**Lista de exercícios Matemática Computacional**
**Parte B – Prof. Dr. Reinaldo Rosa - 2020**

Denis M. A. Eiras

## Exercício 3 - Descrição

Repita o exercício anterior considerando, entretanto, o algoritmo pmodel.py. Neste caso, considere N= 8192 e diversifique os dados em 2 famílias: endógeno e exógeno.

Para a família endógeno (0.4) considerando 3 ou mais valores *p* na faixa de (0,32 - 0,42) e exógeno (0.7) também considerando 3 ou mais valores *p* na faixa de (0,18 - 0,28).
Produza pelo menos 30 sinais para cada família, 10 para cada valor de *p*. Total do Grupo pmnoise: 60.

## Exercício 3 – Detalhes da implementação

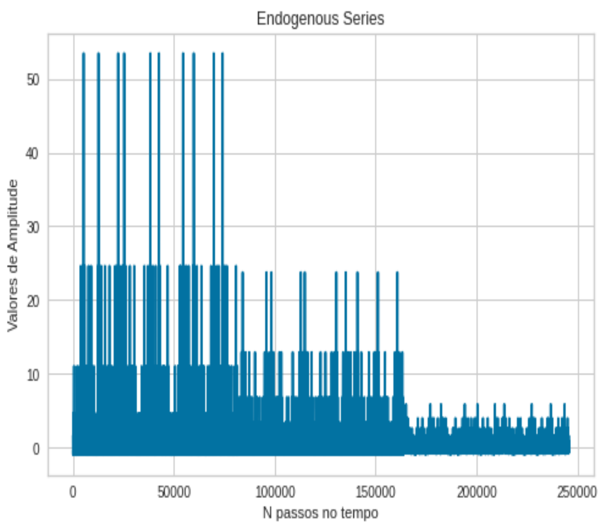
Para a resolução do Exercício 3, foi implementado o programa principal exercicio3.py para executar o exercício por completo, que utiliza funções dos programas do exercício 1: gerador_de_momentos, do exercicio1_2.py e k_means_e_metodo_do_cotovelo, do exercicio1_3.py.

O exercício 3 funciona de maneira análoga ao exercício2, mas implementando a geração de sinais pmodel em uma função própria, onde é possível utilizar os parâmetros que determinam, a quantidade de elementos (famílias), número de sinais e betas.

Mais detalhes podem ser consultados na própria documentação do programa.

## Análise

Foram gerados 10 sinais para cada um dos 3 valores das séries endógenas, 0.32, 0.36, 0.42, e 10 sinais para cada 3 valores para das séries exógenas, 0.18, 0.22, 0.28, que podemos observar nos sinais abaixo.



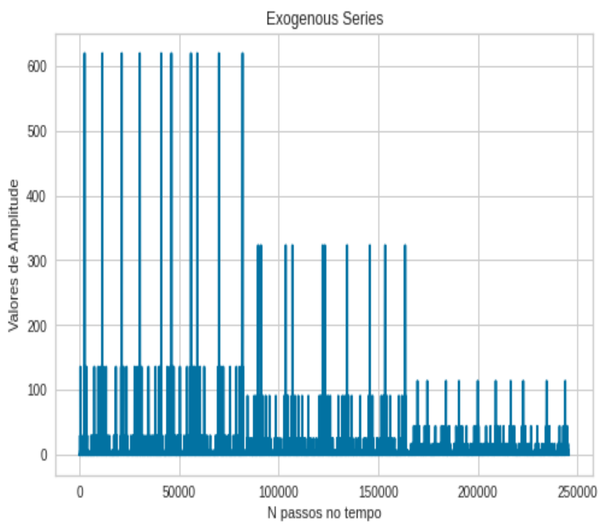




**Figura 1. a) Séries endógenas; b) Séries exógenas.**

Os histogramas da família endógena apresentam uma curva log normal, com pouca variabilidade próxima da média. A cauda da curva aumenta com o aumento do valor de p, onde a variabilidade é mínima, para p=0,32, na figura 2.a, e máxima, para p=0,42, na figura 2.b.

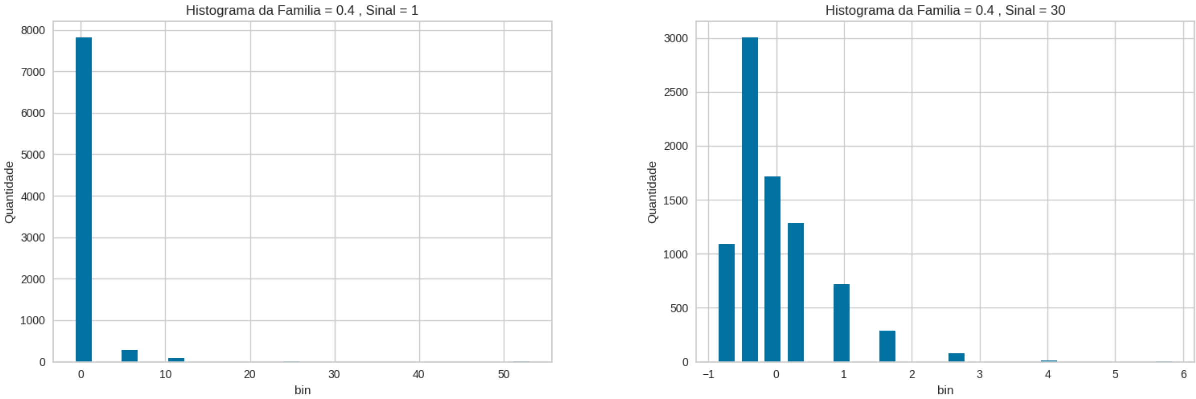



**Figura 2. Dois sinais da família endógena: a) p = 0,32; b) p = 0,42**

Os histogramas da família exógena támbém apresentam uma curva log normal, porém com muito pouca variabilidade próxima da média. A cauda da curva aumenta minimalmente com o aumento do valor de p, onde a variabilidade é mínima, para p=0,18, na figura 3.a, e máxima, para p=0,28, na figura 3.b.

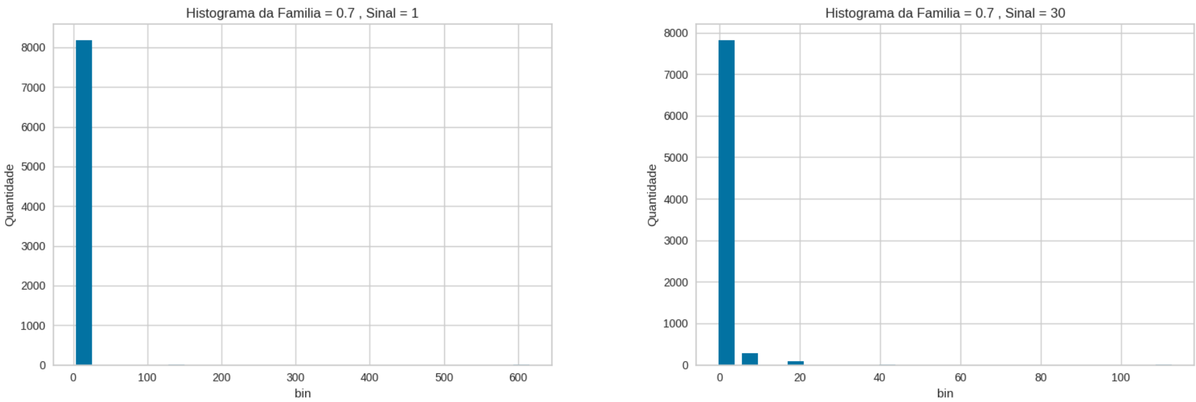



**Figura 3. Dois sinais da família exógena: a) p = 0,18; b) p = 0,28**

Para verificar se há classes nos espaço de parâmetros composto por variância, assimetria e curtose, o algoritmo K-means foi executado com k entre 2 e 14.

Observando a figura 4.a, para k=2, não é possível separar perfeitamente as duas classes endógenas e exógenas, pois se observa que, uma das classes contém 4 pontos, com uma maior variância, e a outra contém apenas 2 pontos, com uma baixa variância, quando o ideas seriam 3 pontos para cada.

Conforme k aumenta, a variância diminui muito, implicando em um *warnign* emitido pelo algoritmo k-means: “*Data must have variance to compute a kernel density estimate*”, a partir de k =4. Com k = 6, os agrupamentos ficam tão próximos que o gráfico não exibe dados na diagonal principal, como pode ser visto na figura 4.b. Em diversas execuções ocorre falhas no método, e os gráficos não são apresentados ou apresentam erros.

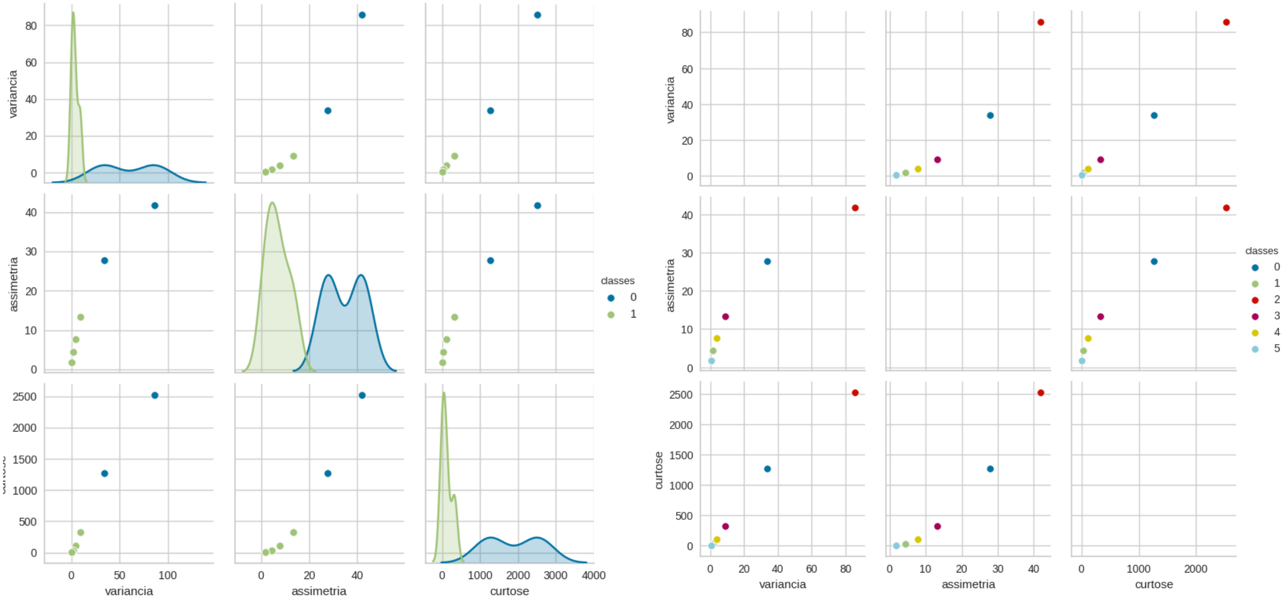



**Figura 4. a) K-means para k=2; b) K-means para k=6**

Como as classes são bem definidas para cada valor de p, na figura 5.a, o método do cotovelo exibe o melhor agrupamento para o primeiro k após 2, onde k = 3, o que se confirma no método da silhueta para esses valores de k. No entanto, o método da silhueta diz que o melhor k é 6, agrupando perfeitamente os 3 valores de p endógenos e 3 valores de p exógenas, como é possível observar na figura 5.b.

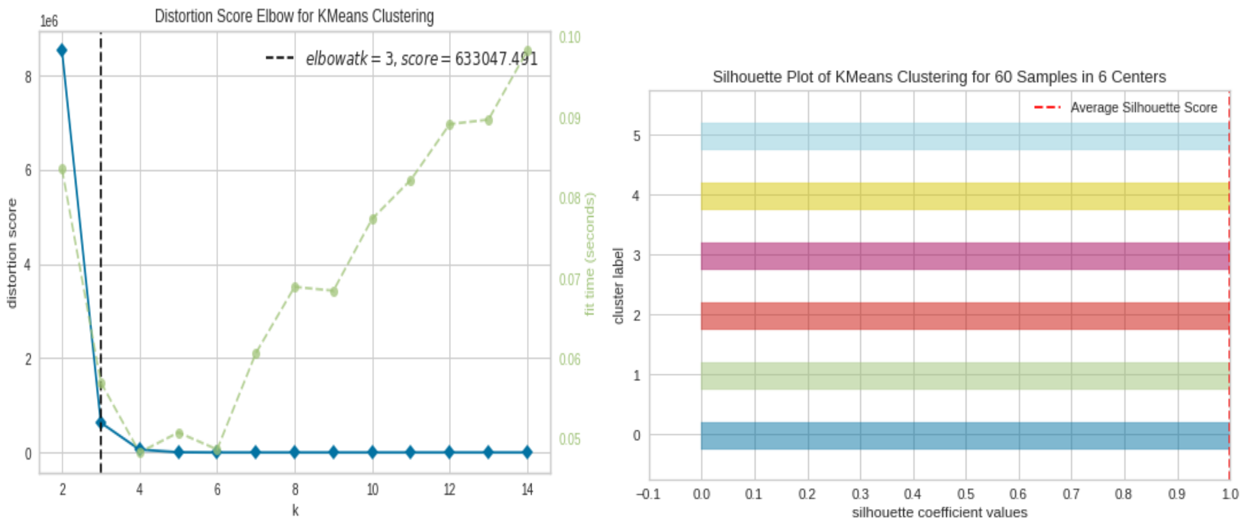



**Figura 5. a) método do cotovelo do pacote yellowbrick; b) método da Silhueta**

## Exercício 3 – Conclusão

A pequena variação dos sinais de um mesmo tipo, isto é, com um mesmo valor de p, fazem com que o K-means agrupe os sinais de valores de p mais próximos, pois têm momentos estatísticos muito semelhantes. Por esse motivo, quando k = quantidade de valores de p, o agrupamento é perfeito, como indica o método da silhueta para k = p = 6, na figura 5.b.