

**Prova de Matemática Computacional – CAP 239-B**  
**Parte A – Prof. Dr. Reinaldo Rosa - 2020**

Aluno: Denis M. A. Eiras

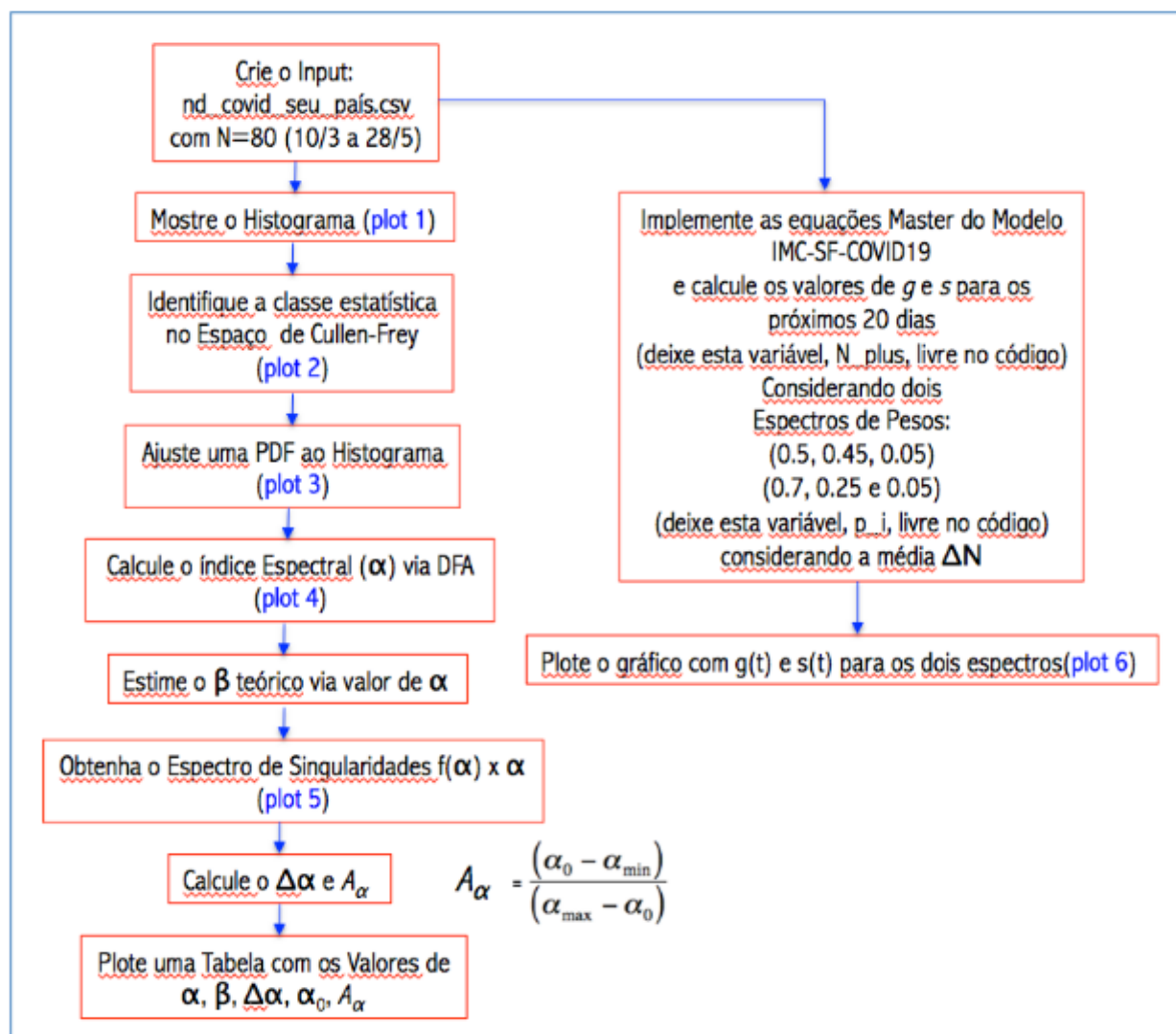
### Parte A (Valor 7)

Implemente dois códigos em Python e obtenha os resultados identificados nos fluxogramas da Figura abaixo.

- Branch esquerdo: 5 pontos
- Branch da Direita: 2 pontos

### Parte B (Valor 3)

Três questões objetivas de Múltipla Escolha em Relação aos Resultados da Parte A (na *live*).



The in-homogeneous multiplicative cascade structural configuration admits as inputs the seven-days rolling average for the number of confirmed cases ( $\langle N_{nb} \rangle_7$ ) and today ( $N_{kt}$ ), to predict the number of confirmed cases tomorrow ( $N_s$ ).

The master formula of the model is as follows:

$$N_{s_{min}} = g(2 \times n_1 + 4 \times n_2 + 5 \times n_3) \quad (1)$$

$$N_{s_{max}} = g(4 \times n_1 + 7 \times n_2 + 10 \times n_3), \quad (2)$$

where  $N_{s_{min}}$  and  $N_{s_{max}}$  are the respective values (minimum and maximum) of daily infection cases estimated for the next day,

with

$$n_1 = p_1 \times N_{kt}, \quad (3)$$

$$n_2 = p_2 \times N_{kt}, \quad (4)$$

$$n_3 = p_3 \times N_{kt}. \quad (5)$$

and

$$g = \frac{\langle N_{nb} \rangle_7}{N_{kt}} \quad \text{if} \quad N_{kt} > \langle N_{nb} \rangle_7 \quad (6)$$

and

$$g = \frac{N_{kt}}{\langle N_{nb} \rangle_7} \quad \text{if} \quad N_{kt} < \langle N_{nb} \rangle_7 \quad (7)$$

The model also allows calculating the suppression factor,  $s(t)$ , based on the derivatives of  $g : \Delta_g$  and  $n : \Delta_{nk}$  as follows:

The derivative  $\Delta_g$  is defined as

$$\Delta_g = (g_0 - g) - q_g \quad \text{if} \quad g_0 < g$$

or

$$\Delta_g = (g_0 - g) + q_{g_0} \quad \text{if} \quad g_0 \geq g,$$

where

$$q_g = (1 - g)^2 \text{ and } q_{g_0} = (1 - g_0)^2.$$

The derivative  $\Delta_{nk}$  is defined as

$$\Delta_{nk} = \frac{(\langle N_{nb} \rangle_7 - N_{kt})}{N_{kt}} \quad (8)$$

so that,

$$s = \frac{2\Delta_g + \Delta_{nk}}{3} \quad (9)$$

## Parte A

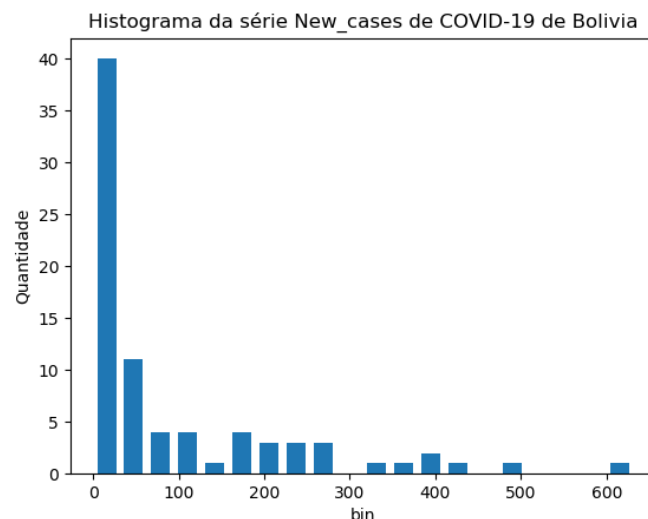
A implementação utilizada na prova pode ser verificada no repositório em [https://github.com/deniseiras/prova\\_cap239\\_b.git](https://github.com/deniseiras/prova_cap239_b.git), no diretório Prova. O programa principal é o `prova_parte_A.py`, que parametriza os dados de entrada e executa as funções `executa_branch_esquerda` (`prova_fluxograma.py`) e `executa_branch_direita` (`prova_modelo_covid.py`).

### Branch Esquerdo

Foram utilizados os dados de casos diários do país Bolívia, entre 10/03 a 28/05 e executadas as tarefas do fluxograma.

Plot 1 - Mostrar o histograma (figura 1).

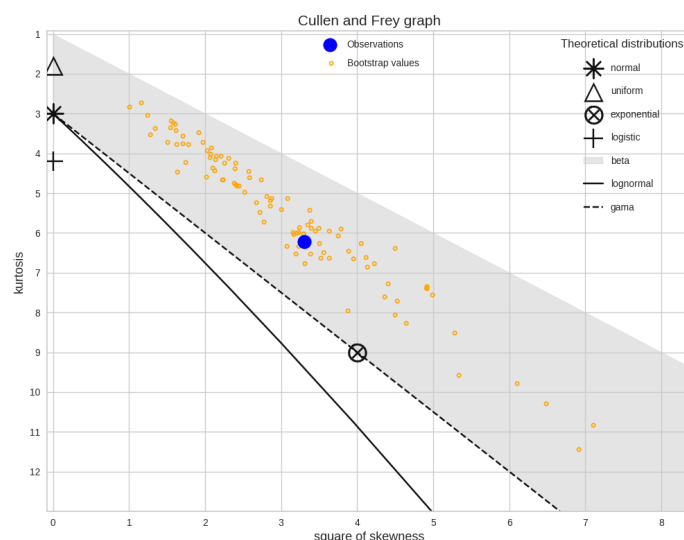
O histograma exibe mais valores para um menor número de casos, e apresenta uma distribuição mais uniforme com o aumento do número de casos. A grande amplitude no início do histograma se dá devido ao baixo número de casos do início da série, onde o primeiro caso é contabilizado em 12/03/2020, dois dias depois do intervalo considerado (10/03/2020 a 28/05/2020).



**Figura 1. Histograma.**

Plot 2 - Identificar a classe estatística no espaço de Cullen-Frey (figura 2).

O plot de Cullen e Frey, na figura 2, nos mostra que, a série se enquadra totalmente dentro do espaço  $\beta$ , indicando uma Função de Probabilidade de Densidade, com uma assimetria e curtose altas.



**Figura 2. Espaço de Cullen-Frey.**

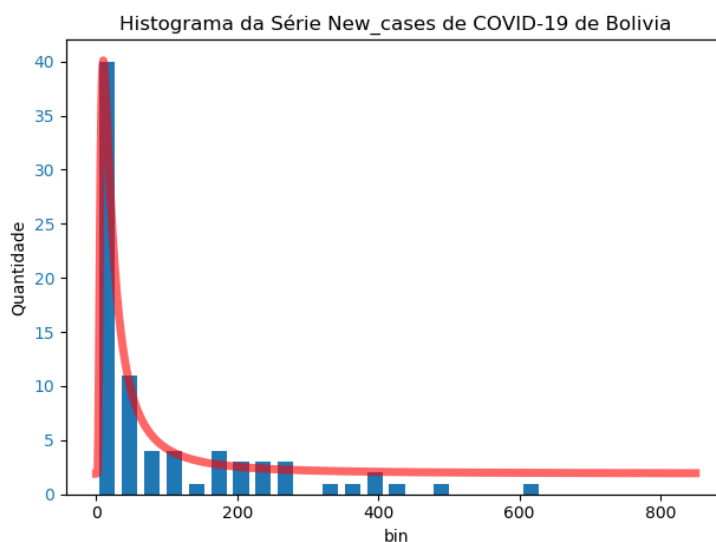
Plot 3 - Ajustar um PDF ao histograma (figura 3).

O Ajuste da PDF pode ser observado na figura 3. Foram utilizados os parâmetros abaixo, da função `genextreme.pdf`, do pacote `scipy.stats`, para gerar uma PDF para o ajuste:

$c = -1$  ;  $loc = 20$  ;  $scale = 20$ , onde a função PDF é calculada da seguinte forma:

$$f(x, c) = \begin{cases} \exp(-\exp(-x)) \exp(-x) & \text{for } c = 0 \\ \exp(-(1 - cx)^{1/c}) (1 - cx)^{1/c-1} & \text{for } x \leq 1/c, c > 0 \end{cases}$$

A densidade de probabilidade acima é definida na forma "padronizada". Para alterar o posicionamento do gráfico no eixo x ou escalar a distribuição, os parâmetros `loc` e `scale` devem ser alterados.



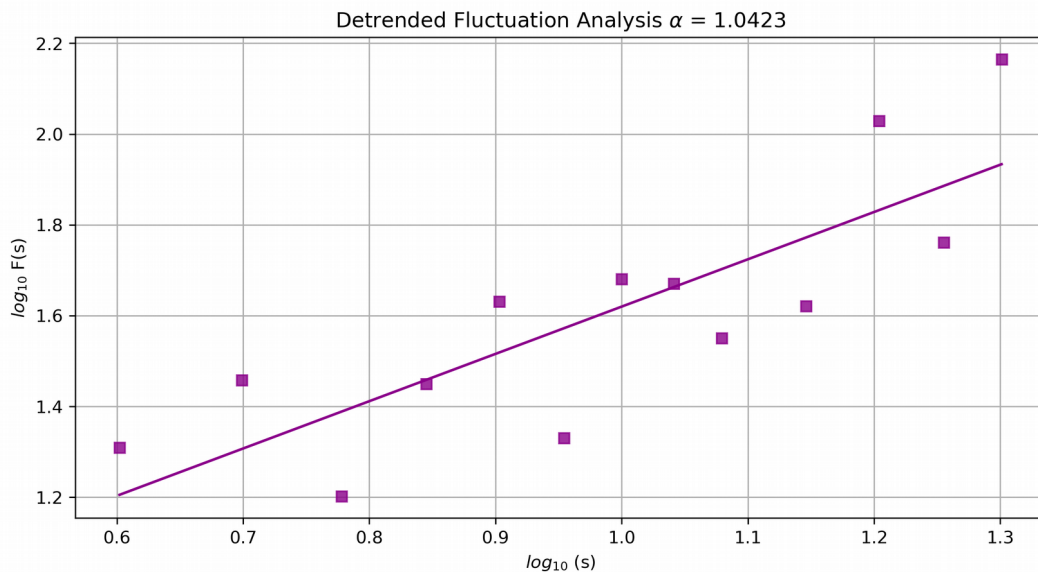
**Figura 3. Ajuste de um PDF ao histograma.**

#### Plot 4 - Índice espectral $\alpha$ via DFA (figura 4).

O valor de  $\alpha$  calculado através da DFA é de 1,0423, gerado através da função `dfa1d(time_series, grau)` do programa SpecPlus, que está na pasta “tools”. A figura 4 exibe o Plot da DFA.

O valor de  $\alpha$ , que determina a auto-afinidade do sinal, é calculado através da utilização de uma função RMSE (Raiz quadrada do erro médio) sobre o sinal. O vetor resultante é colocado em um gráfico  $\log_{10}$ , onde uma reta é traçada interpolando os pontos, e o valor de alfa representa a inclinação desta reta.

Quando o valor de  $\alpha$  é maior que 0,5, indica que há correlação a longo prazo entre os valores da série.



**Figura 4. Índice espectral  $\alpha$  via DFA**

O cálculo do  $\beta$  teórico, via valor de  $\alpha$  (1,0423), é dado por:

$$\beta = 2 * \alpha - 1$$

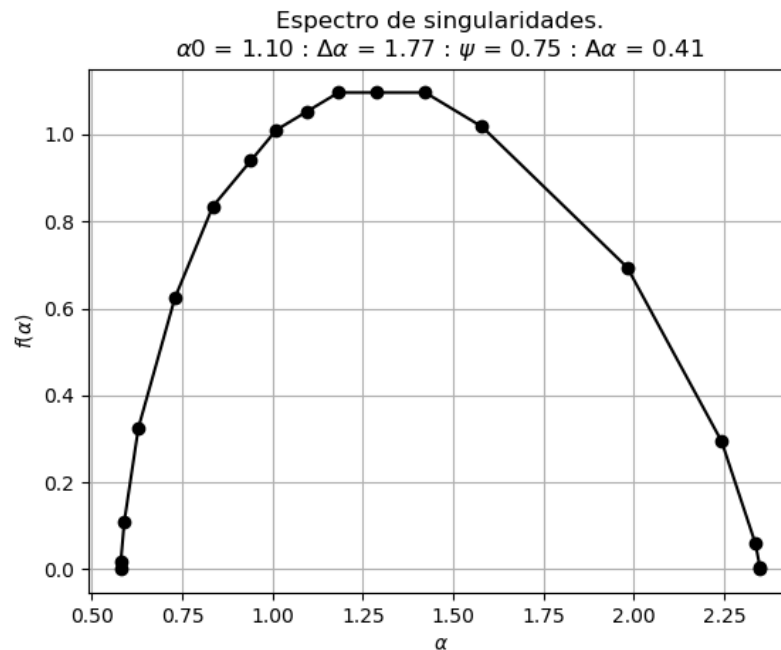
$$\beta = 1,0846$$

O valor de  $\beta$  aproxima-se de 1, indicando um ruído rosa, onde a densidade espectral de potência é inversamente proporcional à frequência do sinal.

#### Plot 5 - Obter o espectro de singularidades $f(\alpha) \times \alpha$ (figura 5).

A figura 5 exibe o espectro de singularidades, onde se pode verificar que  $\Delta_\alpha = 1,7710$ , bem acima de 0,4, indicando aspectos de multifractalidade. Também é possível observar que não aparecem muitos atratores, o que pode se explicar pela pequena quantidade de pontos da série.

A média entre os valores de  $\alpha_{\min}$  e  $\alpha_{\max}$  é 1,46, e  $\alpha_0 = 1,09$ , de onde observamos que a assimetria  $A_\alpha < 1$ , indicando a dominância de flutuações de baixa amplitude.



**Figura 5. Espectro de singularidades**

Os valores de  $\alpha$ ,  $\beta$ ,  $\Delta\alpha$ ,  $\alpha_0$  e  $A\alpha$  podem ser observados na tabela 1.

**Tabela 1. Valores de  $\alpha$ ,  $\beta$ ,  $\Delta\alpha$ ,  $\alpha_0$  e  $A\alpha$**

$\alpha$	$\beta$	$\Delta\alpha$	$\alpha_0$	$A\alpha$
1,04	1,08	1,77	1,10	0,41

## Branch Direito

Foram implementadas as equações do modelo em no arquivo `prova_modelo_covid.py`, com diversas opções de configuração, documentadas no código. Dentre as principais:

- `prob_agent`: (Dicionario: {String: Array de Floats}): Dicionário para configuração dos espectros de probabilidade. Ex: {'Espectro 1': [0.5, 0.45, 0.05], 'Espectro 2': [0.7, 0.25, 0.05]}
- `fator_n_min`: (Array de Floats) Fatores "n" mínimos. Ex = [1.0, 3.0, 5.0]
- `fator_n_max`: (Array de Floats) Fatores "n" máximos. Ex = [2.0, 4.0, 6.0]
- `coluna_agrupadora_covid`: (String) = nome da coluna para agrupar a série. Ex. 'location'
- `coluna_serie_covid`: (String) Nome da coluna da série, ex: 'new\_cases'
- - `data_inicial`: (String) Data inicial da serie. Ex: '2020-05-21'
- - `num_dias_para_media`: (Inteiro) Número de dias de média a ser considerada para inicialização do modelo
- - `data_inicial_previsao`: (String) Data Inicial de previsão. Ex: '2020-05-28'
- - `data_final`: (String) Data final de dados reais e da previsão. Ex: '2020-06-17'
- - `is_atualizar_arquivo_covid`: (Boolean) True para atualizar o arquivo da covid da url do parâmetro
- `estrategia_g`: (String) Estrategia a ser usada para cálculo de g na previsão

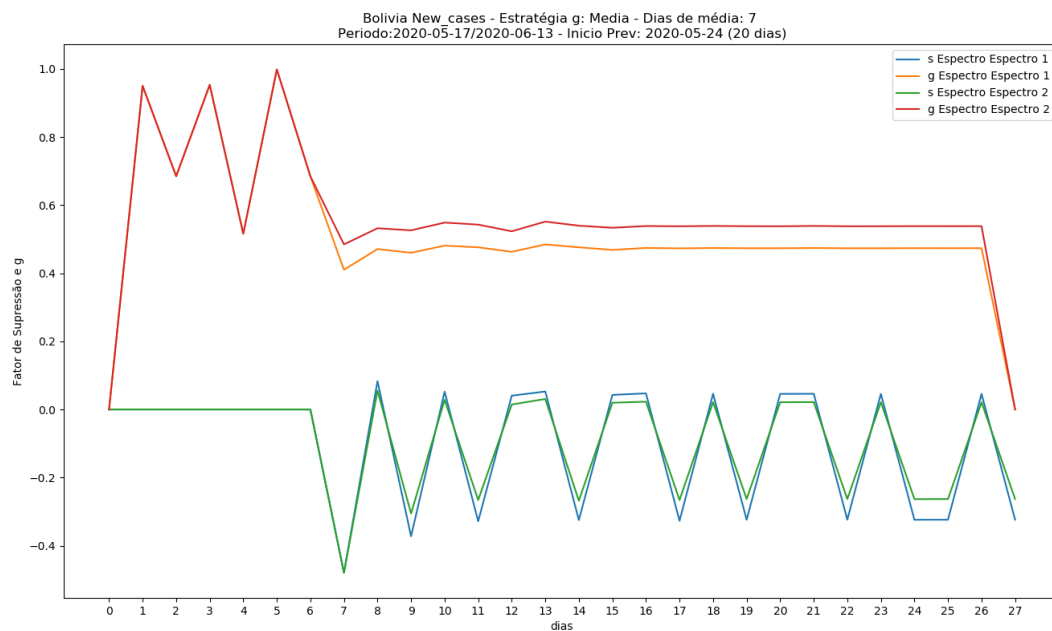
Os parâmetros de entrada foram são configurados em uma instância da classe `DadosEntrada`. Foram consideradas os seguintes parâmetros para a execução do modelo:

```

valor_coluna_agrupador = 'Bolivia'
data_inicial = '2020-05-17'
data_final = '2020-06-13'
data_inicial_previsao = '2020-05-24'
fator_n_min = [1.0, 3.0, 5.0]
fator_n_max = [2.0, 4.0, 6.0]
prob_agent = {'Espectro 1': [0.5, 0.45, 0.05], 'Espectro 2': [0.7, 0.25, 0.05]}
num_dias_para_media = 7
estrategia_g = 'Media' # estrategia do modelo da prova

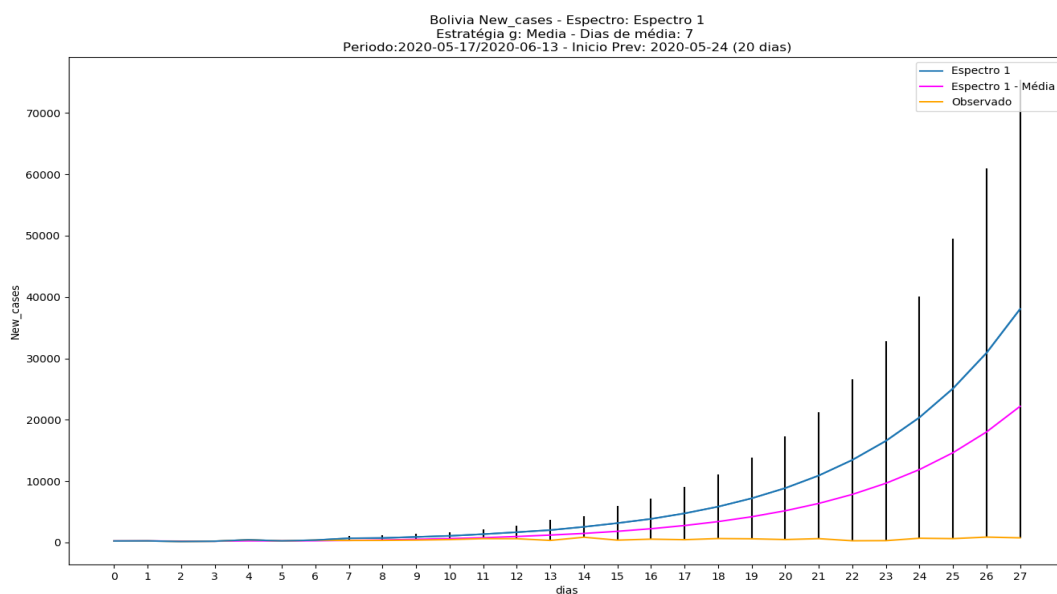
```

A figura 6 exibe o resultado solicitado: Plot do fator de supressão “s” e do fator “g” para os dois espectros, onde o “Espectro 1” é (0,5 ; 0,45 ; 0,05) e “Espectro 2” é (0,7 ; 0,25 ; 0,05).

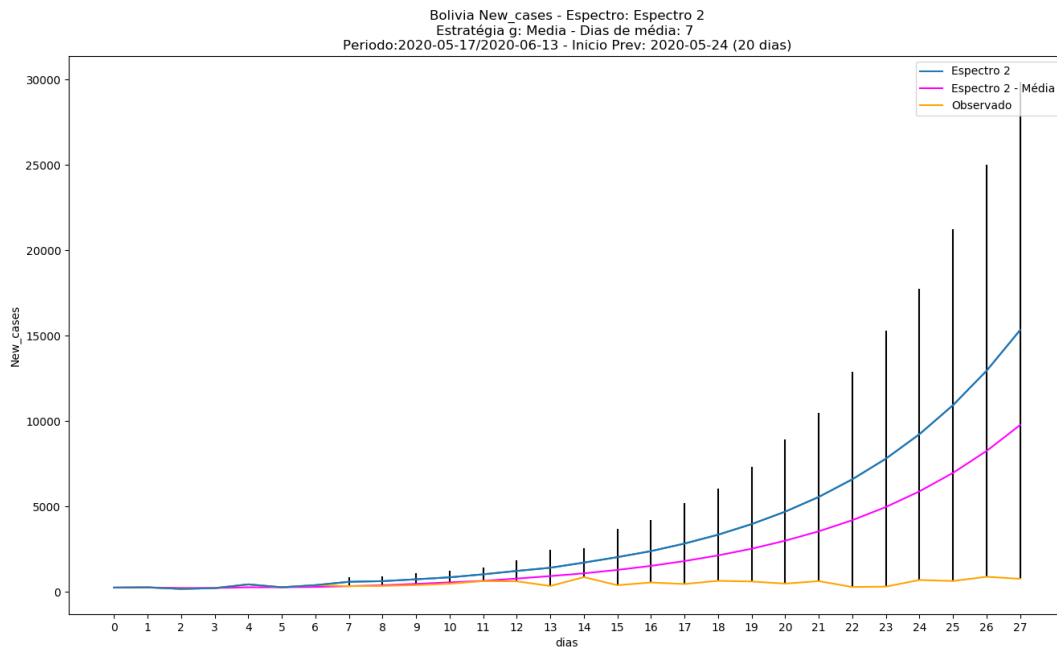


**Figura 6. Fator de supressão “s” e fator “g” para os dois espectros**

Adicionalmente, foram plotados os valores previstos dos dois espectros, observados nas figuras 7 e 8, com barra de erro, juntamente com as médias móveis de 7 dias e o valor observado.

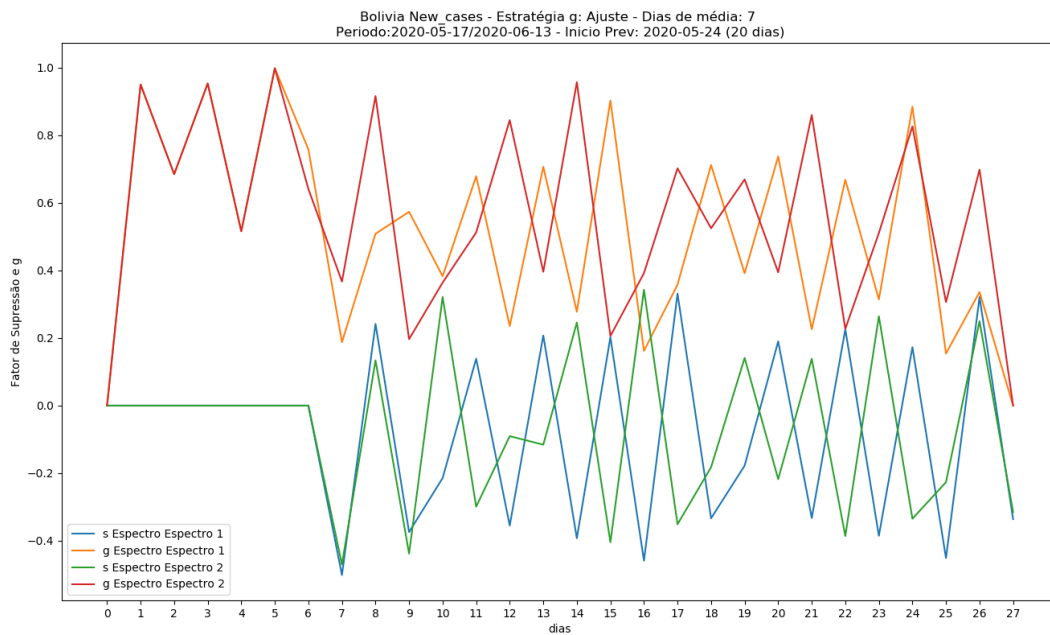


**Figura 7. Previsão para o espectro 1.**



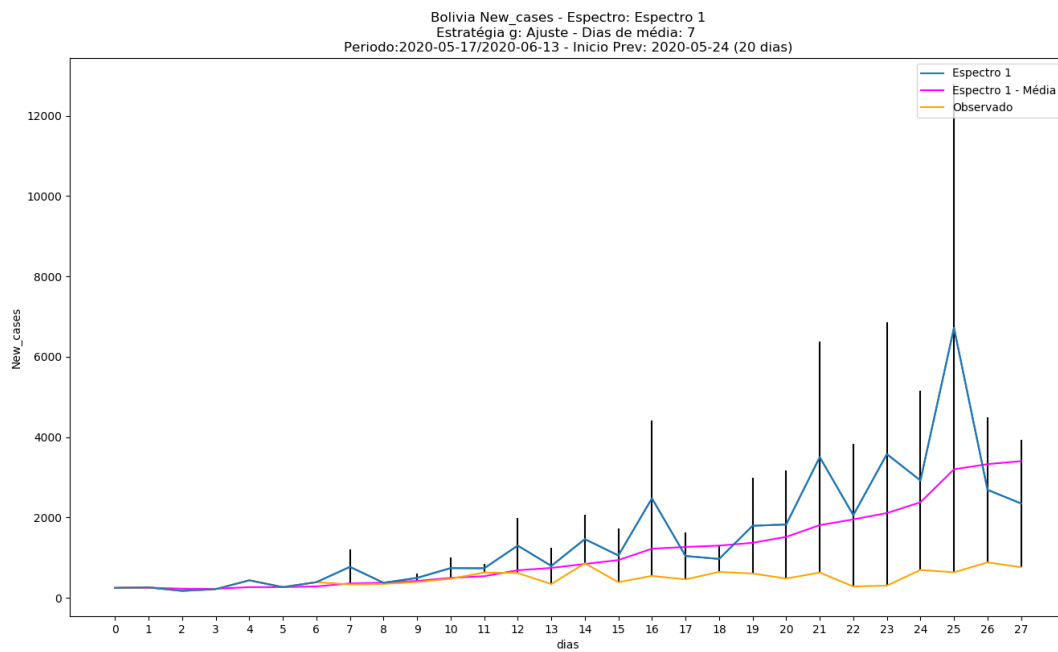
**Figura 8. Previsão para o espectro 2.**

Também foi realizada uma execução com uma estratégia diferente de cálculo de  $g$ , denominada “Ajuste” no modelo, de onde foram geradas as figuras 9, 10 e 11. Essa estratégia consiste em calcular o valor de  $g$  utilizando o valor previsto de amanhã, ao invés de utilizar o valor de hoje. Na rodada de previsão (dia de hoje), o valor de  $g_0$  é utilizado para calcular o dia de amanhã. As figuras 9, 10 e 11 abaixo foram geradas com o  $g_0 = 0.2$ .

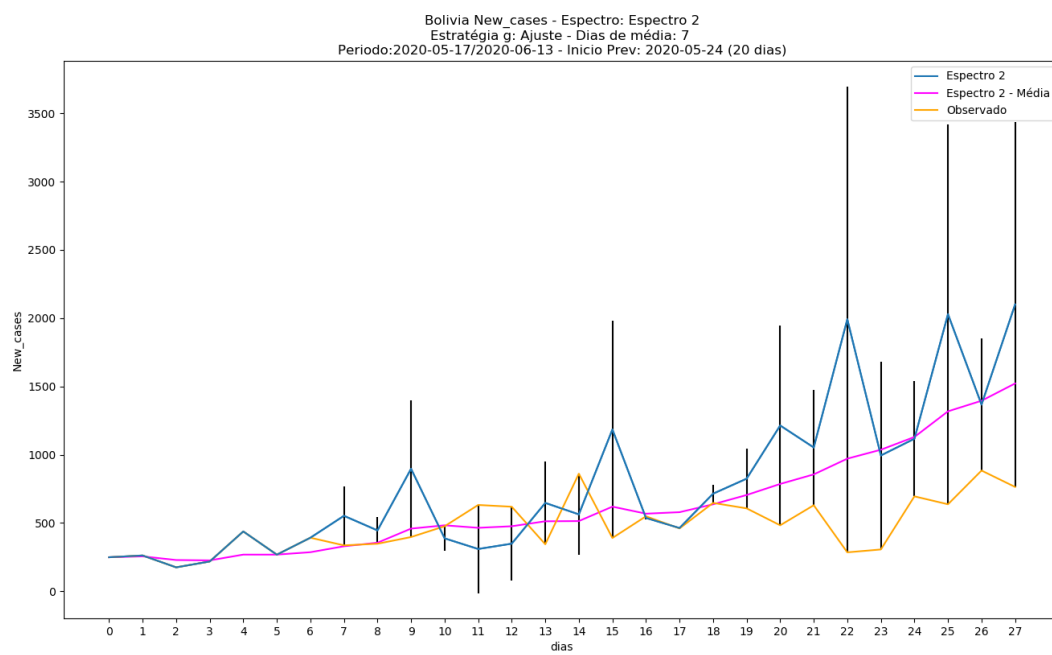


**Figura 9. Fator de supressão “s” e fator “g” para os dois espectros, utilizando a estratégia de g “Ajuste”**





**Figura 10. Previsão para o espectro 1, utilizando a estratégia de g “Ajuste”**



**Figura 11. Previsão para o espectro 2, utilizando a estratégia de g “Ajuste”**