



## CORONAVÍRUS: MODELOS E CENÁRIOS DE UMA PANDEMIA

**Ricardo Borges Gama Neto**

*Essentially, all models are wrong, some are usefull*

*George Box*

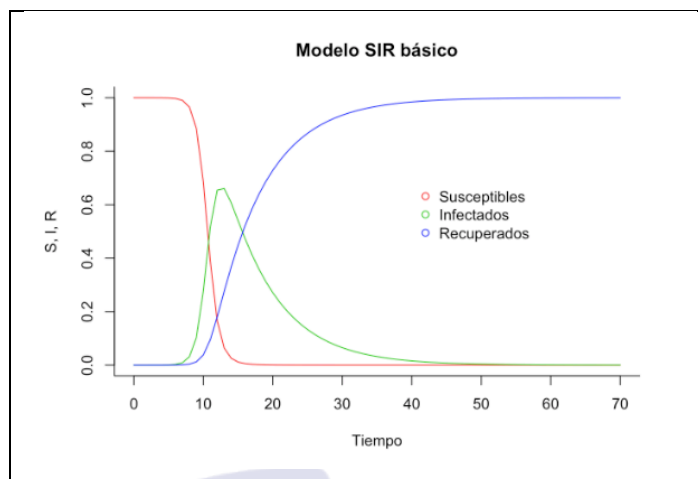
Em 1976, o estatístico britânico George Box escreveu um dos mais famosos aforismos da ciência, no artigo *Science and Statistics*, publicado no *Journal of American Statistic Association*. Sua discussão era sobre o desenvolvimento de modelos parcimoniosos que, ao mesmo tempo, deveriam ser efetivos na estatística experimental<sup>1</sup>. Este aforismo é especialmente importante em tempos de, Covid-19, onde dezenas de estudos apresentam diferentes estimativas de contágios e mortos. Modelos em análise de dados podem ser entendidos, num sentido restrito, como projeções estatísticas (estimativas) baseadas em determinadas suposições ou informações existentes a priori.

A maioria dos modelos de análise, que existem em epidemiologia e na biologia para o entendimento de uma pandemia,<sup>2</sup> se baseiam no modelo SIR (Susceptível, Infectado, Recuperado), elaborado por Kermack e McKendrick (1927), a quase 100 anos. Este se fundamenta na dinâmica da população e equações não-lineares. Cada tipo de população é entendida como um compartimento que variam de acordo com o tempo. O crescimento de uma epidemia ocorre de forma exponencial.

Gráfico 1 – Modelo SIR Básico

<sup>1</sup> <http://www-sop.inria.fr/members/Ian.Jermyn/philosophy/writings/Boxonmaths.pdf>. Acessado em 09 de maio de 2020.

<sup>2</sup> Para melhor entendimento, deve-se definir três conceitos centrais: endemia, epidemia e pandemia. “Endemia é definida como a presença habitual de uma doença em uma determinada área geográfica. Pode-se, também, referir à ocorrência dentro de uma área. Epidemia é definida como a ocorrência em uma comunidade ou região de um grupo de doenças de natureza similar, excedendo, a expectativa norma, derivada de uma fonte comum de propagação. Pandemia refere-se a uma epidemia de dimensões mundiais”. (GORDIS, 2017, p. 22-23).



Fonte: <https://rpubs.com/dsfernandez/422937>

As curvas dependem basicamente dos parâmetros, que são diferentes para cada tipo de enfermidade (taxa de propagação, tempo de incubação, recuperação, etc...). Na primeira fase, a de propagação, ocorre um forte crescimento das pessoas contaminadas. Depois surge a fase de estabilização dos contágios e por fim de recrudescimento da doença. A evolução do covid-19 vai seguir a sequência descrita no gráfico 1<sup>3</sup>.

Cada doença infecciosa tem diferentes parâmetros e, no caso do covid-19, estes foram definidos pela Organização Mundial de Saúde (OMS) a partir das informações da epidemia em Wuhan, em fevereiro de 2020<sup>4</sup>. O coronavírus tem uma taxa de contágio de 1 para 3, ou seja, uma pessoa contamina três e essas mais três em escala. A maioria das pessoas não deve ter maiores complicações, contudo, 20% dos infectados precisarão de acompanhamento médico intensivo. A taxa de mortalidade é fortemente condicionada pela faixa etária, sendo residual nos indivíduos com menos de 9 anos e altamente impactante naqueles com mais de 80 anos.

A partir destes parâmetros foram elaborados vários modelos de contágio e mortes. Em 16 de março, um grupo de cientistas do *Imperial College* apresentou previsões catastróficas a respeito do número de mortos no Reino Unido. Mais de 500 mil pessoas morreriam caso o

<sup>3</sup> Há diversos outros modelos de estudo da evolução do coronavírus, como estatística bayesiana e teoria dos jogos < <https://www.agenciasinc.es/Reportajes/Un-modelo-un-teorema-y-teoria-de-juegos-contra-el-coronavirus> > Acessado em 09 de maio de 2020.

<sup>4</sup> Contudo, há diversas dúvidas sobre confiança dos números chineses. Há forte desconfiança de que os números chineses são subnotificados, dados que foram utilizados pela OMS como parâmetro sobre a evolução da covid-19 e, utilizados por todos os pesquisadores do mundo para as primeiras estimativas de contágio e mortos <https://www.publico.pt/2020/03/27/mundo/noticia/china-duvidas-zero-novos-casos-coronavirus-wuhan-1909853> >. Mas não é somente na China onde existem dúvidas sobre o número da pandemia, o governo norte-americano também desconfia das estatísticas do “Irán, Rusia, Indonesia y especialmente a Corea del Norte, que no ha reportado un solo caso de la enfermedad. Otros, como Arabia Saudí y Egipto, también pueden estar minimizando sus números, según apuntan”. < <https://www.redaccionmedica.com/secciones/sanidad-hoy/coronavirus-un-informe-de-la-cia-dice-que-china-oculta-muertos-y-positivos-7050> >. Acessado em 09 de maio de 2020.



governo não promovesse medidas de isolamento social severo e supressão de toda atividade não essencial, seguindo a estratégia chinesa. O modelo do *Imperial College* se baseia em dois cenários possíveis:

Duas estratégias fundamentais são possíveis: (a) mitigação, que se concentra em retardar, mas não necessariamente em impedir a propagação da epidemia - reduzindo o pico de demanda de assistência médica e protegendo aqueles com maior risco de doenças graves contra infecções; e (b) a supressão, que visa reverter o crescimento da epidemia, reduzindo o número de casos para níveis baixos e mantendo essa situação indefinidamente. Cada política tem grandes desafios. Concluímos que políticas ótimas de mitigação (combinando isolamento domiciliar de casos suspeitos, quarentena domiciliar de pessoas que moram na mesma casa, casos suspeitos e distanciamento social de idosos e outras pessoas com maior risco de doença grave) podem reduzir o pico da demanda de assistência médica em 2 / 3 e mortes pela metade. No entanto, a epidemia mitigada decorrente, provavelmente ainda resultaria em centenas de milhares de mortes e sistemas de saúde (principalmente, unidades de terapia intensiva) sendo sobrecarregados, muitas vezes. Para os países capazes de alcançá-lo, isso deixa a supressão como a opção política preferida (Ferguson *et all*, 2020, p. 01).

Nas primeiras estimativas do *Imperial College*, caso a estratégia fosse de mitigação, o número de mortos poderia chegar a 250 mil e se o governo britânico adotasse o caminho da supressão, as mortes deveriam cair para uma faixa entre 9 a 98 mil mortes. No dia 25 de março, em depoimento *online* ao parlamento, Niel Ferguson, epidemiologista e biólogo matemático do *Imperial College*, mudou sua estimativa para 20 mil. A discrepância não passou despercebida e no dia seguinte a equipe de pesquisa da instituição inglesa publicou o *Report 12: The Global Impact of Covid-19 and Strategies for Mitigation and Supression*.<sup>5</sup> O número do coronavírus no mundo, no caso de não se fazer nada, foi estimado em mais 7 bilhões de infectados, quase toda a população do planeta e mais de 40 milhões de mortos. Em caso de supressão severa, 469 milhões de contaminados e 1 milhão e oitocentos mil mortos. No Brasil, a modelagem apresentou cinco cenários possíveis. Iremos apresentar três: a) numa situação de não intervenção – 187.799.806 infectados, um milhão cento e cinquenta e dois mil mortos; b) com distanciamento social aprimorado para idosos – 120.836.850 infectados, setecentos e dois mil e quatrocentos e noventa e sete mortos e c) adoção de medidas de distanciamento social de forma precoce – 11.495.197 infectados, quarenta e quatro mil e duzentos e doze pessoas mortas<sup>6</sup>.

Tabela 01 – Simulação

	População total	Contaminados	Mortos (4%)
--	-----------------	--------------	-------------

<sup>5</sup> <https://www.imperial.ac.uk/media/imperial-college/medicine/mrc-gida/2020-03-26-COVID19-Report-12.pdf> . Acessado em 09 de maio de 2020.

<sup>6</sup> <https://pfarma.com.br/imperial-college-mortes.html> . Acessado em 09 de maio de 2020.



1º Cenário Não se faz nada (80% de contaminados)	100.000.000	80.000.000	3.200.000
2º Cenário Com isolamento moderado (30% de contaminados)	100.000.000	30.000.000	1.200.000
3º Cenário Com isolamento social forte e quarentena radical (1% de contaminados)	100.000.000	1.000.000	40.000

A tabela acima apresenta um modelo simplificado de contaminação com uma taxa de falecimento de 4%, independente da ação de qualquer variável interveniente como estrutura demográfica, espacial, comportamento do vírus (fatores ambientais como temperatura, umidade altitude, etc., meio de contaminação (direta ou indireta) e tempo (GORDS, 2017). A população do nosso exemplo é de cem milhões de pessoas. A primeira linha, demonstra o cenário sem qualquer tipo de restrição ao comportamento de pessoas e tem como resultado mais de três milhões de mortos, a segunda apresenta uma situação do que denominamos, de isolamento moderado (um milhão e duzentos mil mortos) e o terceiro cenário com um isolamento social forte e quarentena radical, 40.000 mortos.

Um outro modelo matemático, de simulação do comportamento do coronavírus (contagem de infectados, curados e mortos), foi elaborado no dia 26 de março, usando dados do Centro de Ciências de Sistemas e Engenharia (CSSE, em inglês) da Universidade de Johns Hopkins, pelos físicos José Dias do Nascimento (UFRN) e Wladimir Lyra (Universidade do Estado do Novo México), em linguagem *Python*<sup>7</sup>. O modelo é formulado a partir do estudo de sistemas dinâmicos e processos não-lineares, usados em astrofísica<sup>8</sup> e segue a lógica do SIR. De acordo com Lyra:

A astrofísica lida bastante com processos não-lineares, e na pesquisa em dinâmica de fluidos já tinha até usado modelos originalmente desenvolvidos para a dinâmica de populações biológicas, como o modelo predador-presa —modelo que foi

<sup>7</sup> <https://www.python.org> . Ver também <https://www.kaggle.com/unanimad/corona-virus-brazil> . Acessados em 09 de maio de 2020.

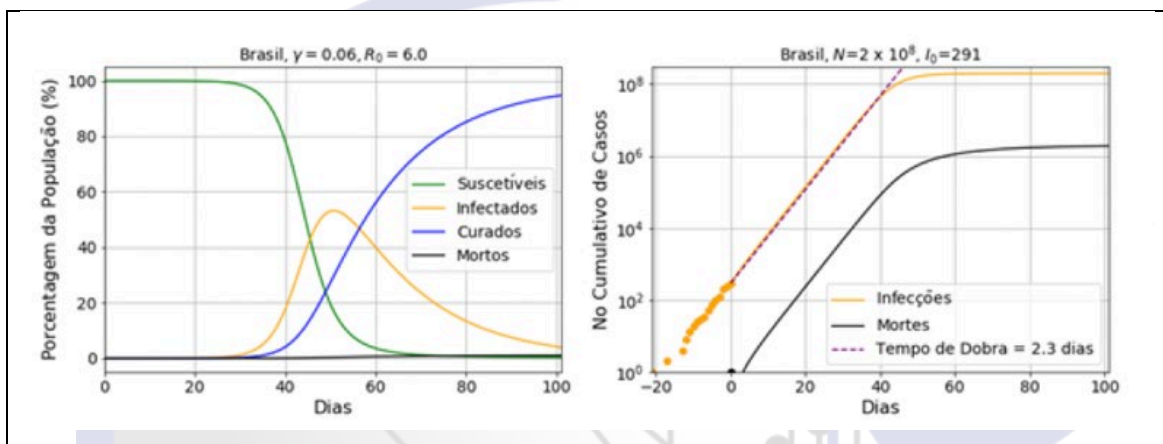
<sup>8</sup> A utilização de modelos da física, para explicar situações complexas não-físicas, não fica restrito às ciências naturais, também são utilizados para explicar questões sociais. Exemplos aplicados a eleições e opinião pública podem ser vistos em: <https://arxiv.org/pdf/1412.4718.pdf> e [https://www.inovacaotecnologica.com.br/noticias/noticia.php?artigo=leis-fisica-explicam-polarizacao-politica-eleicoes&id=010150200131#.XozC\\_S\\_5TaY](https://www.inovacaotecnologica.com.br/noticias/noticia.php?artigo=leis-fisica-explicam-polarizacao-politica-eleicoes&id=010150200131#.XozC_S_5TaY) . Para aplicação da estatística, utilizada na física à ciência política, ver <https://arxiv.org/pdf/cond-mat/0004306.pdf> . Teoria do caos aplicada ao estudo de questões políticas Banerjee; Erçetin e Tekin (2014). Sites acessados em 09 de maio de 2020.



desenvolvido em biologia-matemática para estudar as populações de predadores e presas em um ambiente<sup>9</sup>

O modelo de Lyra e Nascimento, aplicado ao Brasil, defende que cada indivíduo infectado está em média contaminando outros seis, o dobro que estima a OMS. Esta taxa de contágio provoca a duplicação do número de casos, a cada dois ou três dias. O pico da pandemia no Brasil, sem qualquer medida de isolamento social, deverá acontecer com 50 dias e metade da população do país contaminada. Como no modelo do *Imperial College* acreditam no colapso do sistema de saúde, cujo resultado seria dois milhões de mortos.

Gráfico 2 – Predição total de número de mortos no Brasil



Fonte: <https://www.tecmundo.com.br/ciencia/151189-quantos-morrer-coronavirus-brasil-cientistas-respondem.htm>

As estimativas do *Imperial College* e de Lyra e Nascimento<sup>10</sup> divergem bastante. Qual deles devemos considerar? Resposta: os dois. Modelos são aproximações da realidade e nunca serão a realidade em si. Servem como mapas heurísticos de análise, servem para que os pesquisadores e tomadores de decisão adquiram consciência de quais são as possibilidades de cenários futuros, a partir das condições iniciais, e os possíveis processos de evolução da pandemia do Covid-19. Descrever realmente como ocorre um processo complexo, como é o de contaminação em uma pandemia, *ex-ante*, é impossível. São escolhidos os parâmetros e as variáveis mais importantes e com eles se desenvolvem equações que descrevem o comportamento do vírus, pois erros de estimação sempre existem. Por mais erros que possuam, a existência de modelos de predição é fundamental, sem eles o estabelecimento de políticas públicas e de estratégias de enfrentamento a pandemia seriam impossíveis.

<sup>9</sup> <http://www.apufsc.org.br/2020/03/31/como-modelo-cientifico-usado-na-astrofisica-preve-casos-de-covid-19-no-pais/> . Acessado em 09 de maio de 2020.

<sup>10</sup> <https://covid.lais.ufrn.br/#projecoes> . Acessado em 09 de maio de 2020.



Todos os modelos possuem limitações. Àqueles são dependentes tanto do seu próprio desenho, que tem haver com sua lógica interna, quanto das incertezas relacionadas com os dados utilizados para sua elaboração. A primeira destas é o comportamento dos microrganismos em diferentes meio-ambientes. Nunca se sabe exatamente como um vírus novo vai se comportar em função da temperatura, umidade, altitude dos diversos espaços geográficos (especialmente em situações de contaminação indireta) (GORDIS, 2017). Outro problema importante é o tempo de duração da infecção. O tempo de duração do covid-19 ainda é uma questão importante em aberto, inclusive o de latência. O *Imperial College* trabalhou com dois períodos de incubação até o aparecimento da doença, três e sete dias. Os médicos chineses em Wuham e a OMS defendem 15 dias<sup>11</sup>. Quanto mais distante é o tempo proposto na estimativa da evolução de uma pandemia, mais difíceis e menos confiáveis elas serão, da mesma forma que acontece com as previsões meteorológicas. Se a elaboração de estimativas de longo prazo é um problema, quanto maior o tempo mais padrões surgem e estes são introduzidos nos modelos, corrigindo as falhas iniciais e mais informações efetivas os tomadores de decisões políticas possuem.

Em todos os modelos elaborados sobre a taxa de falecidos como resultado do contágio com o covid-19 os números são superlativos, de milhares a milhões, caso os governos não estabelecerem medidas severas de isolamento social e quarentena. O colapso dos sistemas de saúde é outra previsão comum a todos os modelos. Já isto é uma grande contribuição, apesar das enormes diferenças de estimativa.

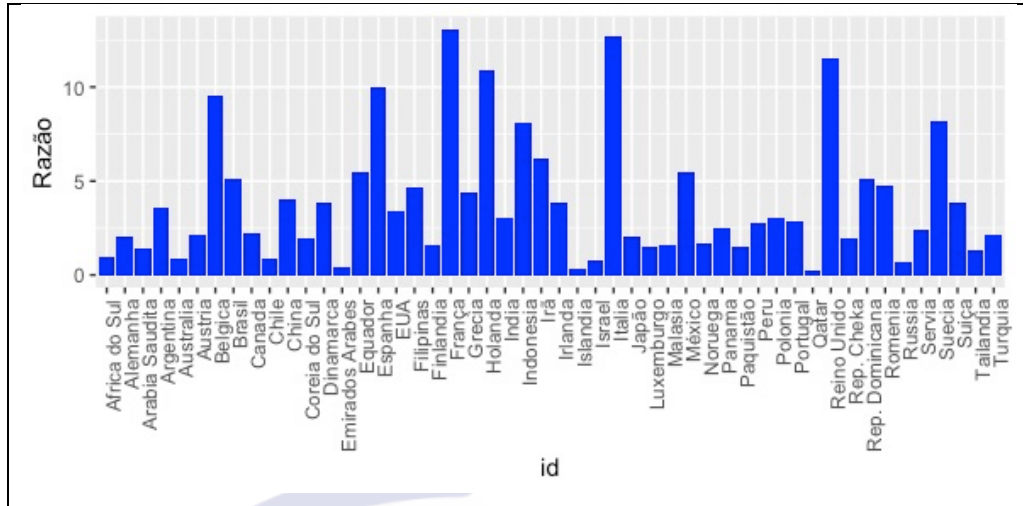
A pandemia do covid-19 tem apresentado dinâmicas extremamente diversas. Podemos afirmar que variáveis demográficas, espaciais, de comportamento social, estratégias de ação de governos e capacidade dos sistemas de saúde de tratar dos doentes são algumas das principais variáveis intervenientes na relação entre contagiados e mortos. A variação na taxa de infectados *versus* falecidos é imensa. Os dados apresentados agora foram retirados do site *Coronavirus Resource Center*<sup>12</sup>, da Universidade John Hopkins, no dia 09 de abril. O total de infectados está contabilizado em 1.484.811 e 88.538 mortos.

### Gráfico 3 – Razão entre Contaminados e Mortos

---

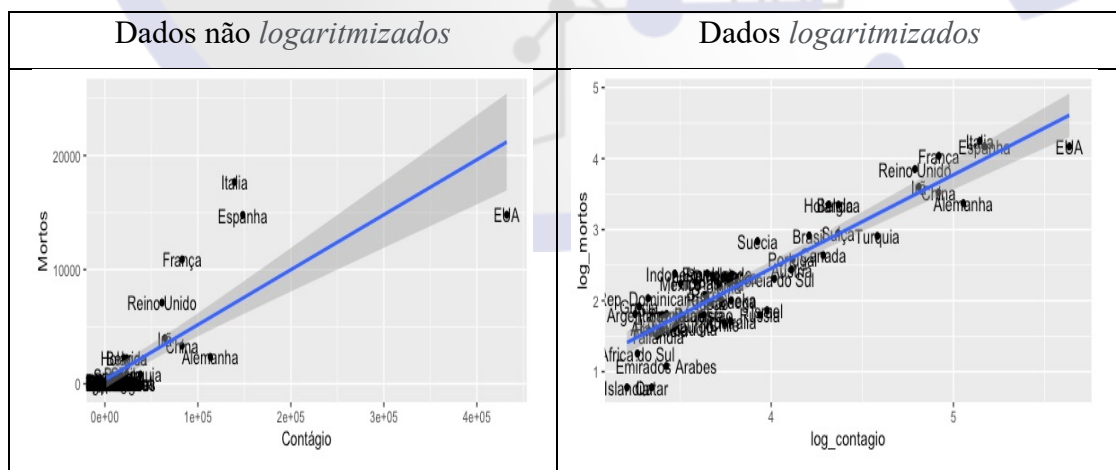
<sup>11</sup> David Uip, coordenador do Centro de Contingência do Coronavírus de São Paulo, afirmou em entrevista que: “O que nós imaginávamos, que o período de incubação ia até 14 dias, nós estamos vendo que o período de incubação [do covid-19] é mais curto. Vai de 3 a 8 dias. Nós vamos sugerir hoje ao Ministério da Saúde que mude o critério de tempo da quarentena, que diminua de 14 para dez dias”. < <https://noticias.uol.com.br/saude/ultimas-noticias/redacao/2020/03/17/tempo-de-incubacao-do-covid-19-e-menor-e-sp-sugere-reduzir-quarentena.htm>. Acessado em 09 de maio de 2020.

<sup>12</sup> <https://coronavirus.jhu.edu/map.html>. Acessado em 09 de maio de 2020.



O gráfico 3<sup>13</sup> foi construído com uma amostra dos 50 países com maior número de pessoas infectadas, abrangendo um total de 97,1% do total de casos. Ele demonstra que há uma variação gigantesca entre o número de contagiados e mortos por países. Itália, França, Espanha, Reino Unido, Holanda e Bélgica têm taxas acima de 10%. Outros países como Islândia, Qatar, Austrália, Rússia, Israel e Chile têm números inferiores a 1%. A média é de 3,90% com um desvio padrão de 3,3. Países europeus, com maior número de idosos em relação ao total da população, parecem ter sido mais fortemente impactados pela mortalidade do covid-19. Contudo, como demonstram os casos da Alemanha (2,07%) e da Áustria (2,11%), ter um grande número de idosos é um fator importante, mas não suficiente para que sejam explicados números tão díspares.

Gráfico 4 – Scatterplots Contaminados vs Mortos



<sup>13</sup> Os gráficos 3 e 4 e as regressões foram feitas no software R versão 3.6.1, utilizando as bibliotecas ggplot2 e AER (Applied Econometric in R).



O gráfico 4 apresenta dois *scatterplots* dos números de contagiados pelo de mortos. No primeiro, os dados estão não-logaritimizados e no segundo estão. Qual a razão deste procedimento? A mais óbvia é que podemos fazer isto, as duas variáveis têm valores maiores que 0. Há sempre diversas razões para utilizarmos o procedimento de re-expressão de dados, alguns bons outros maus, mas como menciona muito bem Wooldrige (2010), nenhuma delas é escrita em pedra. De forma geral, podemos dizer que há três boas razões: a) obter linearidade nos parâmetros, mesmo que as variáveis não sejam lineares; b) resolver problemas de heterocedasticidade e c) quando a teoria sugere que se expresse os dados em forma de logs, como grandes números inteiros (ex: população). Por outro lado, um grande problema é que ele comprime os dados de uma tal forma que esconde *outliers*. No gráfico à esquerda, comparado com o da direita, podemos ver isto facilmente no caso dos EUA, Itália, Espanha, França e Reino Unido.

Logaritimizar resolve o problema da heterocedasticidade<sup>14</sup> (que é um crime muito maior do que usar variáveis que não se distribuem normalmente) e nos permite medir o impacto do contágio sobre o número de mortos. Uma regressão de mínimos quadrados ordinários obteve um  $R^2$  ajustado = 0,83. Podemos inferir que 83% das mortes podem ser explicadas, unicamente, porque os indivíduos foram contaminados. O que demonstra que a estratégia de isolamento social com quarentena é fundamental no processo de superação desta pandemia.

Como dissemos, os modelos de pandemias têm o problema da estimativa temporal. Quanto mais tentam explicar para frente, mais incertos serão. Por este fato, estes modelos são corrigidos a cada novo conjunto de informações que vão sendo disponibilizadas, assim os tomadores de decisão poderão escolher as políticas públicas mais adequadas a cada momento. Há vários instrumentos sendo utilizados para descrever a evolução da pandemia, cada um com sua finalidade, gráficos para demonstrar a curva de evolução da doença, cálculo diferencial para medir a velocidade do crescimento dos contágios e média da curva de Forman-Ricci para perceber mudanças nos comportamento das fases da pandemia<sup>15</sup>.

## Referências Bibliográficas

---

<sup>14</sup> Que foi muito forte num teste de regressão dos dados não-logaritimizados, tanto na análise gráfica quanto no teste Breush-Pagan (bptest).

<sup>15</sup> <https://www.irrd.org/wp-content/uploads/2020/04/covid-19-dash.html#curvatura-forman-ricci> e <https://www.medrxiv.org/content/10.1101/2020.04.01.20047225v1.full.pdf>. Acessado em 09 de maio de 2020.





Rede CTIDC

Pró-Defesa IV - Ciência, Tecnologia e Inovação em Defesa:  
Cibernética e Defesa Nacional

BANERJEE, Santo; ERÇETIN, Sefika e TEKIN, Ali (ed). (2014), *Chaos, Theory in Politics: understanding Complex Systems*. New York: Springer.

BOX, George (1976), "Science and Statistics", *Journal of American Statistic Association*, vol. 71, nº 356, pp. 791-799.

GORDIS, Leon (2017), *Epidemiologia*. 5ª ed. Rio de Janeiro: Thieme Revinter.

KERMACK, Willian O. e MCKENDRICK, Anderson G. (1927), "A Contribution to the Mathematical Theory of Epidemics". *Proceeding of The Royal Society of London. Series A*, Vol. 115, pp. 700-721.

WOODRIDGE, Jeffrey. (2010), *Introducción a la Econometría*. Cidade do México (DF): Cengage Learning Editores.

