**Base monetaria**

**–**

**saldos y promedios**

**(miles de millones de pesos)**

Miriam Quero Gramunt

4rto Estadística

5 enero 2018

Índice

[1. Introducción 3](#_Toc502927904)

[2. Premisas previas 4](#_Toc502927905)

[3. Selección del mejor modelo 7](#_Toc502927906)

[4. Validación del modelo 8](#_Toc502927907)

[5. Predicciones 8](#_Toc502927908)

[6. Avaluación de les predicciones 10](#_Toc502927909)

[7. Comparación de modelos 11](#_Toc502927910)

[8. Conclusión 11](#_Toc502927911)

**Análisis estocástico**

## 1. Introducción

En este primer apartado se realiza un pequeño descriptivo de los datos para situar al lector sin la necesidad de tener que leer la parte anterior de análisis clásico.

Los datos tratados anteriormente y con los que se va a trabajar a continuación es la serie temporal de cifras generadas en base al balance del Banco Central de Chile (en miles de millones de pesos);

Esta base monetaria incluye billetes, monedas y cheques emitidos por el Banco Central, en poder del público, y reservas monetarias de bancos y sociedades financieras; sólo han sido seleccionados los datos de enero de 2007 a diciembre de 2012.

se obtiene en el siguiente link:

(<http://si3.bcentral.cl/Siete/secure/cuadros/arboles.aspx>)

La frecuencia de recogida de datos es mensual; se tienen datos desde enero del año 2007 hasta septiembre de 2017 (pero se decide coger hasta diciembre de 2012), por lo tanto, el número de datos total es de 72 (el periodo extra-muestral serán las últimas 12 observaciones, es decir, las correspondientes al año 2012).

Se pretende averiguar si con el análisis estocástico con los métodos correspondientes a ese tipo de serie hace mejores predicciones que el análisis clásico.

La descriptiva de las cifras mensuales en base al balance del Banco Central de Chile es la siguiente:

Min. :2.954

1st Qu.:3.681

Median :4.564

Mean :4.698

3rd Qu.:5.453

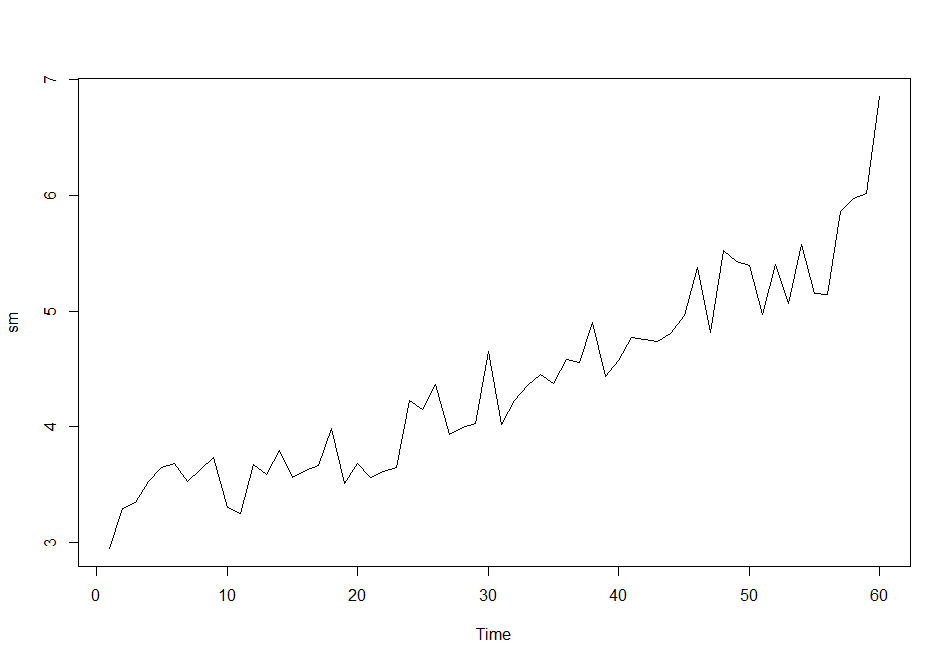
Max. :7.891

Se puede ver que el mínimo de millones de pesos mensuales es de 2.954 y el máximo es de 7.891. Por otro lado, el valor medio de millones de pesos mensuales es de 4.698.

Se recuerda que la serie a tratar es de tipo III, es decir, con tendencia y sin componente estacional.

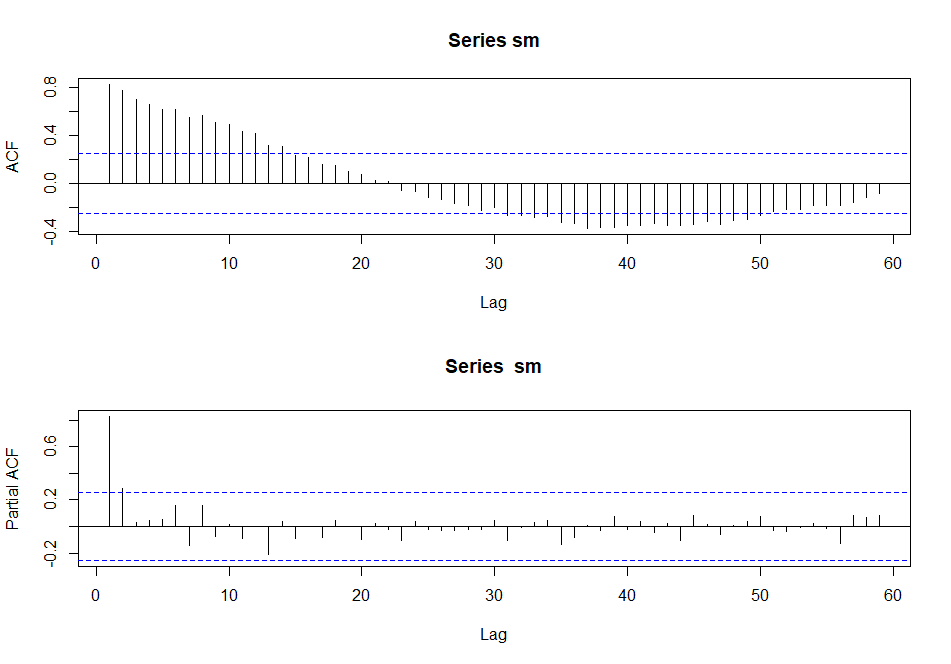
## 2. Premisas previas

Se presenta a continuación el gráfico de la serie temporal para el periodo muestral:



Como se puede ver el gráfico no es horizontal, es decir, muestra tendencia. Se trabajará a continuación para intentar solucionar este problema.

Se muestran, también, las funciones de autocorrelación y de autocorrelación parcial para tener una primera idea del comportamiento de la serie.



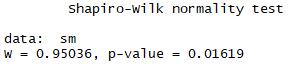
Se puede ver que la función de autocorrelación parece seguir algún tipo de patrón (algún tipo de estacionalidad aunque se ha comprobado que la serie no tiene componente estacional), este hecho supone un problema, ya que la serie no debería presentar ningún tipo de patrón de comportamiento; en cuanto a la función de autocorrelación parcial no tiende de forma rápida al 0, que es lo que debería suceder.

Otra de las premisas es que los datos sigan normalidad; se realiza el test de *Shapiro Wilk* con las hipótesis:

H0: Normalidad

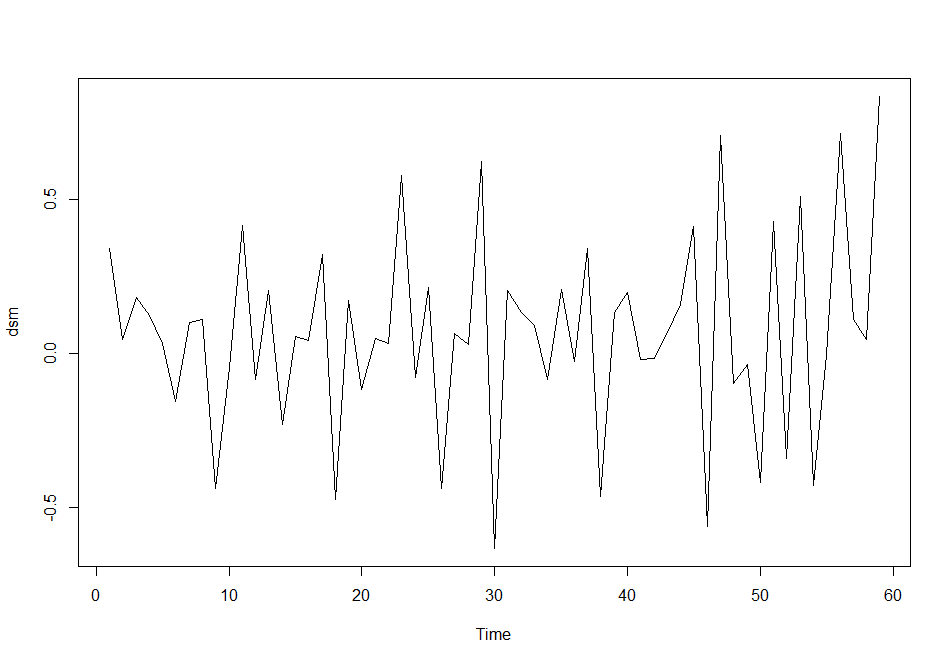
H1: No normalidad

Se obtiene el siguiente resultado:



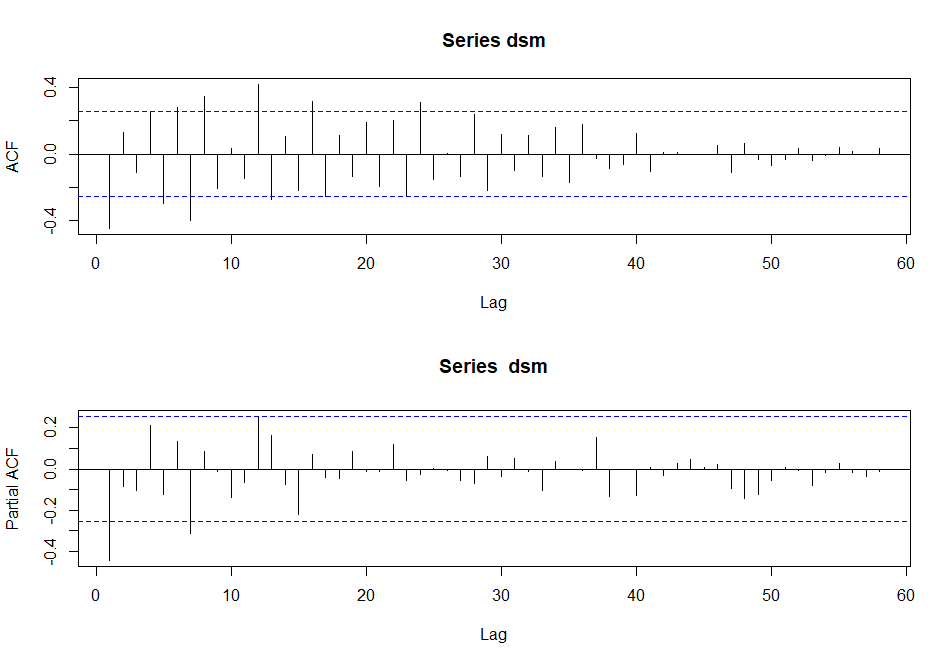
Según el test los datos no son normales (se puede rechazar H0 ya que p-valor es inferior a 0.05).

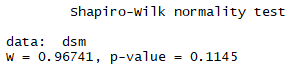
Después de este primer análisis, se decide diferenciar con el objetivo de solucionar algunos de los problemas detectados. Se decide realizar diferencias sin periodicidad, ya que la serie con la se trabaja, como ya se ha dicho no tiene componente estacional.



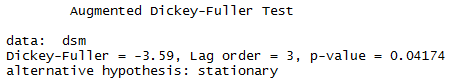
Como se puede ver el comportamiento de la serie después de diferencias es horizontal, ya no parece presentar tendencia.

En cuanto a la funciones de autocorrelación y de autocorrelación parcial se puede ver que la función de autocorrelación ya no parece seguir ningún tipo de patrón, y la función de autocorrelación parcial tiende rápidamente a 0.



Si se comprueba la normalidad de los datos diferenciados se observa que los datos se pueden considerar normales, (p-valor >0.05 -> se acepta H0: normalidad de los datos).  


Se comprueba la estacionariedad:

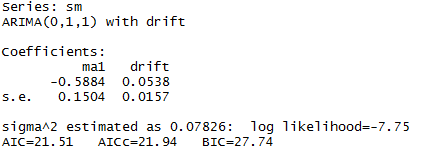


Como el p-valor es inferior a 0.05, se cumple la hipótesis alternativa de estacionariedad, por lo tanto, se puede afirmar que no hay ni tendencia ni estacionalidad. En este caso, se puede proceder con la selección del modelo.

## 3. Selección del mejor modelo

A primera vista, según los gráficos de autocorrelación y autocorrelación parcial parece que un modelo que podría funcionar bien sería un modelo con una componente AR y una componente MA (en el grafico de autocorrelación salen unas líneas significativas (AR(8)) y en el gráfico de autocorrelación parcial (MA(2))).

Según la función de R *auto.arima* del paquete forecast, el mejor modelo sería un ARIMA(0,1,1)



De todos modos, se realizan varias pruebas con diferentes modelos para buscar aquel que se adapte mejor a la serie temporal seleccionada.

Los datos ya han sido diferenciados una vez, con lo cual se probaran modelos en R del tipo ARIMA (X,0,Y) para no diferenciar de nuevo, se muestran algunos de los resultados en la siguiente tabla :

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Modelo | ARIMA en R | Significativo | AIC | Log Likelihood |
| ~~M1~~ | **1,0,1** | **No** | ~~21.09~~ | ~~-7.54~~ |
| M2 | **1,0,2** | **Sí** | **15.51** | **-3.75** |
| ~~M3~~ | 2,0,1 | No | ~~21.91~~ | ~~-6.96~~ |
| M4 | 0,0,1 | Sí | 19.51 | -7.75 |
| ~~M5~~ | 0,0,2 | No | ~~21.22~~ | ~~-7.61~~ |
| M6 | 1,0,0 | Sí | 20.61 | -8.3 |
| ~~M7~~ | 2,0,0 | No | ~~21.64~~ | ~~-7.82~~ |
|  |  |  |  |  |

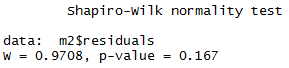
Aquellos modelos que no aparecen como significativos se descartan directamente; aparecen subrayados los dos AIC’s más bajos y los dos ajustes (Log Likelyhood) más altos.

La función *auto.arima* aconseja utilizar un ARIMA(0,1,1) con AIC=19.51, es decir, el modelo 4, pero se decide escoger el modelo 2: ARIMA(1,1,2) ya que su AIC=15.51 es menor que el del otro modelo y a la vez, su Log Likelyhood es mayor con un valor de -3.75. Recordando que se ha diferenciado anteriormente.

## 4. Validación del modelo

A continuación, se valida el modelo. Las premisas necesarias son que los residuos sigan una distribución normal y que sean independientes; Se utiliza R:

Para la normalidad se utiliza el test de Shapiro-Wilk. El p-valor es mayor que 0.05, por lo tanto, se puede aceptar la hipótesis de normalidad (H0).

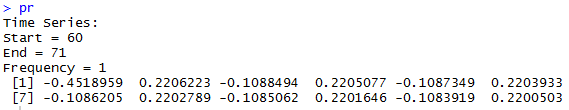


En cuanto a la independencia de los residuos, se utiliza el test de Box-Pierce; el p-valor es mayor que 0.05 por lo tanto se acepta la hipótesis de que los residuos son independientes (H0).

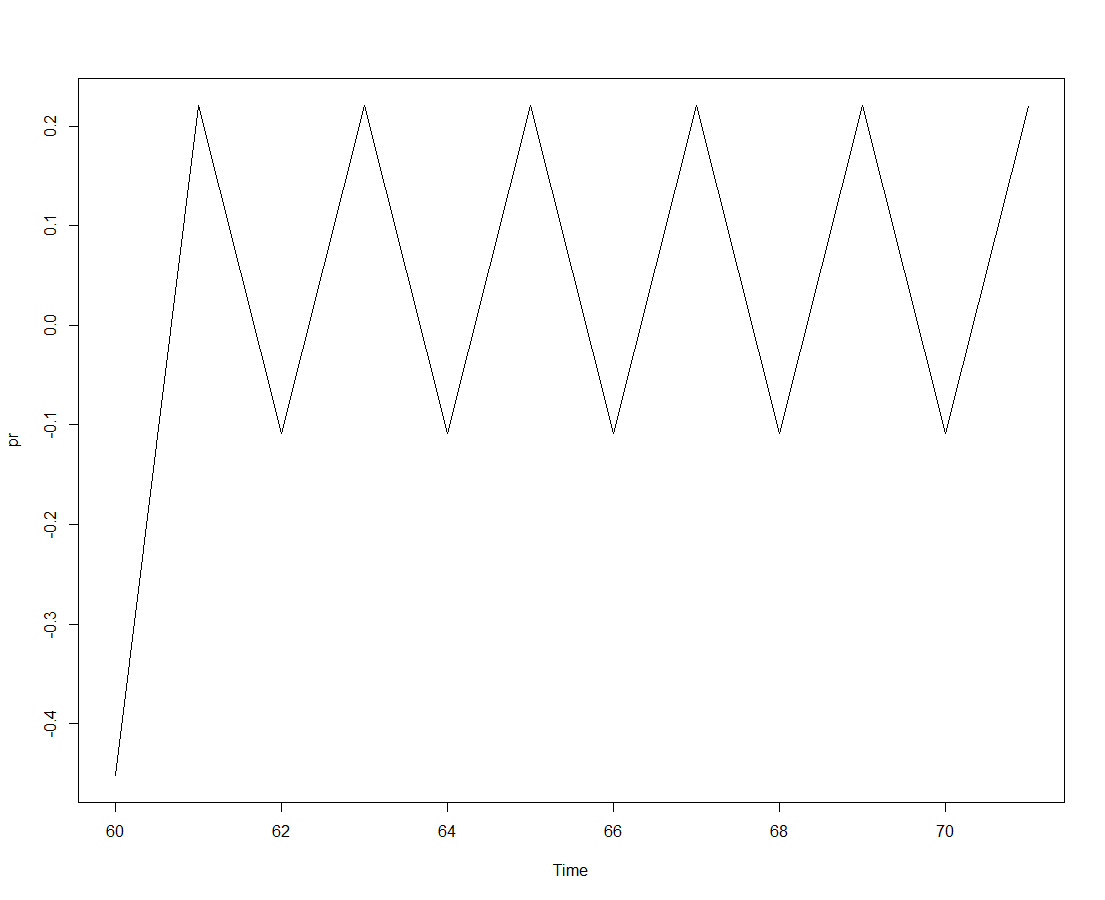


## 5. Predicciones

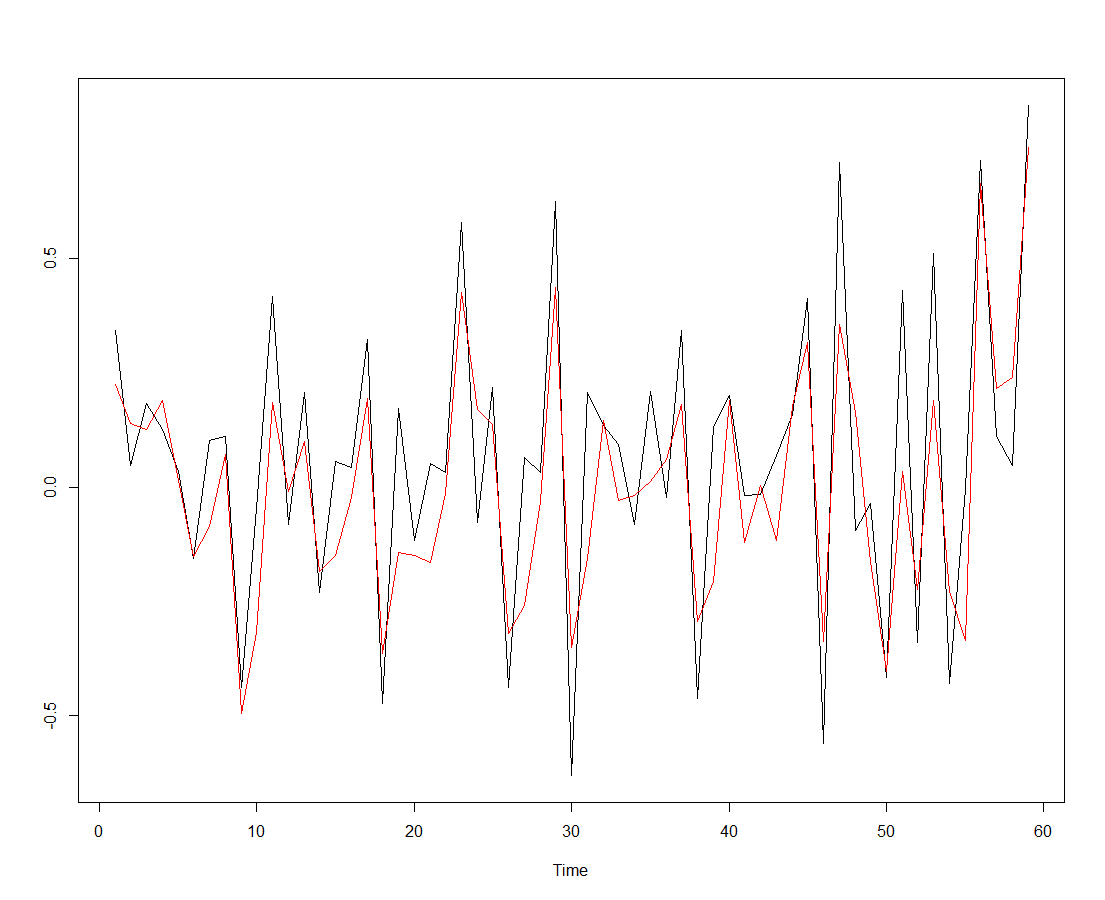
Se obtienen los siguientes valores pronóstico con sus correspondientes desviaciones estándar.



Se muestran a continuación de forma gráfica:



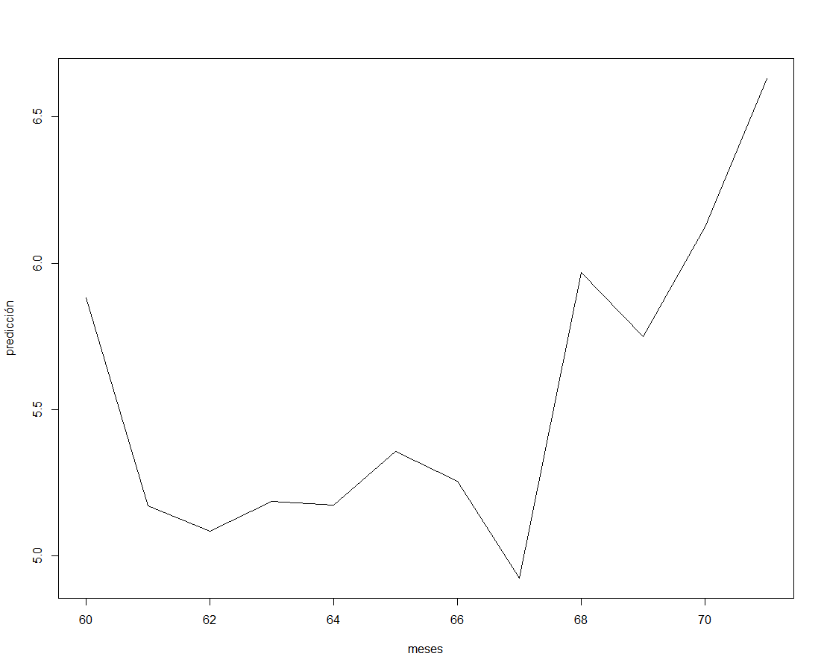
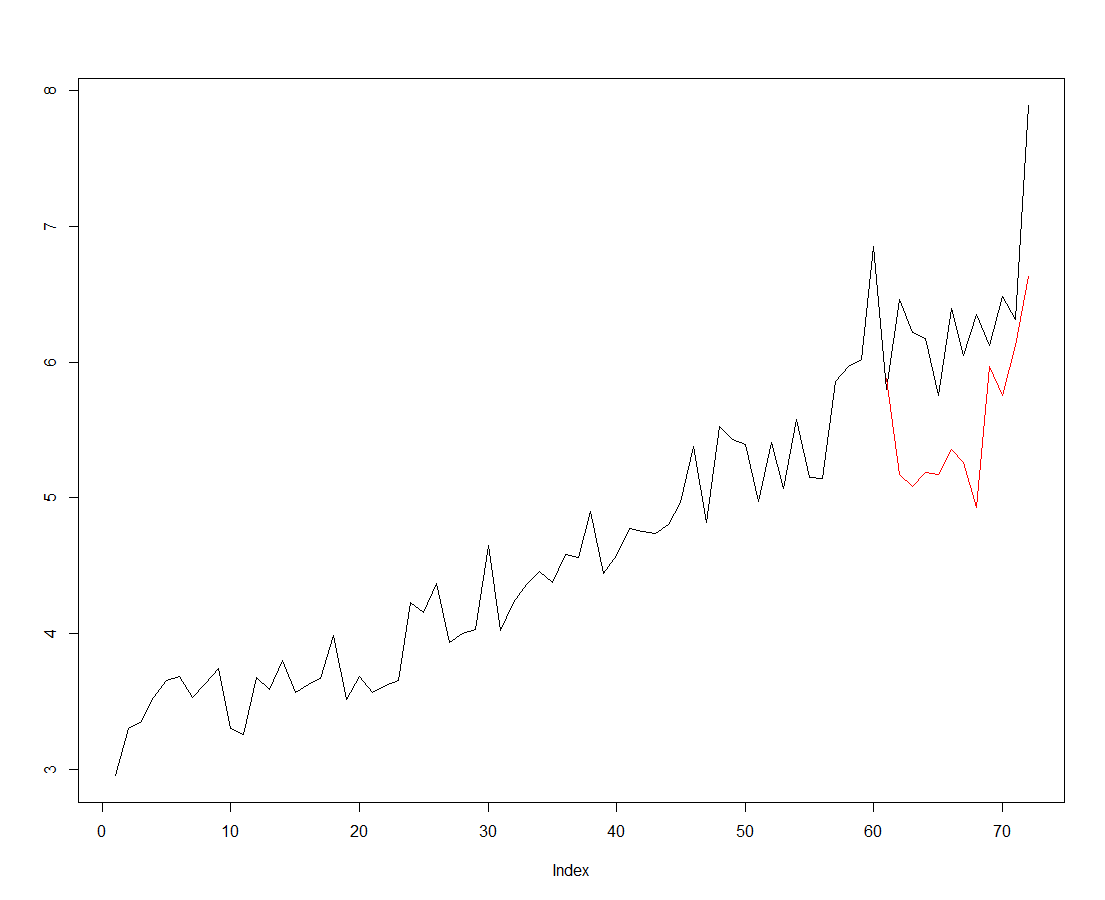
Se muestra también el gráfico de los valores diferencia con los valores de los residuos del modelo; la línea negra corresponde a los valores diferencia, y la roja a los residuos.



Se muestra a continuación las predicciones esperadas para los 12 meses del periodo extramuestral:



El gráfico de la derecha muestra las predicciones, y el de la izquierda muestra la serie temporal completa con las predicciones para el periodo extramuestral marcadas en rojo.

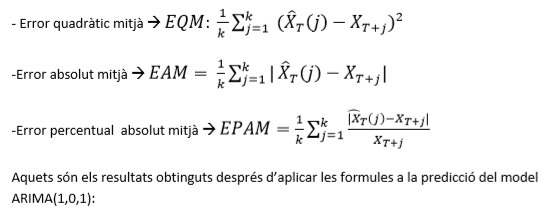


## 6. Avaluación de les predicciones

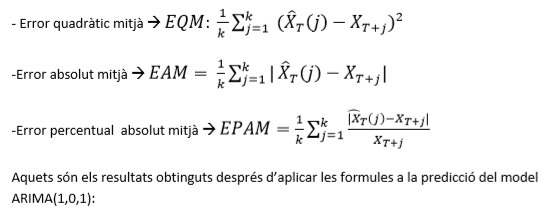
En este apartado, se pretende calcular el error medio de la predicción, para confirmar que realmente el modelo encontrado es bueno a la hora de hacer las predicciones de esta serie temporal.

Esto se hará con el cálculo del error cuadrático medio (ECM), el error absoluto medio (EAM) y el error porcentual absoluto medio (EPAM).

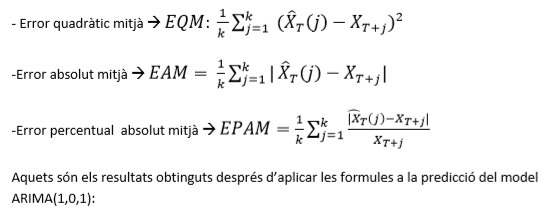
Las fórmulas que se utilizan para el cálculo de estos errores son las siguientes:



Error cuadrático medio:



Error absoluto medio:



Error porcentual absoluto medio:

Estos son los resultados obtenidos:

|  |  |
| --- | --- |
| ECM | 0.8491968 |
| EAM | 0.8054008 |
| EPAM | 0.1252032 |

En esta tabla se puede ver que el EPAM es 0.12. Y ese es un valor que indica, que se tiene una mala capacidad predictiva, el ideal del EPAM es que sea inferior al 3% y en este caso es del 12.52%. Por último se puede decir que el modelo ARIMA (1,1,2) es un mal modelo de predicción para esta serie temporal.

## 7. Comparación de modelos

Finalmente, en este apartado lo que se pretende es comparar los tres modelos estimados a lo largo de los dos trabajos, los dos hechos por el análisis clásico y el del análisis estocástico.

Los resultados de los errores de cada modelo se pueden observar en esta tabla:

|  |  |  |  |
| --- | --- | --- | --- |
|  | ECM | EAM | EPAM |
| Tendencia lineal | 84.121,40 | 240,75 | 5,530% |
| Dobles medias móviles | 90.233,69 | 234,75 | 5,24% |
| **Alisado exponencial de Holt** | **82.714,32** | **217,47** | **4,89%** |
| ARIMA (1,1,2) | 0,8491968 | 0,8054008 | 12,52% |

En esta tabla se puede observar que el mejor modelo que se ha encontrado es el Alisado exponencial de Holt ya que su EPAM es el más pequeño y su valor indica que tiene una buena capacidad predictiva. En cambio, los otros dos modelos tienen valores del EPAM muy elevados e indica que son malos predictores para esta serie temporal.

## 8. Conclusión

Se ha encontrado una serie que no tiene una presencia de estacionalidad. También se ha observado que la serie desde el 2007 hasta ahora presenta tendencia. Así que ha sido clasificada como Serie Tipo 3 en el análisis clásico.

En cuanto al análisis clásico se han realizado predicciones con los 3 modelos que le corresponden a una serie de tipo 3, estos no han sido muy malos, ya que tanto el modelo de tendencia lineal, el de dobles medias móviles y el alisado exponencial de Holt han resultado tener unos EPAMS en torno al 5%, es decir, que de los tres Holt tiene una capacidad predictiva regular (mejor los otros dos) con un EPAM del 4,89%.

En la parte estocástica se ha encontrado un modelo ARIMA (1,1,2), el modelo m2, que ha obtenido unos malos resultados en la predicción de la serie, ya que en los gráficos se ha podido observar que no era muy parecido a los datos originales. Con este modelo se ha obtenido un EPAM del 12,52%, que es peor que los obtenidos en el trabajo anterior. Se calculó con el modelo m4, que era el que recomendaba la función autoarima por si daba un EPAM mejor pero como se obtuvo un valor casi idéntico se decidió quedar-se con el modelo m2.

Por último, se puede decir que el mejor modelo para predecir la base monetaria mensual del banco central de Chile, es el **Alisado exponencial de Holt**, obtenido con el análisis clásico.