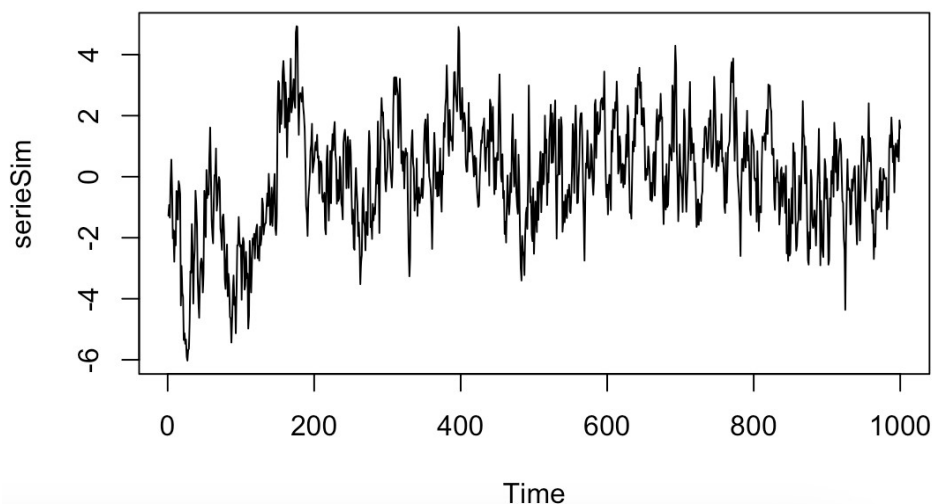


O exercício abaixo propõe simular um modelo ARFIMA(1,d,0) com $\phi = 0.3$ e $d = 0.45$. Em seguida, estimam-se um modelo ARFIMA e outro ARIMA para comparar as diferenças. Esse exercício procura ressaltar que tirar as primeiras diferenças de séries com presença de memória longa não é suficiente para lidar adequadamente com a persistência temporal da série.

I – Modelo ARFIMA

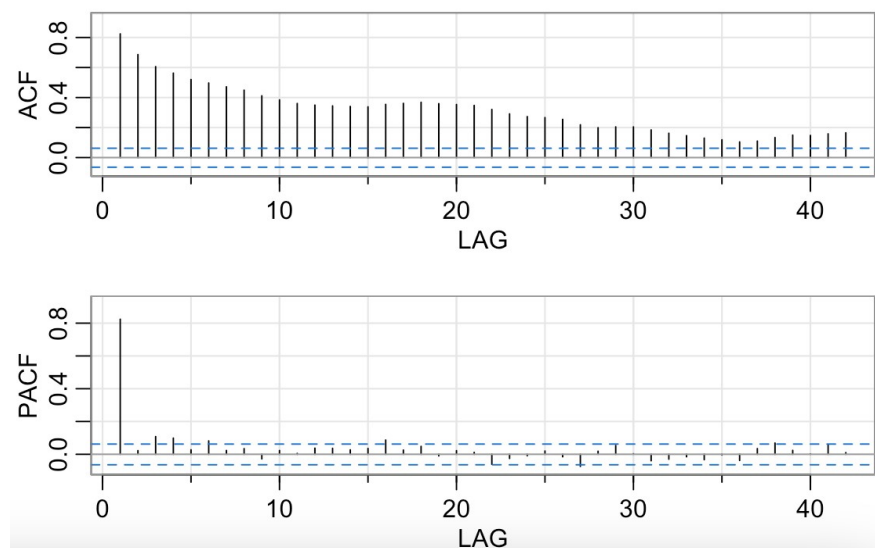
Pela forma como o gráfico se desenvolve ao longo do tempo, podemos notar uma persistência perdurante. Além disso, pelo padrão que se repete bem similarmente entre as subidas e decidas, podemos perceber que uma observação tem correlação com observações a muitos lags atrás. Entretanto, apesar de ser uma série com forte persistência que abarca lags muito distantes, é uma que se dissipa ao longo do tempo, não criando uma tendência no gráfico, como no caso de uma série com raiz unitária.

Figura 1 – Plot da Série



Ao observarmos a ACF, podemos ver a presença da memória longa. O decaimento da ACF não é exponencial, como no caso de um modelo ARMA puro, mas hiperbólico e bem vagaroso ao longo dos lags. Já na PACF, há o primeiro lag altamente sobressaltado e alguns bem pequeninos (muito próximos da linha de intervalo de confiança de 95%) nos lags 3, 4, 6 e outros poucos mais longínquos, o que é característico de processos com memória longa.

Figura 2 – Função de Autocorrelação e Autocorrelação Parcial



Um modelo ARFIMA(1, d, 0) sem constante é estimado nos dados simulados, onde "d" será estimado em conjunto com todos os outros parâmetros do modelo:

Figura 3 – Tabela de Resultados

Call:

```
arfima(z = serieSim, order = c(1, 0, 0), dmean = FALSE)
```

Mode 1 Coefficients:

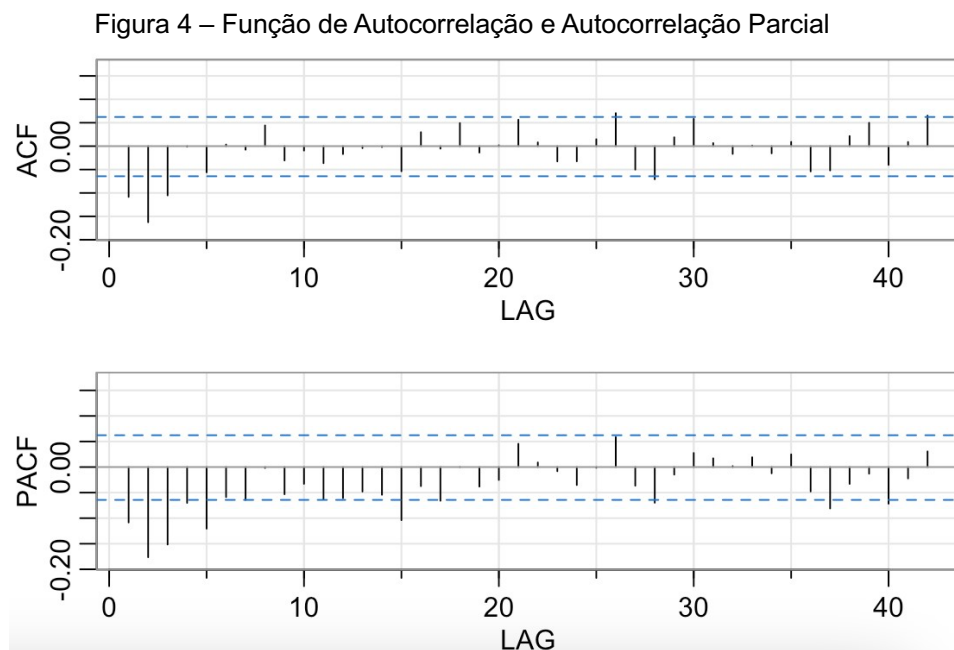
	Estimate	Std. Error	Th. Std. Err.	z-value	Pr(> z)
phi(1)	3.58233e-01	6.10328e-02	6.81469e-02	5.86951	4.3707e-09 ***
d.f	4.10984e-01	4.95076e-02	5.69122e-02	8.30144	< 2.22e-16 ***
zbar	-1.79412e-16	NA	NA	NA	NA

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

sigma^2 estimated as 0.986377; Log-likelihood = 6.58344; AIC = -5.16687; BIC = 14.4641

Pelas estimativas, temos um $\hat{d} = 0.41$. Tanto \hat{d} quanto $\hat{\phi}$ foram significantes com 99% de confiança. No entanto, embora não por muito, percebemos que o modelo arfima ajustado tendeu a subestimar \hat{d} e superestimar $\hat{\phi}$ quando comparados aos respectivos valores escolhidos para simulação.

Por sua vez, um ARMA é estimado após retirar as primeiras diferenças (retornos) da série simulada. São obtidas as seguintes ACF e PACF:



Notamos uma espécie de decaimento exponencial da PACF ao longo dos lags e lags 1,2,3 fortemente sobressaltados na ACF. Vou ajustar, portanto, um ARMA(0,3) sem constante na série de retornos. Segue o valor das estimativas dos parâmetros:

Figura 5 – Tabela de resultado

	Estimate	Std. Error	z value	Pr(> z)	
ma1	-0.204641	0.031453	-6.5062	7.707e-11	***
ma2	-0.249502	0.033433	-7.4629	8.466e-14	***
ma3	-0.173649	0.033024	-5.2582	1.455e-07	***

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

O ajuste do modelo foi muito bom, como veremos no item à frente, por isso, não foi proposto modelos alternativos.

Seguem os diagnósticos do fit 1 - ARFIMA(1, 0.41, 0) - sobre a série simulada:

Figura 6 – Plot dos Resíduos

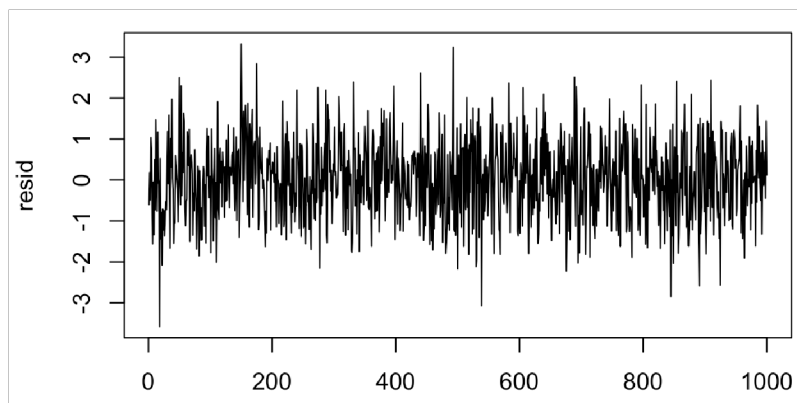


Figura 7 – Função de Autocorrelação e Autocorrelação Parcial dos Resíduos

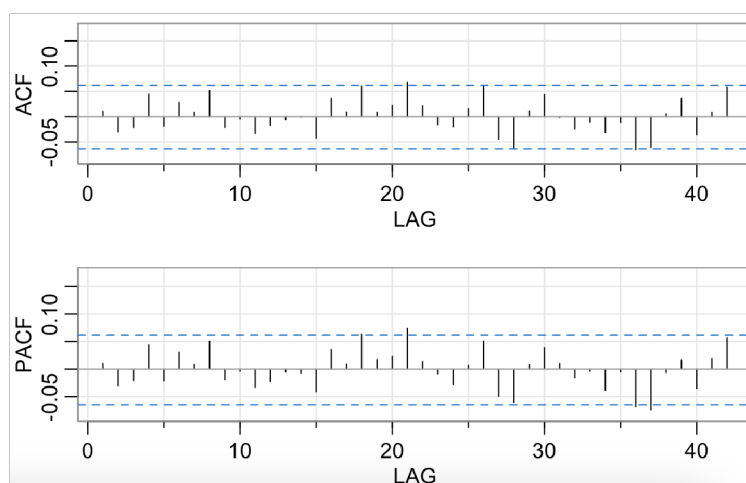


Figura 8 - Q-Q Plot Normal sobre os Resíduos

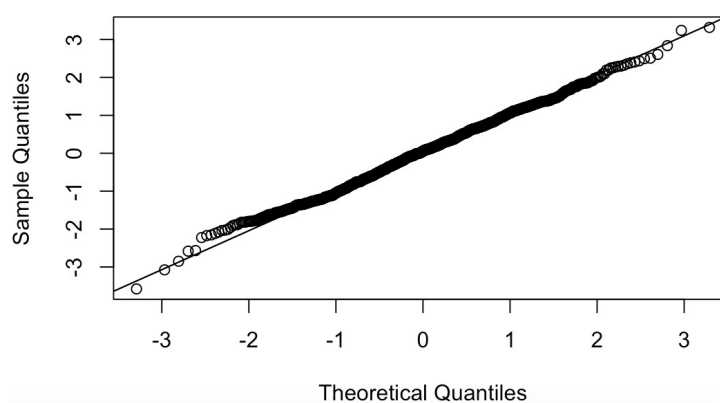
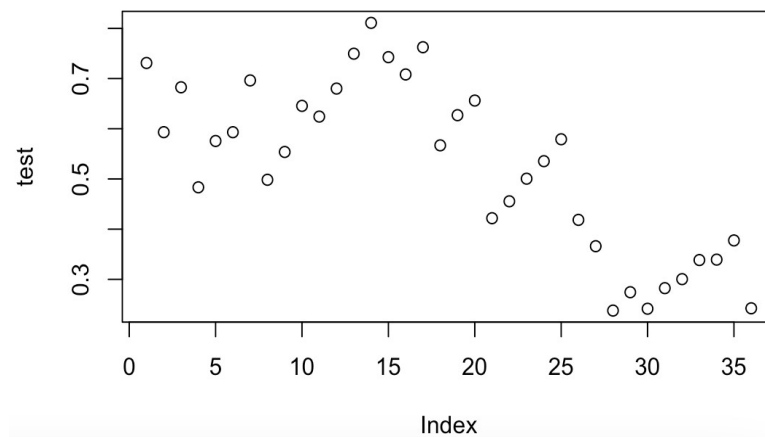


Figura 9 – Teste de Ljung-Box sobre os Resíduos



Comparativamente, é analisado como o do fit 2 - ARMA(0, 3) - sobre a série de retornos se comporta:

Figura 10 – Plot dos Resíduos

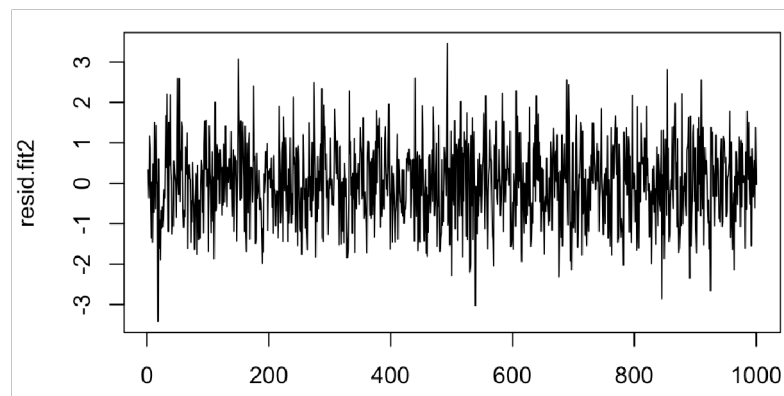


Figura 11 – Função de Autocorrelação e Autocorrelação Parcial dos Resíduos

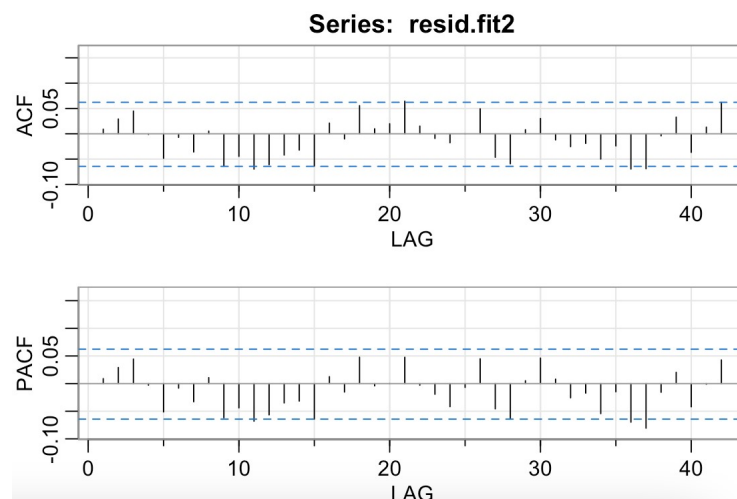


Figura 12 - Q-Q Plot Normal sobre os Resíduos

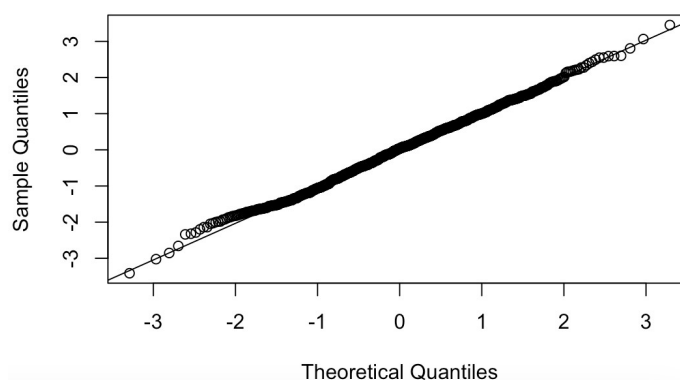
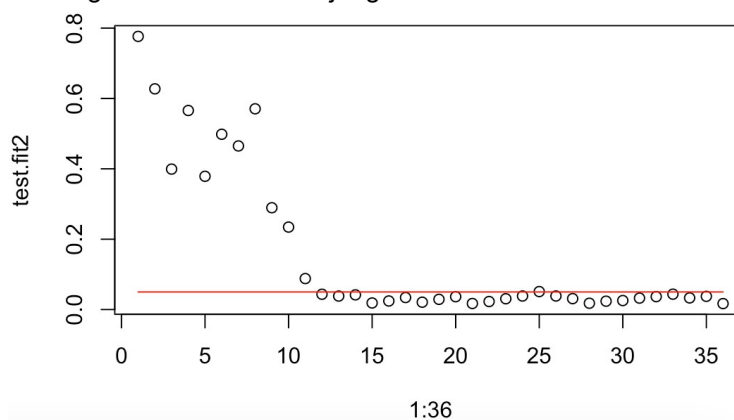


Figura 13 – Teste de Ljung-Box sobre os Resíduos



Quando comparamos os 2 modelos, percebemos que ambos os resíduos apresentam tanto plot, ACF e PACF muito próximas do que seria esperado de um ruído branco e a performance dos 2 modelos foi muito próxima. Quanto à questão da normalidade de seus respectivos resíduos, ambos os modelos capturam bem o meio da distribuição, mas deixando um pouco a desejar quanto às caudas. Apesar disso, o fit 1 se saiu melhor no teste de Ljung-box, não rejeitando a hipótese nula de não haver autocorrelação em nenhum dos 36 lags. O teste de Ljung-box para o fit 2 mostrou que o modelo extraiu as autocorrelações de lag menores, mas não foi capaz de eliminá-las em lags de maiores ordens (hipótese nula é rejeitada a partir do lag 11), indicando que o modelo ARIMA(0,1,3) não capturou bem a memória longa presente na série simulada. Outro aspecto interessante de destacar é que o AIC do primeiro modelo foi muito melhor (-5.16), ganhando do segundo (2822). O modelo escolhido é, portanto, o fit 1.

II – Modelo ARFIMA

Por último faz-se as previsões 12 períodos à frente usando o modelo 1.

Tabela 1 - Tabela com os valores das previsões 12 lags à frente e seus respectivos IC a 95%

lags	Valores previstos	IC inferior a 95%	IC superior a 95%
1	1.23	-0.7104	3.1704
2	0.99	-1.46	3.44
3	0.84	-1.8452	3.5252
4	0.73	-2.0924	3.5524
5	0.65	-2.2508	3.5508
6	0.59	-2.3696	3.5496
7	0.55	-2.4488	3.5488
8	0.51	-2.528	3.548
9	0.48	-2.5776	3.5376
10	0.45	-2.6468	3.5468
11	0.42	-2.6964	3.5364
12	0.4	-2.736	3.536

Figura 14 - Gráfico das previsões 12 lags à frente

