. Sin embargo, es importante considerar que la decisión final sobre cuál fondo elegir debe ir más allá de los aspectos estadísticos de los rendimientos y tomar en cuenta otros factores clave, como la preferencia por el riesgo, los costos de administración y otras características específicas de los fondos (por ejemplo, liquidez, historial de rendimiento, etc.).

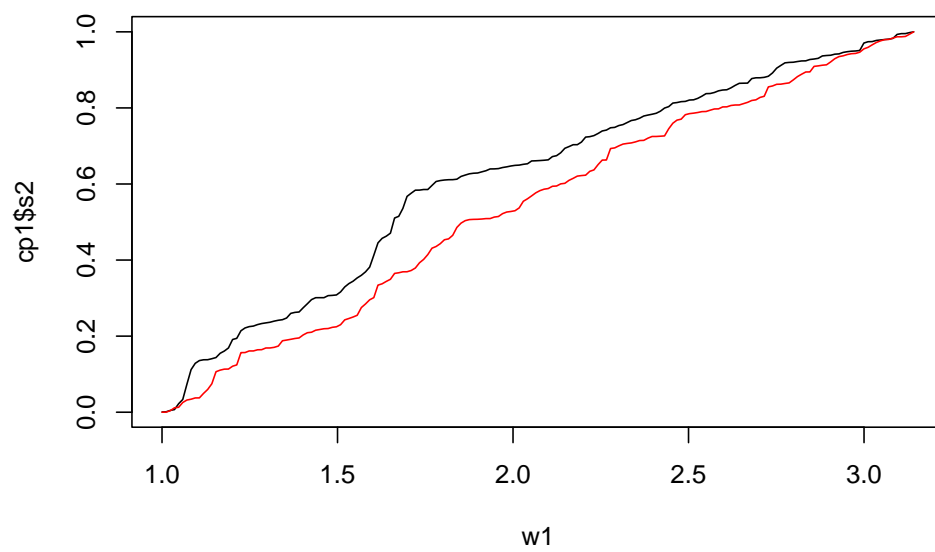


Figura 11: Gráfica de Periodogramas acumulados

5 Apéndice-Código usado

5.1 Llamado de librerías

```
library(descomponer)
library(readxl)
library(tseries)
library(aTSA)
library(TSA)
library(forecast)
library(kableExtra)
library(lmtest)
```

5.2 Punto a)

```
#Lectura bases de Datos
Dafuturo<-read_xlsx("Dafuturo.xlsx")
Popular<-read_xlsx("Popular.xlsx")

#Grafica de la serie 1
y1<-Dafuturo$`Valor de la unidad`
x1<-diff(log(y1),1,1)
fechas1=as.Date(Dafuturo$FECHA)

np = length(y1)
ejex.mes = seq(fechas1[1],fechas1[np], "months")
ejex.year = seq(fechas1[1],fechas1[np], "years")

par(mfrow=c(1,2))

#Gráfica de los valores originales
plot(fechas1,y1, xaxt="n", panel.first = grid(),
     type='l',ylab='valor unidad Fondo')
axis.Date(1, at=ejex.mes, format="%m/%y")
axis.Date(1, at=ejex.year, labels = FALSE, tcl = -0.2)

#Gráfica de las diferencias logarítmicas
plot(fechas1[-1],x1, xaxt="n", panel.first = grid(),
     type='l',ylab='retorno valor unidad Fondo')
axis.Date(1, at=ejex.mes, format="%m/%y")
axis.Date(1, at=ejex.year, labels = FALSE, tcl = -0.2)

#Gráfica FAC Dafuturo
TSA::acf(x1, 310, ci.type='ma', drop.lag.0=TRUE,
```

```

main="fac de precio")

#test Dafuturo
kpss.test(x1)

#fondo 2 Popular
y2<-Popular$`Valor de la unidad`
x2<-diff(log(y2),1,1)
fechas2=as.Date(Popular$FECHA)

np = length(y2)
ejex.mes = seq(fechas2[1],fechas2[np], "months")
ejex.year = seq(fechas2[1],fechas2[np], "years")

#Valor original
par(mfrow=c(1,2))
plot(fechas2,y2, xaxt="n", panel.first = grid(),
     type='l',ylab='valor unidad Fondo')
axis.Date(1, at=ejex.mes, format="%m/%y")
axis.Date(1, at=ejex.year, labels = FALSE, tcl = -0.2)

# Variaciones porcentuales
plot(fechas2[-1],x2, xaxt="n", panel.first = grid(),
     type='l',ylab='retorno valor unidad Fondo')
axis.Date(1, at=ejex.mes, format="%m/%y")
axis.Date(1, at=ejex.year, labels = FALSE, tcl = -0.2)

# Gráfica FAC fondo Popular
TSA::acf(x2, 80, ci.type='ma', drop.lag.0=TRUE,
main="fac de precio")

# test Popular
kpss.test(x2)

# Filtro Loess-Dafuturo
N1 = length(x1)
t1 = seq(1, N1)
mloess = loess(x1 ~ t1, span = 0.75,
               control = loess.control(surface = "direct"))
f1 = mloess$fitted

x1.l = x1 - f1

kpss.test(x1.l)

```

```

#Filtro Loess-Popular
N2 = length(x2)
t2 = seq(1, N2)
mloess = loess(x2 ~ t2, span = 0.75,
               control = loess.control(surface = "direct"))
f2 = mloess$fitted

x2.1 = x2 - f2

kpss.test(x2.1)

```

5.3 Punto b)

```

#posible orden para MA(q) de Popular
res <- armasubsets(x2, nar = 0, nma = 20,
                  y.name = "Variaciones logarítmicas", ar.method = "ols")

plot(res)

# Ajustamos el modelo MA con el orden q
q=7
ma.7 = arima(x2, order = c(0, 0, q))

summary(ma.7)
coeftest(ma.7)

# Gráfica de residuos del modelo MA
residuals = ma.7$residuals
ts.plot(residuals, main = "Residuos del modelo MA(q)")

# Función fac de los residuos
TSA::acf(residuals, 280, ci.type = 'ma', drop.lag.0 = TRUE,
         main="Residuales")

# Periodograma de los residuos
stats::cpgram(residuals)

# Prueba de Ljung-Box para ruido blanco
Box.test(residuals, lag = 7, type = "Ljung-Box")

```

5.4 Punto c)

```
# Gráfica de los periodogramas acumulados
par(mfrow=c(1,2))
Iwk1 = periodogram(x1.l,plot=TRUE,log="no")$spec
abline(h=0)
Iwk2 = periodogram(x2.l,plot=TRUE,log="no")$spec
abline(h=0)

ecdf_1 <- cumsum(Iwk1)/sum(Iwk1)
ecdf_2 <- cumsum(Iwk2)/sum(Iwk2)

w1 = seq(0,0.5,length.out=length(Iwk1))
w2 = seq(0,0.5,length.out=length(Iwk2))

Z <- c(Iwk1,Iwk2)
n <- length(Iwk1)
m <- length(Iwk2)
N <- length(Z)

# Test
TS <- function(Iwk1,Iwk2){
  ecdf_1 <- cumsum(Iwk1)/sum(Iwk1)
  ecdf_2 <- cumsum(Iwk2)/sum(Iwk2)
  abs_dif <- abs(ecdf_1- ecdf_2)
  distancia_ks <- max(abs_dif)
  return(distancia_ks)
}
T0 <- TS(Iwk1,Iwk2)

# Vector del estadístico de cada permutación
TT <- vector()
K = 100000
# Test con permutaciones
for(i in 1:K){
  set.seed(i)
  Z.pi <- sample(Z, N, replace = FALSE)
  TT[i] <- TS(Z.pi[1:n], Z.pi[(n+1):(n+m)])
}

mean(TT>T0) #Valor p

# Histograma
hist(c(T0,TT))
```

```

abline(v = T0, lwd = 2, col = "red")
box()

# Densidades espectrales
cp1=td(x1.1)
cp2=td(x2.1)

w1 = seq(1,pi,length.out=length(cp1$s2))
w2 = seq(1,pi,length.out=length(cp2$s2))

plot(w1,cp1$s2,type='l')
lines(w2,cp2$s2,col='red')

```

Referencias

Gomez, N.D.G. *Introduccion a Series de Tiempo con aplicaciones en R*. Universidad Nacional de Colombia.