

Proyección sobre la evolución de la incidencia del virus COVID-19

Motivación

Ante la situación de alarma actual en relación al denominado COVID-19, se han puesto en marcha diferentes iniciativas que tratan de anticipar algunos de los efectos negativos que la pandemia actual genera, de esta manera es posible planificar diferentes escenarios y tomar decisiones con el objetivo de paliar las consecuencias más negativas de esta enfermedad. Así, el CEMat se encarga de coordinar las iniciativas de la comunidad matemática española relacionadas con la crisis creada por el COVID-19. Así, un grupo amplio de investigadores trata de encontrar una respuesta al problema actual de salud pública mencionado.

En particular, desde este grupo de investigación de la Facultat d'Economia de la Universitat de València, también tratamos de realizar una pequeña contribución, teniendo en cuenta las dificultades que el problema analizado presenta y que, por tanto, limita la respuesta de los diferentes modelos matemáticos que se utilizan.

Introducción Metodológica

Existen diferentes maneras de afrontar este tipo de problemas: (i) con modelos predictivos, haciendo uso de técnicas estadísticas como los denominados GLM (Modelo Lineal Generalizado); (ii) con modelos epidemiológicos establecidos, como los denominados SIR (Susceptibles-Infectados-Recuperados) y otros modelos derivados de este; y, (iii) técnicas de Series Temporales, en las que se analiza una o más variables de interés y se establece una relación estructural de evolución temporal que se asume 'persistente' en el tiempo. (iv) Por supuesto, existen otros enfoques pero he pasado a describir los más utilizados actualmente.

El grupo de investigación está trabajando en los tres tipos mencionados (i) a (iii), no obstante este documento se centra en el tercer tipo, modelización de alguna de las variables como serie temporal, por su gran interés de cara a realizar acciones organizativas si fuera el caso.

Metodología utilizada

En relación a la metodología utilizada, esta es ampliamente conocida en el ámbito de las series temporales. Se ha seguido el siguiente esquema:

- (1) Se consideran Tasas de Variación entre valores diarios consecutivos.
- (2) Para estimar la tendencia se inicia el proceso mediante una media geométrica móvil.

- (3) La estimación de la tendencia se realiza mediante un ajuste funcional de la serie obtenida en el paso anterior. En este caso logarítmico.
- (4) Se utiliza la estimación anterior para predecir los valores en momentos futuros.
- (5) Con los valores de tendencia y los valores observados, mediante encadenamiento, se obtienen los valores que conforman la predicción inicial.
- (6) La predicción inicial obtenida en el paso anterior es re-estimada mediante una suavización temporal (*modificación metodológica respecto del informe del 23 de marzo*)
- (7) Se obtiene el error de la predicción a posteriori, es decir, una vez observado el verdadero valor.
- (8) La estimación ‘a priori’ del error se basa en el que se introduce mediante el ajuste funcional (no aparece en este documento).
- (9) La generación de escenarios alternativos y plausibles se está diseñando y ajustando para que pueda ser utilizado en la práctica (no aparece en este documento).

Frecuencia de actualización del análisis

Los resultados se amplían diariamente con los valores observados y se recalibran los parámetros del modelo y los ajustes utilizados.

Horizonte de predicción y variables analizadas

Las estimaciones son útiles en el corto plazo (1-4 días), las variables analizadas son el número de fallecidos, y el número de ingresos en UCI (no aparece en este documento), para el total acumulado en el conjunto del territorio nacional.

Datos y Fuentes de Información

Los datos utilizados son los publicados por el Gobierno Español, aunque en la fase inicial se utilizaron datos recopilados y depurados por el grupo de trabajo *Datadista* (Github) y datos proporcionados por el *Johns Hopkins CSSE*.

Principales Resultados obtenidos

En esta sección se presentan: (i) una tabla numérica con las diferentes predicciones realizadas en días previos, utilizando la misma metodología que en la fase de validación- no de calibración (Tabla 1) y algunos de los errores estimados (Tabla 2); y, (ii) la evolución de los valores observados de fallecimientos acumulados, junto con la predicción realizada (Figura 1); el ajuste sobre la tendencia de las tasas de variación (Figura 2); así como una comparación entre las predicciones realizadas en días previos (Figura 3).

Resultados Numéricos obtenidos**Tabla 1 (Predicción con la metodología utilizada hasta el 23 de marzo)**

Fecha	Pred 20 marzo	pred 21 marzo	Pred 22 marzo	pred 23 marzo	pred 24 marzo
03-mar	1	1	1	1	1
04-mar	3	3	3	3	3
05-mar	5	5	5	5	5
06-mar	16	16	16	16	16
07-mar	35	35	35	35	35
08-mar	47	47	47	47	47
09-mar	84	84	84	84	84
10-mar	120	120	120	120	120
11-mar	136	136	136	136	136
12-mar	288	288	288	288	288
13-mar	309	309	309	309	309
16-mar	491	491	491	491	491
17-mar	598	598	598	598	598
18-mar	767	767	767	767	767
19-mar	1002	1002	1002	1002	1002
20-mar	1345	1326	1326	1326	1326
21-mar	1694	1697	1720	1720	1720
22-mar	2092	2099	2146	2182	2182
23-mar	2534	2547	2616	2670	2696
24-mar	3015	3036	3132	3213	3323
25-mar	3526	3558	3690	3805	3967
26-mar	4058	4103	4278	4436	4666
27-mar	4597	4659	4885	5097	5410
28-mar	5132	5213	5500	5557	5972



Figura 1

Errores Relativos (%) por predicción y horizonte.

Tabla 2

Error Relativo	Pred 20 marzo	pred 21 marzo	Pred 22 marzo	Pred 23 marzo	Pred 24 marzo
a +1 día	1.4%	-1.33%	-1.65%	-0.96%	-
a 2 días	-1.49%	-3.81%	-2.99%		
a 3 días	-4.12%	-5.53%			
A 4 días	-6.0%				

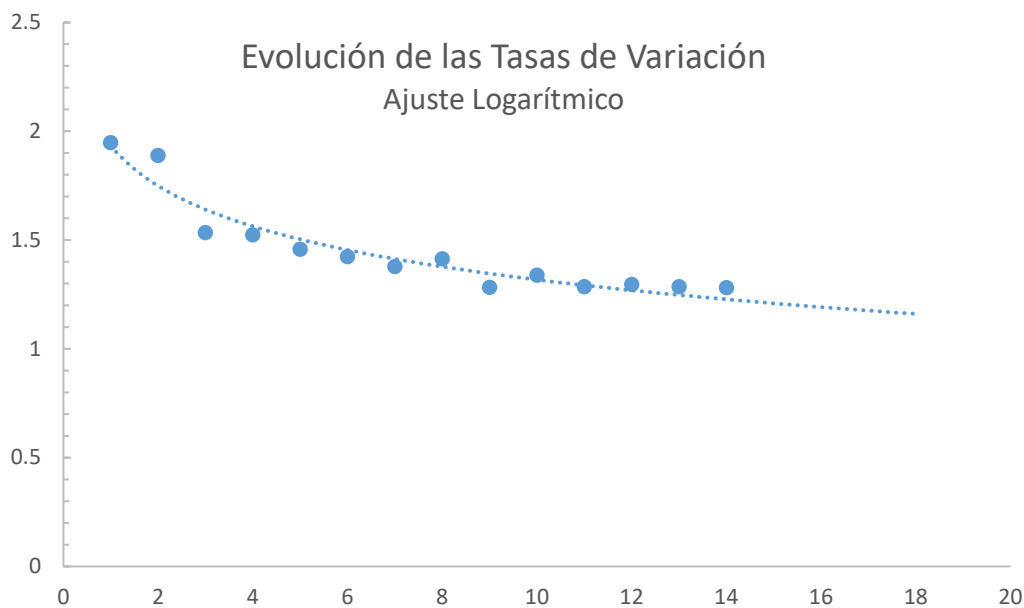


Figura 2



Figura 3

ANEXO

La Tabla A1 muestra la predicción con la metodología utilizada hasta el 23 de marzo. Sin el cambio metodológico introducido en este documento.

Tabla A1

Fecha	Pred 20 marzo	pred 21 marzo	Pred 22 marzo	pred 23 marzo	
03-mar	1	1	1	1	1
04-mar	3	3	3	3	3
05-mar	5	5	5	5	5
06-mar	16	16	16	16	16
07-mar	35	35	35	35	35
08-mar	47	47	47	47	47
09-mar	84	84	84	84	84
10-mar	120	120	120	120	120
11-mar	136	136	136	136	136
12-mar	288	288	288	288	288
13-mar	309	309	309	309	309
16-mar	491	491	491	491	491
17-mar	598	598	598	598	598
18-mar	767	767	767	767	767
19-mar	1002	1002	1002	1002	1002
20-mar	1345	1326	1326	1326	1326
21-mar	1699	1700	1720	1720	1720
22-mar	2105	2106	2240	2182	2182
23-mar	2559	2560	2753	2834	2696
24-mar	3056	3057	3326	3457	3453
25-mar	3588	3589	3953	4151	4175
26-mar	4214	4147	4628	4912	4976

La Tabla A2 muestra los Errores Relativos (%) por predicción y horizonte con la metodología utilizada hasta el 23 de marzo. Sin el cambio metodológico introducido en este documento.

Tabla A2

Error Relativo	Pred 20 marzo	pred 21 marzo	Pred 22 marzo	Pred 23 marzo
a +1 día	1.4%	-1.2%	2.7%	5.1%
a 2 días	-1.2%	-3.5%	2.1%	
a 3 días	-3.5%	-5.1%		
	-5.1%			

24 de marzo de 2020