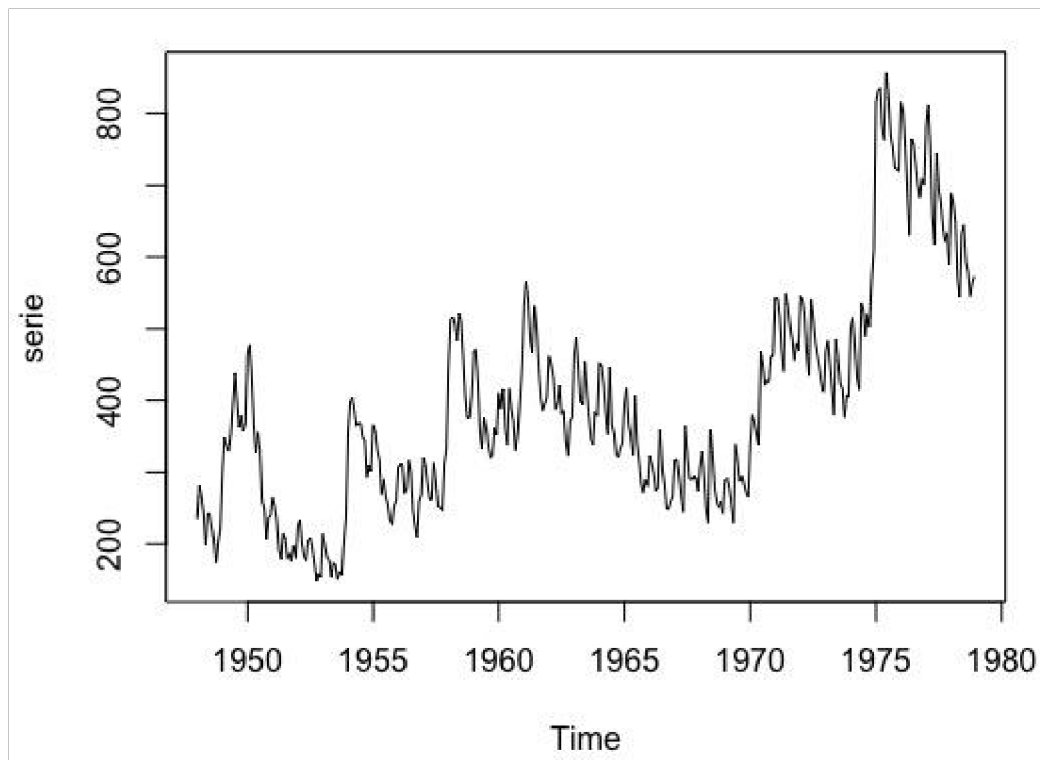
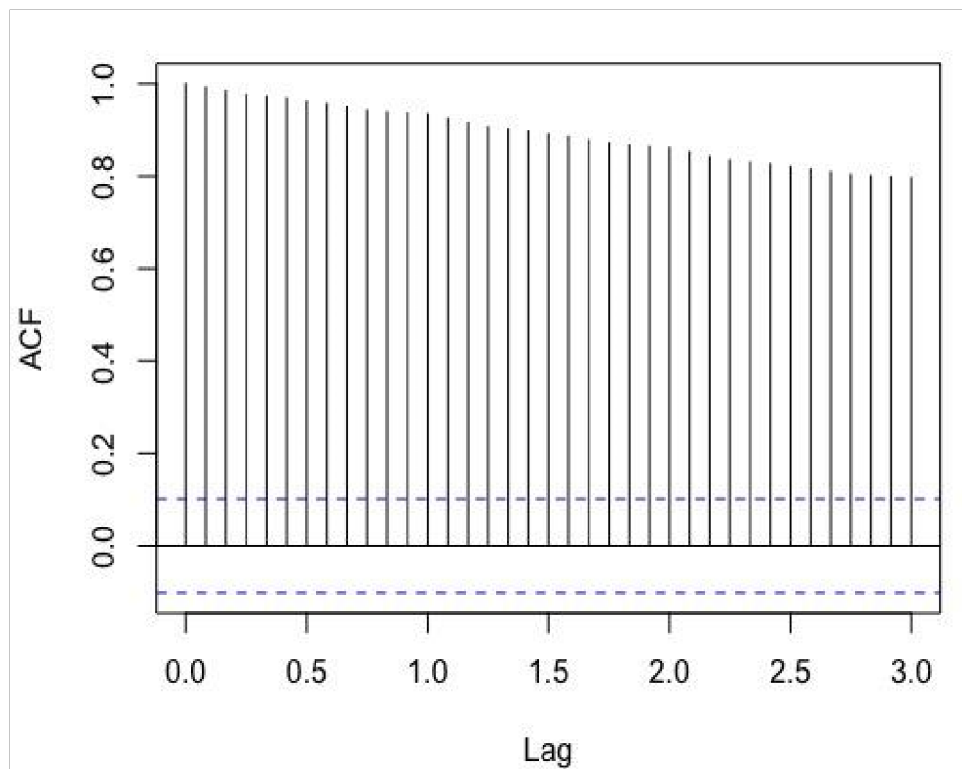
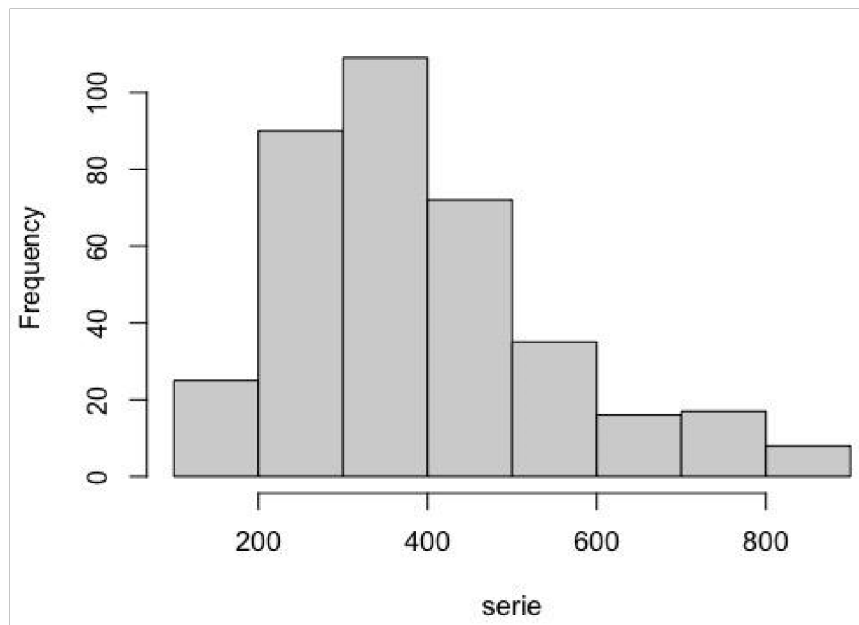


A série 5 possui 372 observações, frequência mensal, começo em 1948 e término em 1978. Abaixo encontra-se o plot da série:



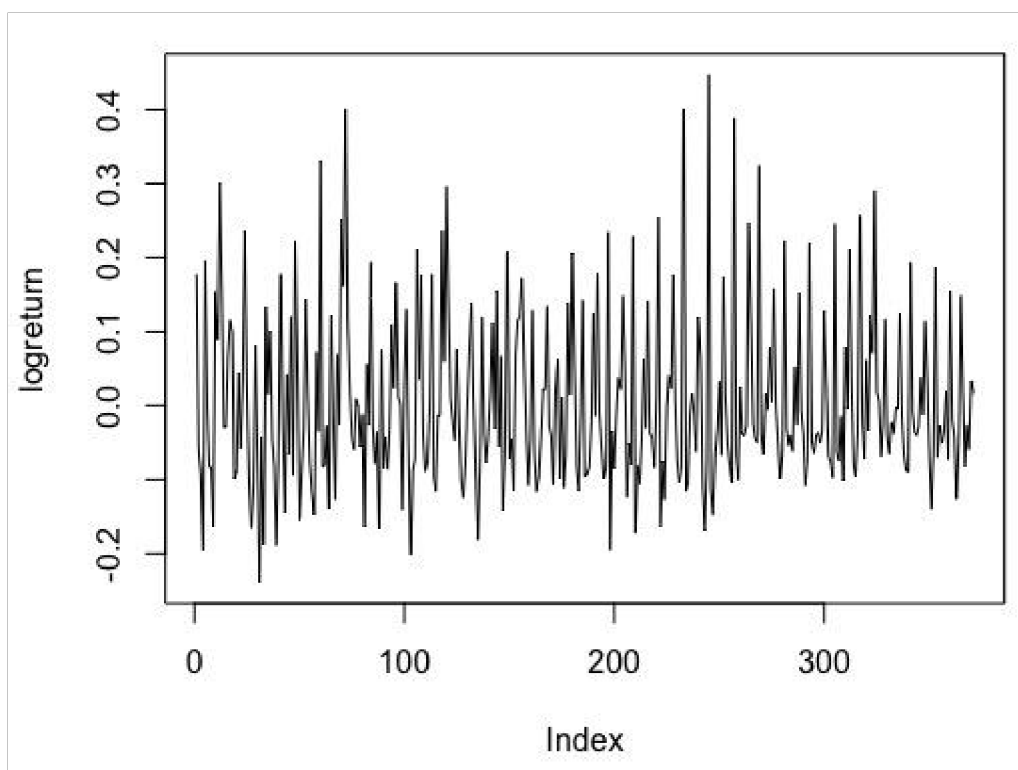
Algumas estatísticas descritivas da série podem ser encontradas abaixo:

Estatísticas	Valores
Média	393.05
Variância	24051.49
Assimetria	0.91
Curtose	3.45
Máximo	857.42
Mínimo	147.80

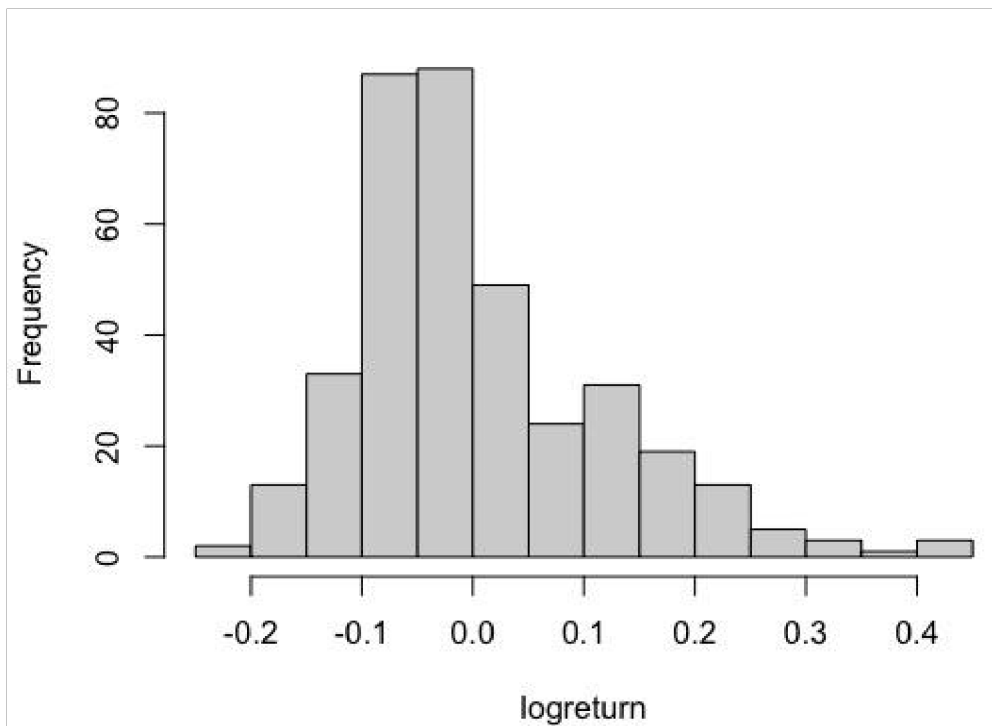


Podemos notar, pela função de autocorrelação da série, que há tendência temporal e que sua distribuição é muito próxima de uma normal. Observando o histograma da série, podemos notar uma leve assimetria para esquerda e caudas levementes pesadas, que são confirmadas pelo coeficiente de assimetria de 0.92 e curtose de 3.45. Além disso, a série possui alta variabilidade.

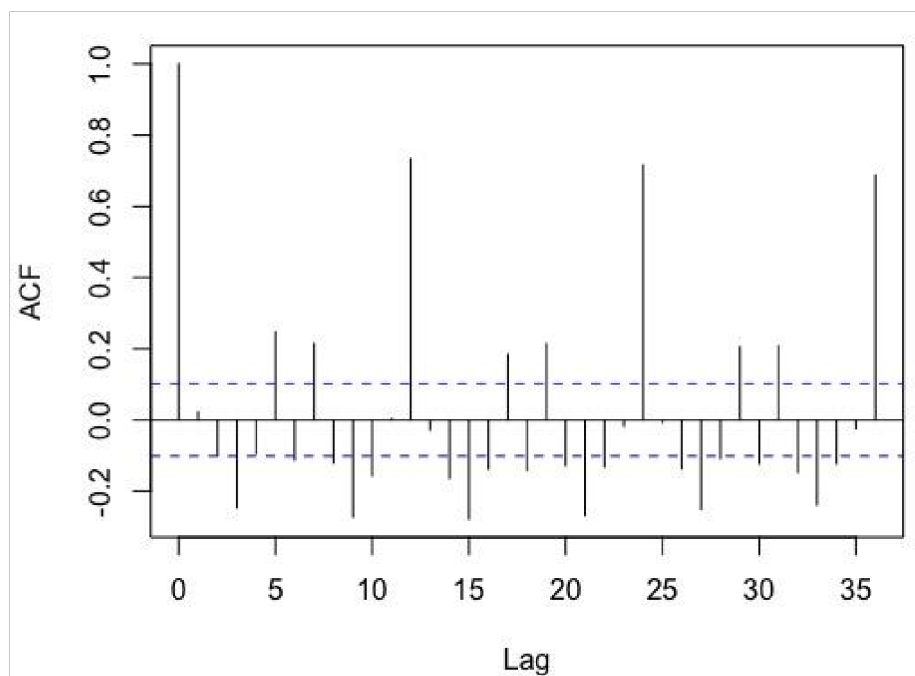
Como transformação da série é indicado fazer os log-retornos, em que $x_t = \log(y_t)$ e Δx_t , com intuito retirar a tendência temporal, diminuir a variabilidade e aproximar mais sua distribuição a de uma normal. Segue abaixo o plot dos log-retornos e suas estatísticas descritivas:



Estatísticas	Valores
Média	0.002
Variância	0.012
Assimetria	1.04
Curtose	4.13
Máximo	0.44
Mínimo	-0.23

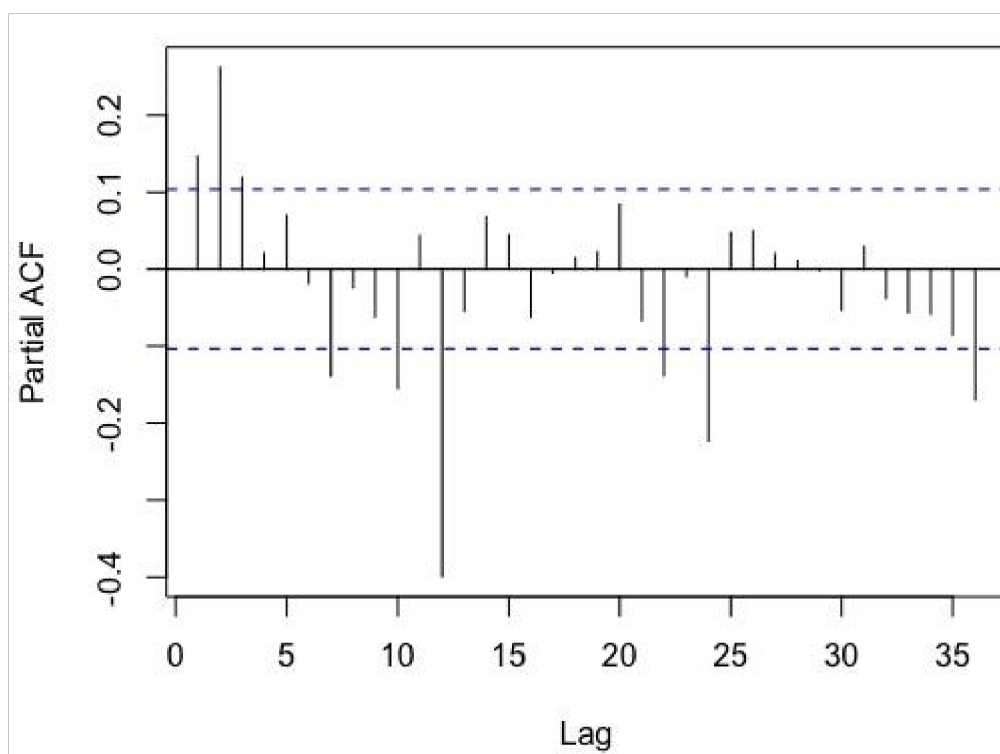
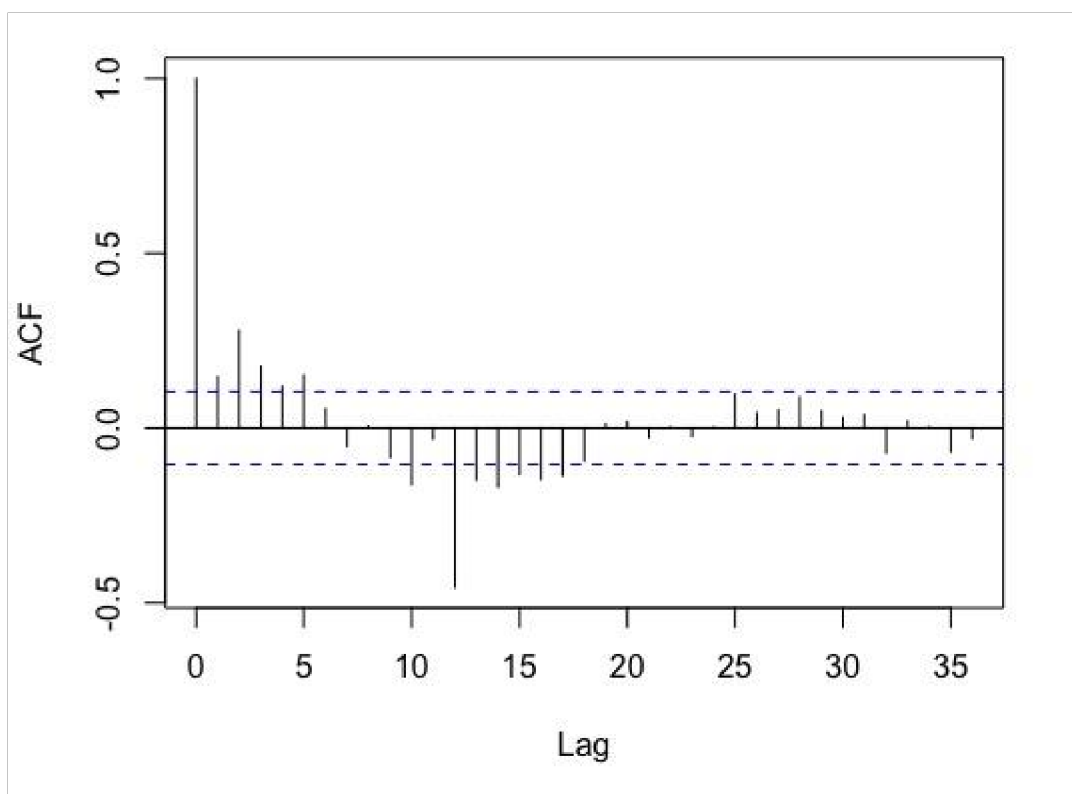


Embora os log-retornos ainda possuam uma leve assimetria para esquerda (coeficiente de assimetria = 1.04) e caudas levemente pesadas (coeficiente de curtose = 4.13), a variabilidade da nova série diminuiu muito. Além disso, devido às primeiras diferenças, a tendência temporal foi eliminada, no entanto, ainda restou uma tendência sazonal, como pode ser visto na função de autocorrelação temporal dos log-retornos:



Para tanto, tiramos as primeiras diferenças sazonais no período 12, ou seja, temos $\Delta\Delta_{12}x_t$.

O resultado é a eliminação da tendência sazonal:



Chegamos em uma série transformada da original sem nenhum tipo de tendência e, então, podemos encontrar o modelo SARIMA apropriado. Quanto ao componente ARMA sazonal, devido a PACF cair exponencialmente ao longo dos lags múltiplos de 12 e a ACF possuir uma correlação alta no lag 12 e nenhuma estatisticamente significativa no seus múltiplos de 12, escolhemos um SMA(1). Quanto ao ARMA típico, iremos contemplar inicialmente um AR(3), pois há um decaimento exponencial na ACF e os 3 primeiros lags estatisticamente significantes na PACF.

A primeira alternativa, um SARIMA (3,1,0)x(0,1,1)[12], apresenta os seguintes resultados:

z test of coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
ar1	0.06042277	0.05311103	1.1377	0.2552587	
ar2	0.18658675	0.05206442	3.5838	0.0003387	***
ar3	0.09092996	0.05295591	1.7171	0.0859631	.
sma1	-0.71050417	0.04526890	-15.6952	< 2.2e-16	***
intercept	-0.00032533	0.00153394	-0.2121	0.8320386	

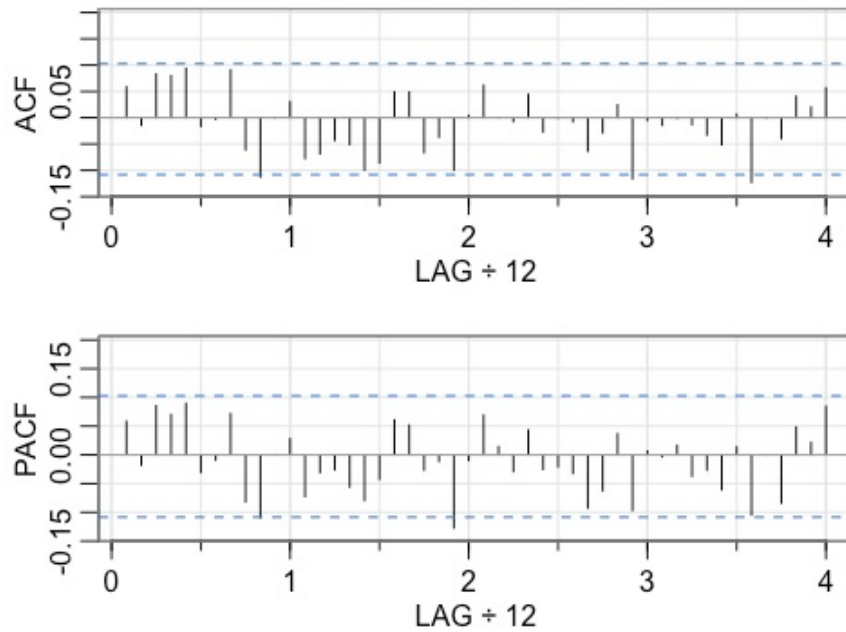
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Vemos os coeficientes ar3, ar1 e a média não são significantes a 5%. Por essa razão, iremos tentar outro modelo, um SARIMA (2,0,0)x(0,0,1)[12] sem média e com o componente ar1 zerado. Seguem os resultados:

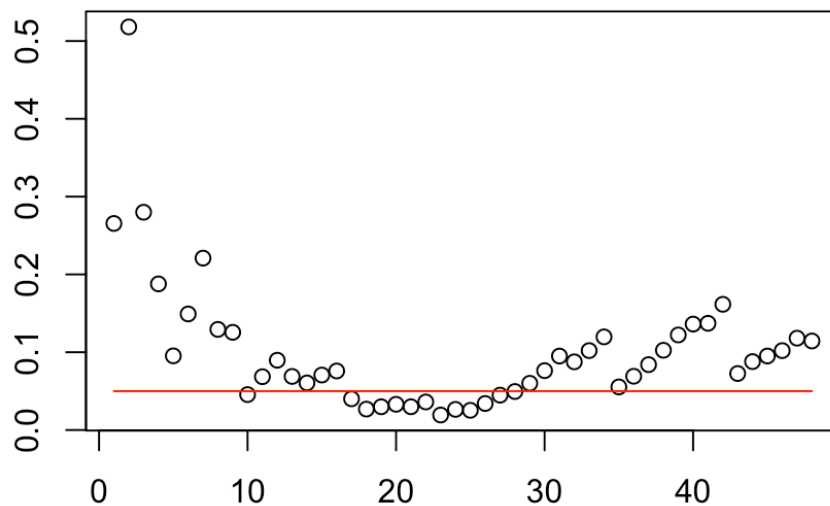
z test of coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
ar2	0.198886	0.052301	3.8027	0.0001431	***
sma1	-0.727253	0.044867	-16.2089	< 2.2e-16	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1



Temos então resíduos que aparentam ser de um resíduo branco, no entanto, ao realizarmos o teste de Ljung-Box, notamos a presença de autocorrelações estatisticamente significantes nos resíduos do modelo SARIMA (2,0,0)x(0,0,1)[12], como podemos ver na figura abaixo:

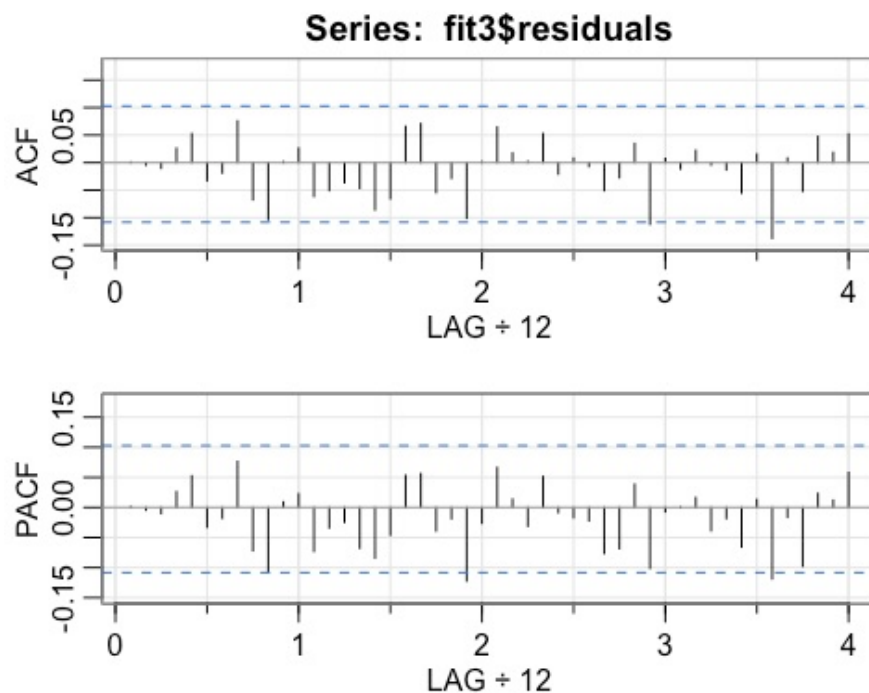


Resta-nos então tentar estimar outro modelo. Já que a estimativa do $ar(3)$ foi não significativa, uma outra interpretação das ACF e PACF pode nos levar a acreditar que ambas “aparentam” ter um decaimento exponencial. Característica de modelos ARMA. A próxima tentativa será um SARIMA(2,1,1)x(0,1,1)[12] sem média. Abaixo estão dispostos os resultados:

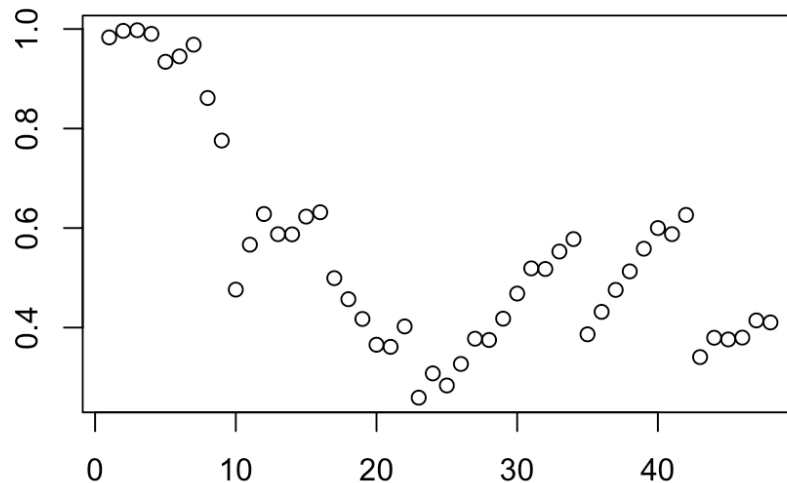
z test of coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
ar1	0.560613	0.144694	3.8745	0.0001069	***
ar2	0.149951	0.060644	2.4726	0.0134117	*
ma1	-0.509591	0.140540	-3.6260	0.0002879	***
sma1	-0.712666	0.044587	-15.9836	< 2.2e-16	***

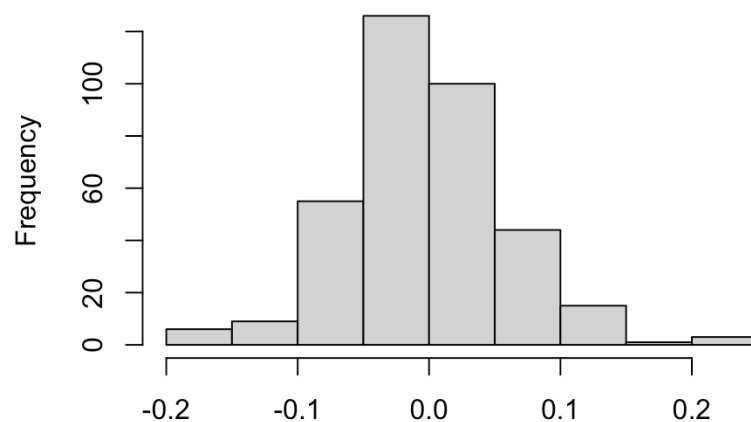
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1



Todos os coeficientes estimados são significantes a 5% e tanto a ACF quanto a PACF dos resíduos do modelo aparentam ser de um ruído branco. Ao olhar os resultados do teste de Ljung-Box sobre os resíduos temos a confirmação que são, de fato, ruídos brancos.



Apesar disso, os resíduos ainda não têm distribuição normal, como pode ser corroborado pela rejeição da hipótese nula do teste Jarque-Bera, onde a estatística chi-quadrado foi de 15.717 com p-valor 0.0003 a 95% de significância. No histograma dos resíduos, podemos notar caudas pesadas:



Para fazer previsões, elas serão feitas utilizando-se do modelo SARIMA(2,1,1)x(0,1,1)[12] sem média. Abaixo encontra-se as estimativas para 12 períodos à frente mais os intervalos de confiança de 80% (1 desvio padrão) e de 95%.

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
Jan 1979	6.511767	6.433974	6.589561	6.392792	6.630742
Feb 1979	6.524815	6.411957	6.637674	6.352213	6.697418
Mar 1979	6.476534	6.328592	6.624476	6.250276	6.702792
Apr 1979	6.370266	6.189405	6.551127	6.093663	6.646869
May 1979	6.313763	6.101650	6.525877	5.989365	6.638162
Jun 1979	6.504453	6.262750	6.746156	6.134801	6.874105
Jul 1979	6.478155	6.208421	6.747888	6.065632	6.890677
Aug 1979	6.418753	6.122442	6.715064	5.965585	6.871921
Sep 1979	6.392716	6.071170	6.714262	5.900954	6.884478
Oct 1979	6.349206	6.003654	6.694758	5.820730	6.877682
Nov 1979	6.383225	6.014791	6.751660	5.819754	6.946697
Dec 1979	6.373043	5.982748	6.763337	5.776139	6.969946