



PREDICCIÓN DE CARACTERÍSTICAS COMPLEJAS EN LOS TIEMPOS DE COVID 19¹

María Dolores Benavente

En nombre de la Academia Nacional de Economía, les doy las buenas tardes a todos; sabemos que están desde diferentes puntos del planeta lo cual nos pone muy felices, son las maravillas de la tecnología.

Agradezco al panelista y a la comentarista por su generosidad, así como a los miembros de las Academias amigas: tenemos participantes de la Academia de Ciencias, de la de Medicina, de la de Ingeniería, de la de Derecho.

La Academia Nacional de Economía, tiene ya 62 años, y se ocupa de temas sociales, económicos, tiene un Centro de Pensamiento y Propuestas llamado Pharos, que ya tiene cinco publicaciones que se pueden consultar en el sitio de internet de la Academia, www.acadeco.com.uy.

En especial, en la Academia, el tema de la pandemia motivó varias acciones, una de ellas fue el Premio Academia Nacional de Economía: este año hicimos una edición especial donde les pedimos a los jóvenes estudiantes y profesionales que hicieran propuestas de política económica y social para enfrentar la pandemia. Y la verdad es que se recibieron trabajos muy buenos, los dos seleccionados ya están publicados en el sitio de internet de la Academia y fueron estupendos. En uno de ellos, que tenía un modelo bastante complejo, Daniel Gianola nos ayudó en su corrección porque nos superaba en algunas cosas, así que resultó en una sinergia bien interesante y en un evento de premiación que si bien fue remoto, tuvo un brillo especial con los comentarios de los trabajos a cargo del Cr. Ariel Davrieux.

La charla de hoy cuenta con un expositor y una comentarista de renombre, con credenciales que realmente impresionan:

Daniel Gianola nació en Montevideo, Uruguay, se graduó como ingeniero agrónomo en la Universidad de la República en 1970, estudió en las Universidades de Wisconsin, obteniendo título de master y es PhD en Genética cuantitativa aplicada a la ganadería. Después de un periodo en el Banco Mundial fue catedrático en las Universidades de Illinois y Wisconsin. De esta última es profesor emérito. Ha sido premiado en varias partes del mundo incluyendo seis doctorados honoris causa conferidos por Universidades en Europa, Estados Unidos y la Universidad de la República. Además de ser profesor en Universidades de China, es miembro de la diversas Academias en varios países. Ha dado cientos de cursos de post grado a miles de estudiantes en más de 30 países, especialmente Alemania, Argentina, Australia, China,

¹ 10 de junio de 2020

Dinamarca, Italia, Noruega y Japón. Ha publicado 350 trabajos científicos, ha publicado libros. Introdujo la metodología bayesiana en la genética cuantitativa en animales y plantas en 1982. Recientemente adaptó el método de matching learning de aprendizaje automático. Además, él no lo puso en el currículo, pero nos ha nutrido con sus análisis del COVID – 19 a través de Juan Berchesi, a quien le agradecemos, por haber acercado a Daniel a la Academia.

Vamos a tener los comentarios de la Dra. Alicia Carriquiry. Alicia nació en Montevideo, es ingeniero agrónomo de la Universidad de la República, tiene una Maestría en cría de animales en la Universidad de Illinois, Maestría en estadísticas en la Universidad estatal de Iowa, doctorado conjunto en estadísticas y genética animal, desde 1990 está en la Facultad del Departamento de Estadística y es profesora distinguida de artes y ciencias liberales. Ocupa la silla de presidencia en estadística y es directora del Centro de estadísticas y aplicaciones de evidencia forense, un centro de investigación que está financiado con fondos federales. Es miembro electo de la Academia Nacional de Medicina, de la Asociación estadounidense de Estadística, del Instituto de Estadísticas y Matemáticas, de la Sociedad Internacional de Análisis Bayesiano, de la Asociación Americana para el avance de las ciencias y miembro electo del Instituto Internacional de estadística. Sus intereses de investigación incluyen; modelado de roles de medición, muestreo de encuestas y métodos bayesianos, y en los últimos años se ha interesado en aprender algoritmos y su aplicación en varias disciplinas, en particular en ciencias forenses y justicia penal. Y en sus ratos libres le hace caso a Daniel y a María Dolores que le piden que generosamente nos apoye en esta actividad.

Así que muchísimas gracias a ambos y sin más los dejo con Daniel.

Daniel Gianola

Muchas gracias a todos por acompañarme en este evento, gracias María Dolores por la invitación, y a la Academia por la amable presentación.

En primer lugar me gustaría darles una idea del ámbito, del ambiente en el cual vivo, y acá tenemos tres fotos de la ciudad de Madison, la primera a la izquierda es una vía aérea, Madison es la capital del estado de Wisconsin, con un área metropolitana de aproximadamente 350 mil personas y en la foto del centro tenemos una vista aérea del campus. La Universidad de Wisconsin tiene aproximadamente 43 mil estudiantes, de los cuales buena parte está en programas de posgrados y la investigación es un componente muy importante en la Universidad, en realidad el 80% del presupuesto de la Universidad está en proyectos de investigación financiados por el gobierno federal de los Estados Unidos, por el gobierno estatal de Wisconsin y también por el sector privado.



Y como el tema de hoy tiene mucho que ver con medicina quiero presentar en la foto que se ve a la derecha, el campus de ciencias biomédicas, que consiste de tres hospitales, más las escuelas de farmacia, y de enfermería.

Más abajo vemos fotos más locales. Como ustedes saben en las Universidades públicas en Estados Unidos, sobre todo en el medio oeste, los deportes y las bandas universitarias son muy importantes. En el ángulo inferior izquierdo está la banda de la Universidad de Wisconsin, marchando en Pasadena, California, donde este año jugamos, antes de la pandemia, contra la Universidad de Alabama, Wisconsin “marchó” pero la banda acompañó muy bien la derrota del equipo.

Aparte de estudiar hay esparcimiento y la foto del medio es la terraza de uno de los centros de estudiantes de la Universidad, un lugar muy agradable para tomar cerveza y conversar y cuando el tiempo lo permite, bailar.

Y naturalmente, Wisconsin es un estado que se encuentra al norte y con inviernos particularmente duros y acá tenemos alguna de las clásicas batallas de nieve entre los estudiantes.

Ahora les voy a recordar una famosa película que estuvo muchos años y también en teatro de Broadway, “Violinistas sobre el tejado” (Fiddler on the Roof) y lo traigo a colación porque uno de los personajes principales de la película, interpretada por el inolvidable Topol, cada vez que tenía un problema familiar consultaba con su Dios personal y trataba de evaluar todos los ángulos del problema.

Entonces cuando esto comenzó, fue que empecé a hacer gráficas para compartir entre amigos. Las gráficas siguiendo la lógica de internet llegaron a manos de la Academia y entonces a algunos miembros les pareció razonable e interesante que yo hablara al respecto. Pero yo no soy médico, no soy virólogo, no soy especialista en evolución molecular de virus, en fin, me dedico al mejoramiento genético ganadero.

Entonces como decía, Topol en la película hablaba con su Dios y decía “on the other hand.” Pero como tampoco se sabe mucho sobre el virus y por cierto no sabía que se iban a inscribir tantas personas que saben muchísimo más que yo sobre el tema, para resumir, ya que el

tiempo apremia simplemente voy a hablar de temas que no conozco y me animé a hacerlo porque hice el supuesto que los economistas tampoco saben mucho de los temas que les voy a hablar. De manera que ese es un poco el contexto en el cual agradecía esta gentil invitación.

TOPICOS A DISCUTIR

- Perspectiva: metodos empleados en mejoramiento genetico animal
- Descripcion de la pandemia de COVID-19: 16 paises
- Uruguay y contexto regional
- Estimacion de "R" (numero de infecciones/infectores)
- Modelo predictivo (simple) del numero de casos diarios
- Evaluacion de predicciones via "RANDEMIAS"
- Inmunidad de "rodeo" [si Uruguay fuera Suecia...]
- Comentarios miscelaneos

Voy a hablar los siguiente, primero les quiero dar un poquito de información sobre de dónde vengo, cuál es el área mía de trabajo porque eso influye la perspectiva con la cual voy a abordar el tema. Después voy a hacer una descripción de la pandemia me voy a focalizar en 16 países para que vean las formas de la curva epidémica, y posteriormente pasaría a describir la situación de Uruguay y el contexto regional.

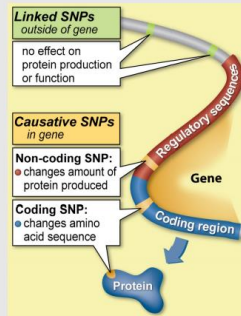
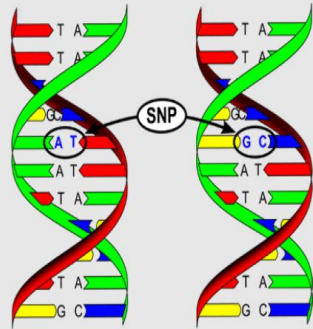
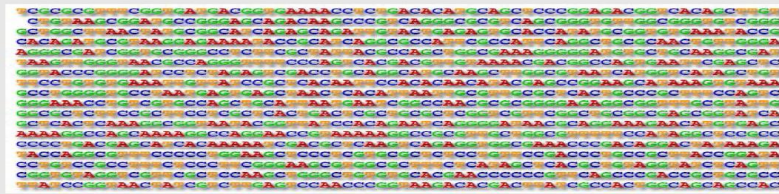
Después me ocuparé de algunos temas técnicos pero que son importantes, que es el de la estimación de un parámetro que es muy importante en epidemias, que se llama el número de reproducción del virus (R) y voy a desarrollar un simple modelo predictivo del número de casos diarios que ha sido bastante exitoso, después voy a agregar algunos comentarios sobre la evolución del modelo y qué es lo que implica ese relativo éxito en predecir los casos que van a ocurrir.

Después vamos a ver un concepto que yo le llamo "Randemias" y voy a explicar de qué se trata y después voy a hacer un pequeño análisis hipotético que hubiera pasado en Uruguay si hubiéramos adoptado la estrategia sueca, la cual le llamo inmunidad de rodeo.

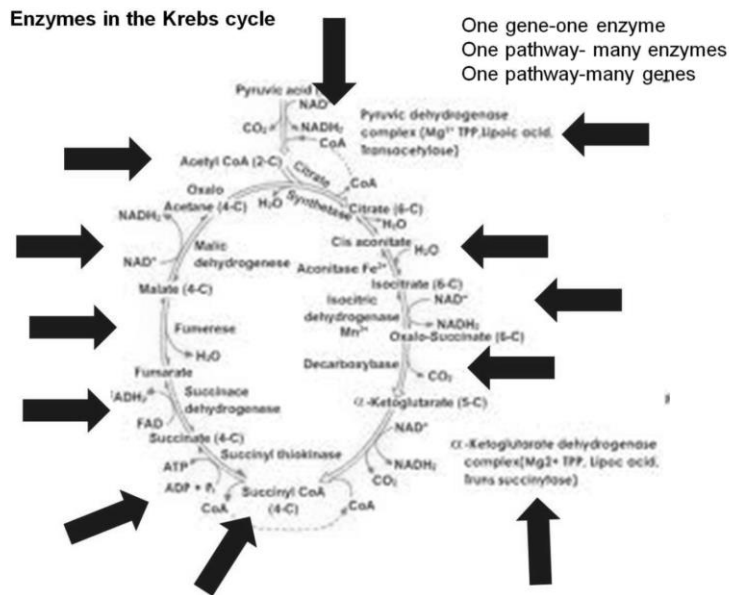
Finalmente haré una serie de comentarios misceláneos.

Hoy en día, los que trabajamos en mejoramiento genético de animales y plantas, tratamos de relacionar variantes genéticas a nivel de ADN con características productivas y la idea es no solamente localizar zonas genómicas que afectan características como crecimiento animal, producción de leche, propensión, susceptibilidad a enfermedades en animales, en plantas y humanos, la zona del ADN. Hay varios enfoques que se pueden hacer en las ciencias médicas. En su enorme mayoría este tipo de estudios no ha tenido valor clínico, más bien los que son la excepción a la regla y en agricultura y ganadería son de valor prácticamente nulo, porque nuestras características son extraordinariamente complejas.

MASSIVE NUMBERS OF MOLECULAR MARKERS AVAILABLE: DNA sequences
(cattle: 3 Gb; maize: 2500 Mb)

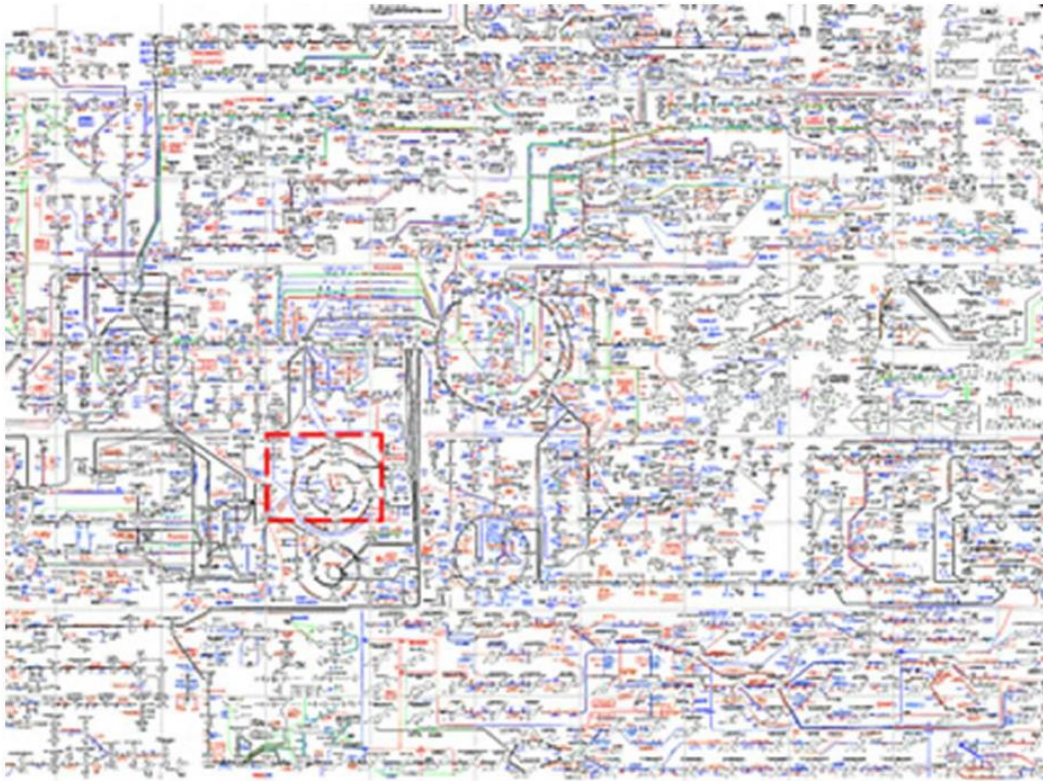


Para comenzar -esto es un recordatorio para los que tienen formación en biología- este es un ciclo bioquímico, el famoso ciclo de los ácidos tricarbóxicos en el cual intervienen en este solo ciclo dos enzimas y hay un viejo paradigma que es un gen - una enzima, de manera que en un ciclo relativamente sencillo como éste intervienen por lo menos doce genes.



El resultado de lo que nosotros evaluamos: la presencia o la ausencia de una enfermedad en una persona, o el rendimiento de una variedad del trigo o la producción de leche de una vaca durante 305 días de lactación, es el resultado de un número enorme de reacciones bioquímicas que están exquisitamente concertadas las unas con las otras y en franca interacción.

A “complex” trait involves many metabolic pathways: Roche’s Chart



Esta es una carta de reacciones bioquímicas que no se ve muy bien, pero para darles una idea este recuadro confinado en ese rectángulo acá es aproximadamente equivalente a la complejidad del sistema ferroviario de la ciudad de Frankfurt. Si ustedes han estado en Frankfurt, leer el mapa la intersección de trenes, es un enorme desafío.

De manera que esto es lo que llamamos una característica compleja

En biología tenemos dos tipos de enfoques: un enfoque deductivo, analítico de tratar de entender la base mecánica de los fenómenos y cuando los problemas son demasiado complejos como para usar ese tipo de reducción, usamos abstracciones.

BIOLOGIA-ABSTRACCIONES



A la izquierda tenemos una foto de Pablo Picasso es una foto real, excepto por las manos y a la derecha tenemos una abstracción hecha por el propio Picasso que es un autorretrato.

Traigo de esta colación para que interpreten esa abstracción que hace Picasso como un modelo y la foto de Picasso como la realidad. Entonces los modelos se aproximan a la realidad, pero no la describen necesariamente de manera fidedigna, inclusive no es deseable desde el punto de vista predictivo que los modelos reproduzcan la realidad de manera exacta porque entonces pierden capacidad predictiva cuando se encuentran con nuevas situaciones.

Coping with complexity

(WELCOME TO THE WORLD OF ABSTRACTIONS)

First assumption: there is a genetic signal and an environmental signal

Second assumption: the joint effect translates into a phenotype y

$$Y = f(G, E) \quad \text{For some UNKNOWN function } f$$

$$\text{Choices? } \left\{ \begin{array}{l} Y = G^E? \\ Y = E^G? \\ Y = G + E + GE? \quad \longrightarrow \text{Is an assumption} \\ Y = (G + E)^{GE}? \\ Y = G + E? \quad \longrightarrow \text{Is an even a stronger assumption} \end{array} \right.$$

En la manera en que tratamos con la complejidad vamos introduciendo más modelos matemáticos y hay un sinfín de modelos y especificación, aquí hay simplemente algunas totalmente idealizados de cómo podemos ligar variables genéticas, resaltadas con G y variables ambientales resaltadas con E a algo que observamos y representamos con Y. El número infinito de representaciones que también depende de la información que pongamos.

El campo de trabajo en la mejora genética animal, la genética cuantitativa, que ahora ha entrado con mucha fuerza en genética humana se basa también en trabajos del padre de la

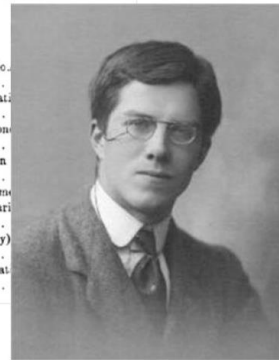
estadística que fue su gran obsesión, Fisher, esta es una foto de Ronald Fisher cuando era estudiante en Cambridge y a los 20 años envió un trabajo en el cual sentó, nada menos, que las bases de la genética cuantitativa y del análisis de varianza y regresión y tres años más tarde, en otro trabajo, introdujo los principios modernos de inferencia estadística y en particular las ideas centrales de un principio, que se llama principio de la verosimilitud y un famoso estimador que se usa mucho en economía también, que se llama un estimador de máxima verosimilitud.

Fisher, R. A. 1918. The correlation between relatives on the supposition of Mendelian inheritance. Transactions of the Royal Society of Edinburgh 52:399-433.

XV.—The Correlation between Relatives on the Supposition of Mendelian Inheritance. By R. A. Fisher, B.A. Communicated by Professor J. ARTHUR THOMSON. (With Four Figures in Text.)

(MS. received June 15, 1918. Read July 8, 1918. Issued separately October 1, 1918.)

CONTENTS.	
	PAGE
1. The superposition of factors distributed independently	402
2. Phase frequency in each array	402
3. Parental regression	403
4. Dominance deviations	403
5. Correlation for parent; genetic correlations	404
6. Fraternal correlation	405
7. Correlations for other relatives	406
8. Epistasy	408
9. Assortative mating	410
10. Frequency of phases	410
11. Association of factors	411
12. Conditions of equilibrium	412
13. Nature of association	413
14. Multiple allelomorphism	415
15. Homogamy and multiple allelomorphism	415
16. Coupling	416
17. Theories of marital correlations	416
18. Ancestral correlations (second theory)	417
19. Numerical values of association	417
20. Fraternal correlation	418
21. Numerical values for environmental ratios; analysis of variance	418
22. Other relatives	419
23. Numerical values (third theory)	419
24. Comparison of results	420
25. Interpretation of dominance ratios	420
26. Summary	421



Fisher, Young_Mat_Photo courtesy of Professor A W F Edwards by kind permission of Joan Fisher Box.jpg

Sobre la base de estos fundamentos cabalgamos aproximadamente durante 50 - 60 y en esa trayectoria se desarrollaron modelos lineales que han tenido un enorme impacto en particular en la mejora genética animal.

Charles R. Henderson y DG
University of Illinois
Urbana, 1987

En la foto del centro a la izquierda hay un profesor con el cual yo trabajé un año en la Universidad por lo tanto es colega mío, en la Universidad de Illinois, que se llamaba Charles Roy Henderson y Henderson desarrolló una serie de modelos que han sido enormemente útiles en agricultura, en medicina humana, en economía, que se llaman modelos mixtos.

Abajo tenemos al profesor Searle junto a un estadístico neozelandés que hizo el trabajo con Henderson y otros estudiantes, Dale Van Vleck, que hoy es profesor emérito en la Universidad de Nebraska. Y sobre el trabajo de estos tres gigantes se desarrollaron en cantidades de métodos de evaluación genética y reproductora que hoy se usan para comparar por ejemplo un ganado con otro.

A mediados del siglo 20 se produjo una revolución en el campo de la estadística que fue - en economía bien conocido dos economistas- econometrista, mejor dicho, James - Stein, que mostraron en un trabajo los modelos lineales y en particular el estimador de máxima verosimilitud fallaba cuando uno tenía, por ejemplo, un modelo lineal con más de dos parámetros. Esto fue ignorado durante muchos años y después fue revisado desde la perspectiva que llama bayesiana – laplaciana.

Bayes era un sacerdote que le gustaban las matemáticas y tiene un famoso teorema que hemos aprendido todos los que hemos hecho cursos introductorios de estadística.

Y Pierre Simon Laplace, que nació unos 15 años después de Bayes, no se conocían uno con el otro por supuesto, no había email, había palomas mensajeras, pero creo que no, no se conocían y Laplace se dedicó a usar ideas muy parecidas a las de Bayes, pero en lugar de tratar de problemas estadísticos abstractos como ¿cuál va a ser la posición que va a tomar una bola de color rojo en el juego de billar? Laplace se dedicó a analizar por ejemplo el ratio entre niños y niñas al nacimiento.

Como comercial y ya lo dijo María Dolores, al principio la década de los 80 a mí me interesó de sobremanera la influencia bayesiana y vi que usando ese enfoque se podían resolver muchos problemas en el campo de la mejora genética animal que no habían sido resueltos y hoy en día aplicamos en lo que llama selección genómica en animales y plantas y muchos de los métodos que se emplean están basado en ideas bayesianas.

Más recientemente nuestro (método) no ha sido indiferente a lo que llaman la era de machine learning y de inteligencia artificial, que es un campo que esencialmente prescinde de los modelos cuando menos de la interpretación mecanística de los modelos y trata de extraer patrones a través de los cuales se pueden construir máquinas clasificatorias o predictivas. Recientemente seguro que han escuchado hablar del método *deep learning*, son métodos sumamente efectivos en procesamiento de imágenes, por ejemplo, ¿cómo es posible cuando se busca una plancha Google les da una plancha de 25 dólares que se ajusta bastante bien de acuerdo a sus necesidades? Todo esto está basado en algoritmos que no tienen interpretación fácil desde el punto de vista mecanístico, pero que sí tienen una enorme habilidad clasificatoria y predictiva.

Ustedes recordarán el famoso libro el Código Da Vinci de Dan Brown, el último que se llama Origen, comienza en el museo Guggenheim de Bilbao y el personaje central es una computadora que se llama Winston que utiliza métodos de aprendizaje profundo, *deep learning*, que son redes neuronales, o sea representaciones esquemáticas y matemáticas de cómo funciona el cerebro para relacionarse y para los seres humanos. Al final termina en que todo es un complot para matar al Rey de España, pero eso no viene a colación.

Así que entonces en Estados Unidos esencialmente, hay evidencia de que el virus había entrado a este país, el gobierno demoró en actuar, pero empezamos a ver lo que estaba ocurriendo en Europa y algunos gobiernos estatales - hay que recordar que Estados Unidos tiene un gobierno federal- tomaron acciones rápidamente. El gobierno federal o sea la

administración de Trump al principio minimizó la importancia, dijo que el virus iba a desaparecer por un milagro, bueno lamentablemente no ha sido así.

Pero ese tipo de división resultó en demora, en inacción, en fallos, en falta de coordinación, en crear competencias falsas entre el Estado y el gobierno federal. De manera que en algunos Estados tuvimos órdenes de quedarnos en casa, Wisconsin fue uno de ellos, prácticamente se cerró la actividad económica, la actividad cultural, como saben ha ocurrido mucho en otros países.

Entonces el tema es qué hacemos estando en casa y hemos estado invadidos por sugerencias en internet. En el caso mío lo primero que hice fue leer libros sobre enfermedades, les recomiendo La plaga de Albert Camus, es una epidemia que transcurre hipotéticamente en 1948 al norte de África y es formidable porque describe situaciones y conceptos que se aplican hoy en 2020, en el contexto de la pandemia de COVID-19.

También hay un libro interesante de Daniel Defoe que más bien una descripción de lo que ocurrió con la plaga en Londres. Y también he estado leyendo un libro que se llama "Hombres y microbios", es formidable el libro, pero es a la vez aterrador porque nos hace tomar conciencia que vivimos en una especie de meta universo rodeado de virus y bacterias y todo tipo de organismos peligrosos.

También gracias a Netflix nos pudimos entretener bastante, les recomiendo la serie Califato que es sueca. El terror que me inspiró al libro de Hombres y microbios después se traduce en terror físico, cosas que ocurren en Suecia en una situación hipotética, pero que se aproxima a algunas cosas que han pasado en pandemia.

Y naturalmente me dediqué a hacer epidemiología casera, tratando de entender qué estaba pasando por lo menos cuantitativamente con la enfermedad y en particular en Uruguay, empecé a producir gráficas y las gráficas llegaron a manos de amigos, de ahí terminó en la Academia y por eso que les estoy hablando hoy en día sobre este tema.

Ayer la situación en el mundo era la siguiente de acuerdo de la página de la Universidad John Hopkins: había más de 7 millones de casos confirmados en todo el mundo con un poco más de 400.000 muertes globales constatadas. Estados Unidos como ven en la columna a la izquierda lidera el número de casos, por lo menos hasta ayer por la noche, Brasil venía segundo, Rusia tercero, Reino Unido cuarto.

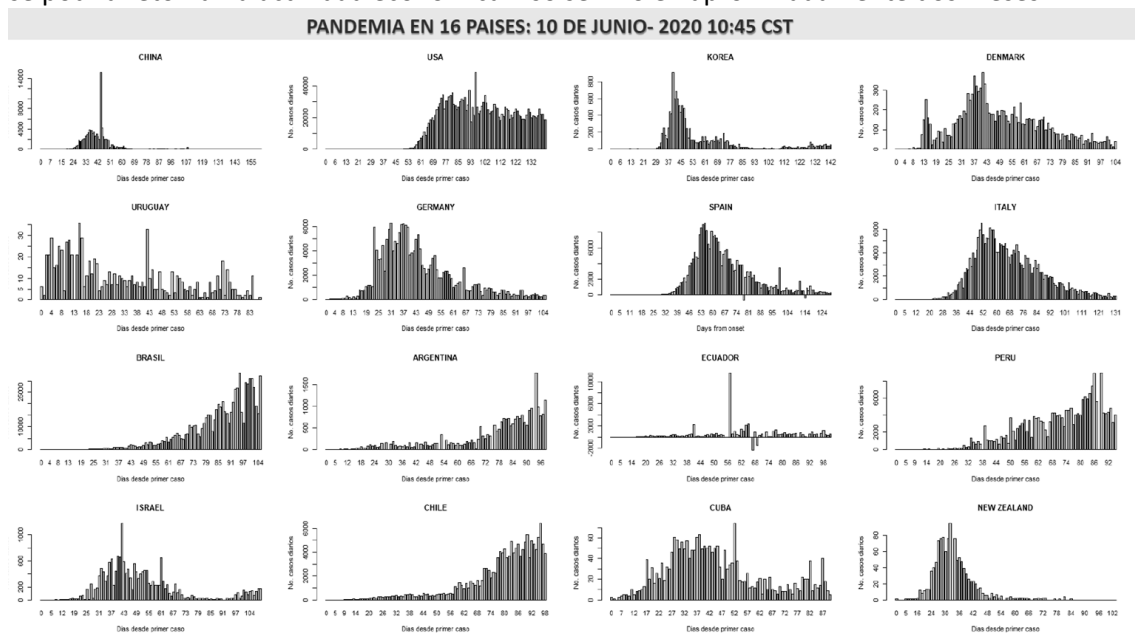


Y en las muertes, en este momento estamos en aproximadamente 112 mil muertes en Estados Unidos, 41 mil en Reino Unido, 37 mil en Brasil -esta probablemente es una cuota inferior del número de muertes- y hay algoritmos que predicen que el 3 de agosto el número de muertes en Estados Unidos sea probablemente alrededor de 135 mil, que es el estimado puntual, con un intervalo de confianza que llegar hasta casi 200 mil personas.

De manera que estamos ante un fenómeno evidentemente global y que tiene consecuencias muy serias sobre la vida humana.

Acá les presento la situación en 16 países, no están en orden alfabético.

Primero vamos a hablar de China, ustedes saben que el virus se transmitió de murciélagos probablemente a pangolines y de pangolines saltó a seres humanos en un mercado de Wuhan. No hay evidencia científica creíble que permita pensar que el virus no es resultado de evolución natural. En realidad, todo apunta que efectivamente la secuencia genética del virus es perfectamente consistente con lo que se sabe de evolución molecular. Y en China en la ciudad de Wuhan – que tiene aproximadamente 10 millones de habitantes, yo estuve en el año 2008 cuando se estrenaron los trenes balas que van desde Wuhan a Cantón, en cuatro horas o cinco horas-, la epidemia que se resolvió de manera muy rápida, por lo menos de acuerdo a la información que disponemos. Luego si la información es creíble o no, no voy a opinar al respecto. Pero la estrategia que usó el gobierno chino fue esencialmente cerrar Wuhan, limpiar la ciudad del virus y reabrirla en el momento en que consideraron que se podría retomar la actividad económica. Eso se hizo en aproximadamente dos meses.



En otros países la curva ha seguido trayectorias típicas de epidemia, por ejemplo, tenemos el caso de Brasil, Argentina, Chile, y Perú también, con una trayectoria muy consistente con el momento en que las epidemias crecen.

Otros casos bastante notorios son: Italia, España y Alemania. En España hubo una enorme mortalidad y Alemania es considerado como un país que ha tenido un éxito bastante remarcable con respecto al control de la epidemia y hace pocos días, Alemania está reabriendo de manera paulatina.

Esto es más o menos la trayectoria típica de una epidemia la fase de crecimiento exponencial, luego un crecimiento más o menos lento, dependiendo de las medidas de control epidemiológico que se tome.

Voy mencionarles en el caso de Suecia, que está en un estadio estacionario. Y en Estados Unidos donde la epidemia creció sobre todo debido al fenómeno de la New York City y New Jersey, ciudades que han tenido éxito en reducir la epidemia, pero ocurre que en Estados Unidos hay 10 epidemias paralelas, que comienzan a diferentes ritmos, progresan a diferentes ritmos y depende fundamentalmente de la situación política en cada uno de los Estados. Y aparte en días recientes, con el problema de la violencia policial, ha habido manifestaciones en todo el país, que inevitablemente se van a trasuntar en un enlentecimiento de la tasa de mejoramiento, si no en un empeoramiento.

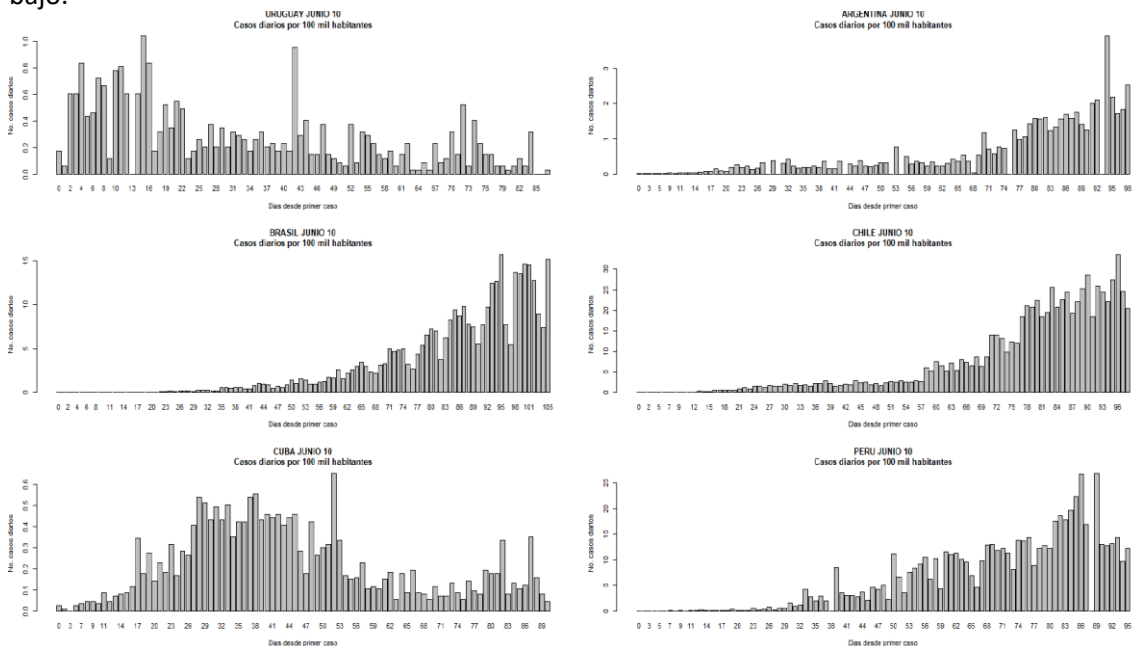
Otro país que ha sido bastante elogiado por la manera en que llevó a cabo el control de epidemias es el caso de Corea. Vean acá la fase epidémica del decrecimiento y todavía hay casos que se constatan son focos relativamente aislados que el país aparentemente está controlando con éxito.

Permítanme volver a China. En China se siguen constatando casos sobre todo en la parte noreste, en la provincia de Heilongjiang que tiene frontera con Rusia. Tengo colegas en Heilongjiang que argumentan que la culpa la tienen los rusos, China había resuelto el tema de la epidemia pero que después entraron a Rusia y hay mucha porosidad en la frontera Rusia – China, o al menos en esa zona -porosidad relativa nosotros tenemos bastante más.

De Uruguay me voy a ocupar con más detenimiento, acá tenemos una gráfica de la situación de Uruguay. Está gráfica la hice en la mañana así que está bastante actualizada.

Ahora una comparación regional expresado cada mil habitantes. Tal vez el país que ha seguido una trayectoria más consistente con lo que se esperaría de una epidemia es Cuba, pero vean que no se ha controlado de virus de manera definitiva y todavía hay fluctuaciones en el número de casos.

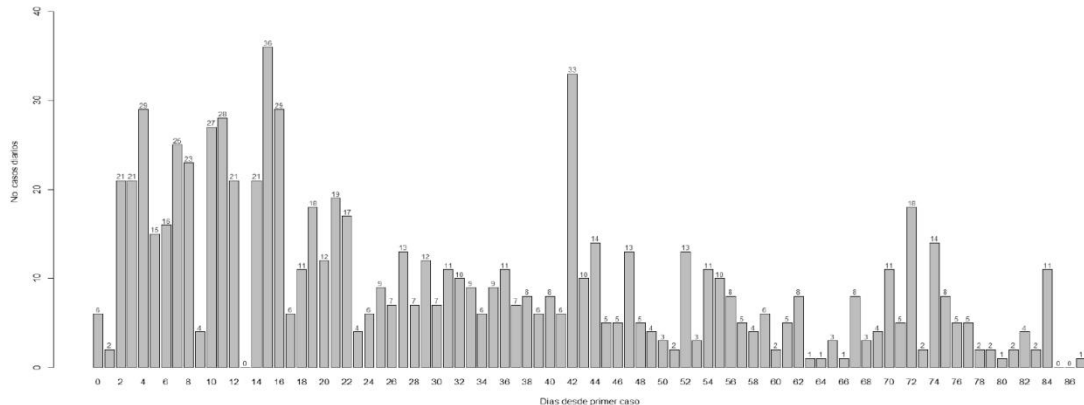
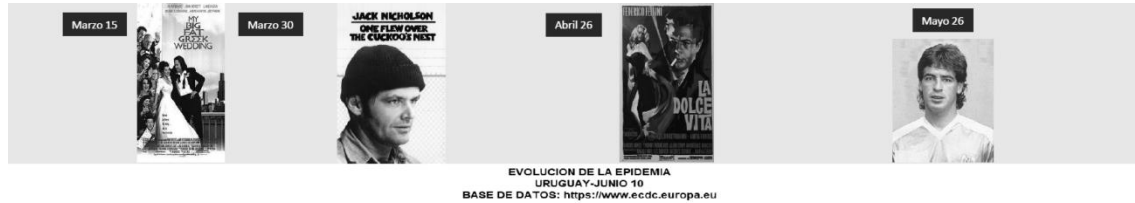
Uruguay y Cuba de estos seis países son los que tienen el número de casos per cápita más bajo.



Una cosa que tal vez es sorprendente, hablamos mucho de la gravedad de la situación de Brasil, pero cuando expresamos la epidemia por mil habitantes en realidad la situación de Chile y Perú es bastante más seria que la de Brasil, por lo menos de acuerdo a las cifras estas que son de una de una base de datos de European Center for Disease Control.

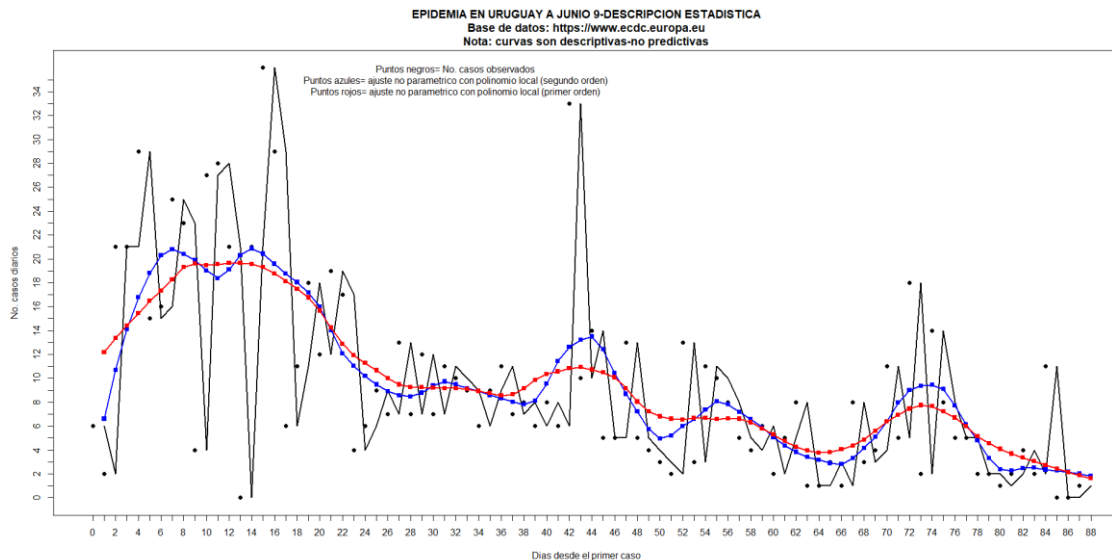
De manera que, en términos relativos, la epidemia en Chile y en Perú es bastante seria.

Esta ha sido la evolución de la epidemia y acá tengo una gráfica de barras con el número de casos en Uruguay hasta el día de ayer, que el Sistema Nacional de Emergencias reportó un caso, veníamos de dos días sin casos detectados.



Ahora voy agregar algunos mojones de la epidemia de Uruguay. No tiene una forma clásica, los epidemiólogos pueden opinar mucho más que yo al respecto, pero tenemos esencialmente el primer foco que se debió a un casamiento, no voy a hacer más comentarios al respecto ya que ha sido hablado ampliamente por las redes sociales. Después tuvimos el problema del hospital psiquiátrico y con todo el respeto, pongo la formidable película Jack Nicholson, *One Flew Over the Cuckoo's Nest*, simplemente para que ustedes lo ubiquen en términos temporales. Después tuvimos otro foco que fue en un residencial de ancianos, una cosa bastante grave. Uno de los residenciales se llamaba la Dolce Vita, como la película de Marcelo Mastroianni y Anita Ekberg, con el famoso baño en la Fontana de Trevi. Y después tenemos otro foco en el cual ustedes reconocerán al jugador del fútbol uruguayo Pablo Bengoechea, lo puse con la camiseta celeste para que no se produzcan conflictos con los hinchas de Nacional y los hinchas de Peñarol, pero es para recordarles que fue un foco que se produjo en la ciudad de Rivera.

Les comento porque sé que hay personas de otros países, Rivera un Departamento Uruguay y hay una ciudad que tiene unos habitantes con acento muy distintivo que llama Vichadero. Un día, estando en España, fui a un kiosco a comprar el diario y el diariero dice "buenos días" y yo le pregunto si es de Vichadero y me dice "¿cómo sabe?" Así hablan los de Vichadero.



Vamos entonces a analizar esas gráficas, usando lo que se llaman polinomios locales. Las líneas azules y rojas son dos formas de describir una trayectoria, pero sin un modelo explícito, tratar de alisar, de eliminar las variaciones que son claramente fluctuaciones estadísticas fuertes y las dos curvas muestran que estamos en un enlentecimiento gradual de la epidemia, pero a veces se han dado focos que han sido efectivamente controlados.

Ahora vamos a pasar de la descripción, al análisis y lo que voy a hablar es prácticamente todo lo que es epidemiología.

En epidemias hay un parámetro que se llama R, seguramente ustedes han leído todos en la prensa que es una de las definiciones el número promedio de infectados por infectador.

NUMERO DE REPRODUCCION DEL VIRUS: R= numero promedio de infectados por infectador

Basic Reproduction Number (R0) = Attack Rate * Contacts

AR = 60% AR = 40% AR = 100%
 RO = 3 RO = 2 RO = 5

Basic reproduction number, R_0

Covid-19 (high), $R_0 = 3$

Total infected: 40

N_0
N_1	X_{01}
N_2	X_{02}	X_{12}
N_3	X_{03}	X_{13}	X_{23}	.	.	.
N_4	.	X_{14}	X_{24}	X_{34}	.	.
N_5	.	.	X_{25}	X_{35}	X_{45}	.
N_6	.	.	.	X_{36}	X_{46}	X_{56}
.	X_{47}	X_{57}
.	X_{58}

Distribucion del intervalo serial del virus
 Probabilidad/número de días entre ataques de personas entre infectado e infectos

La idea es la siguiente: los tres casos en donde una persona infectada de color rojo se expone a cinco posibles infectados. En el primer caso de esos cinco se infectan tres, se llama una tasa de ataque 60%, de los 5 se infecta el 60 % y el R vale 3 en ese caso. En el panel del medio, se infectan dos personas y el R vale 2, la tasa de infectados es el 40%, y en el tercer ejemplo se infecta a las 5 personas. Esa tasa de ataque o ese coeficiente es peculiar a las enfermedades, a

los virus que producen el síndrome, pero también depende de factores ambientales, políticas epidemiológicas, en ausencia de intervenciones farmacológicas.

El sarampión, por ejemplo, tiene un R un poco mayor que 14, la poliomielitis de aproximadamente 6 y acá tenemos el Coronavirus, cuyas estimaciones al principio de la epidemia varían desde 1.5 hasta 4 dependiendo de la base de datos.

En el panel de la derecha tenemos un poco la lógica de inflexión. Si suponemos que una persona en promedio infecta a 3, ese infectado infecta a tres más, tres más después de uno dos tres ciclos de infección hay un total de 40 infectados, en ausencia de políticas de intervención. Y esa es una de las cosas que le confiere peligrosidad al COVID en adición al tema de que hay un alto índice de asintomáticos, que tienen capacidad de contagiar y de manera diferencial, dependiendo de la carga viral de esa persona asintomática. Es un tema en discusión.

Ahora vamos a suponer que al comienzo de la epidemia se generan un número cero de casos. Pero esos casos comienzan a incrementar en el transcurso de los días, en este caso suponemos que esos casos generan infecciones en tres días sucesivos, después al día siguiente se generan casos que producen infecciones en tres días subsiguientes y así sucesivamente. Hay un concepto que es muy importante, que interviene en los cálculos que efectuamos, que se llama la distribución del intervalo serial del virus. Esa es la distribución de probabilidades del periodo de tiempo que media entre la aparición de síntomas en un infectado y en un infectador. El número de días puede ir desde 1 hasta más de 18 y estas son las probabilidades de que la infección se produzca en el día 1, en el día 2, en el día 4, en el día 5. Esta distribución del intervalo serial es muy importante porque se utiliza en el modelo de estimación del parámetro R.

ESTIMACION DE R Y PREDICCIÓN DE NUMERO DE CASOS

Inferencia
Teorema de Bayes

$$= \frac{p(R_t | \text{Casos "nuevos", Casos "viejos", supuestos iniciales})}{\int \frac{\Pr(\text{Casos "nuevos"} | R_t, \text{Casos viejos, supuestos}) p(R_t | \text{Casos "viejos", supuestos})}{\Pr(\text{Casos "nuevos"} | R_t, \text{Casos viejos, supuestos}) p(R_t | \text{Casos "viejos", supuestos})} dR_t}$$

Predicción
Distribución Predictiva

$$= \int \Pr(\text{Casos futuros} | R_t, \text{Casos "N" y "V", supuestos}) p(R_t | \text{Casos "N" y "V", supuestos iniciales}) dR_t$$

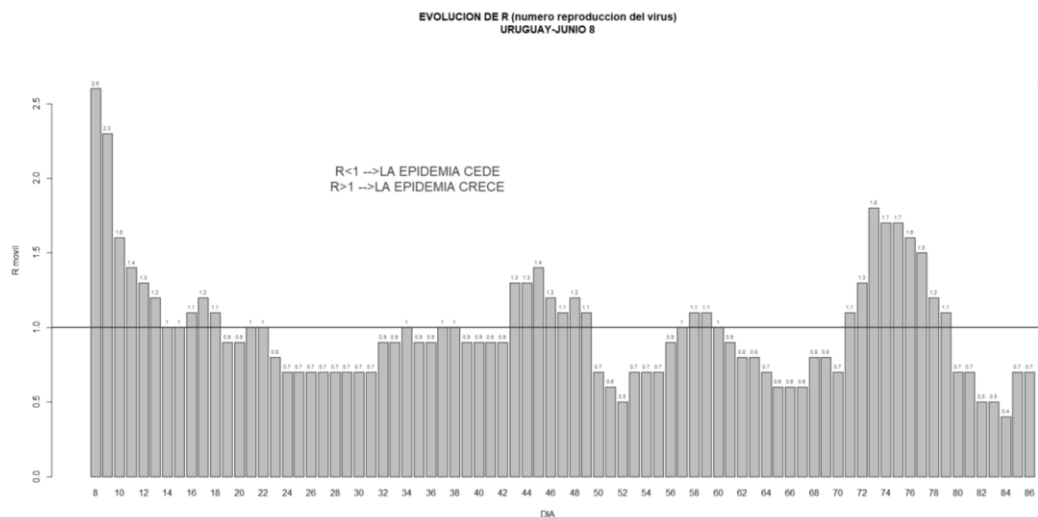
Esta es la gráfica más complicada simplemente está basada en la lógica bayesiana. Este modelo es exactamente el que usa el software Imperial College y la idea es que se parte de unos supuestos iniciales, esos supuestos iniciales son de dos parámetros que describen la incertidumbre inicial con respecto a la tasa de contagios del virus y otro supuesto inicial es la distribución del intervalo serial que vimos en la diapositiva previa.

Entonces comienza la epidemia, se observan casos y a través de esos casos se calcula una

distribución de incertidumbres de ese parámetro y esto es una aplicación del teorema de Bayes.

Una cosa que no hace el Imperial College es usar el teorema para construir una máquina predictiva y esto es un subproducto de la inferencia bayesiana, que una vez que se cuenta con esta distribución de incertidumbre, que se llama la distribución posterior es bastante fácil construir lo que se llama una distribución predictiva. Y nos dice que por ejemplo mañana se producirán 0, 1, 2, 3, 4, 10, 20, 30, 40, casos y esta es la forma en que se calcula esa distribución predictiva.

Y con estos dos elementos yo he jugado, construyendo y tratando estimar el parámetro R diariamente y a la vez que pasar una especie de pronóstico, que hasta el momento parece funcionar mejor que el de Inumet según dicen mis amigos, para ver qué podemos esperar al día siguiente. Es posible por supuesto hacer predicciones más allá del día siguiente y si contáramos con información con suficiente granularidad, por ejemplo, si hubiera un número suficiente de casos que nos permitieran separar entre edades, sexos, si pudiéramos, por ejemplo, tener los genotipos a nivel las secuencias genéticas, podríamos construir modelos predictivos mucho más ricos que este que está simplemente basado en observar el número de casos.



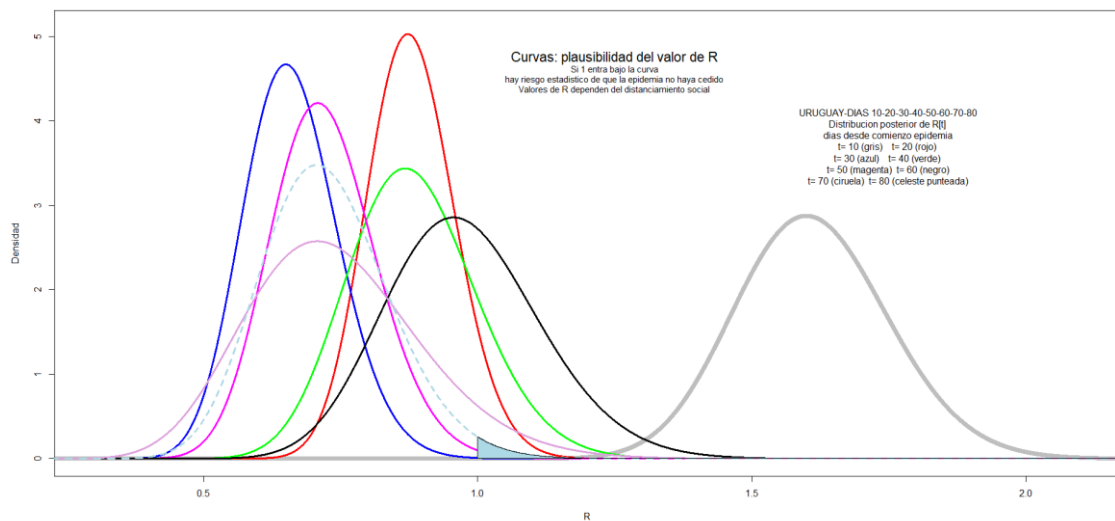
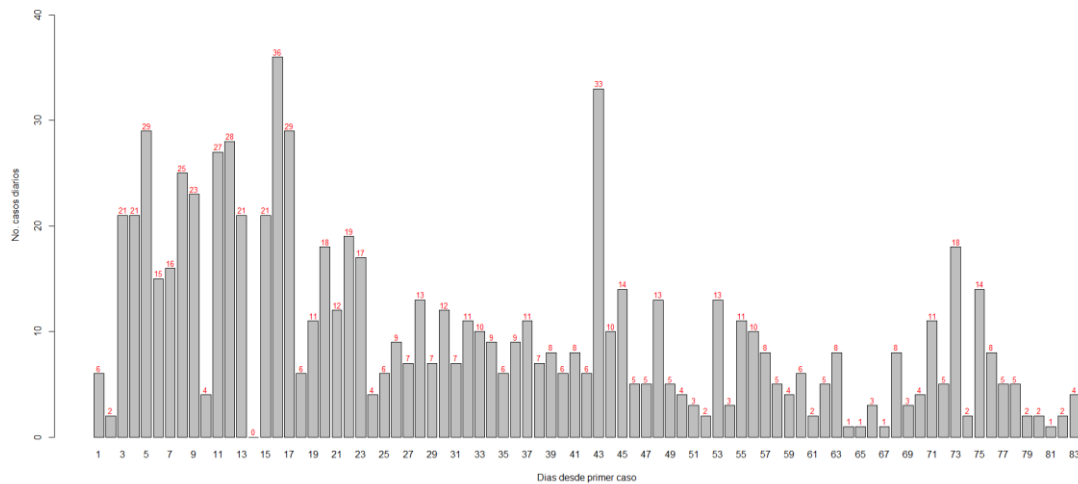
Esta es la evolución del parámetro R estimado, estas son las medianas de lo que llamamos la distribución posterior. Como ven, la epidemia en Uruguay comenzó de valores bastante altos consistentes con lo que sabemos de Coronavirus, después del foco casamiento y del foco del hospital psiquiátrico, pasamos a valores de R inferiores a 1, cosa que muy bueno, porque cuando R es menor que 1, la epidemia cede y hay unos modelos que se llaman S y R que permiten ver con claridad por qué eso ocurre. Después se produjo el foco en los residenciales, R volvió a aumentar, nuevamente vinieron los bomberos apagaron el incendio y después hemos estado transitando entre valores bastante bajos de R y nuevamente subimos a la zona peligrosa debido al foco de Rivera,

Pero es importante anotar que estos números que están asociados con incertidumbre y el Programa lo que hace es, aparte de calcular las medias y las medianas, calcula intervalo de credibilidad de 2.5 % y 7.5 %. Pero si yo les digo a ustedes, ¿Ustedes se subirían un avión si tiene solamente 97?5 de probabilidades de llegar al destino? Creo que no lo harían. De manera que es importante caracterizar la incertidumbre y eso lo podemos hacer usando la lógica

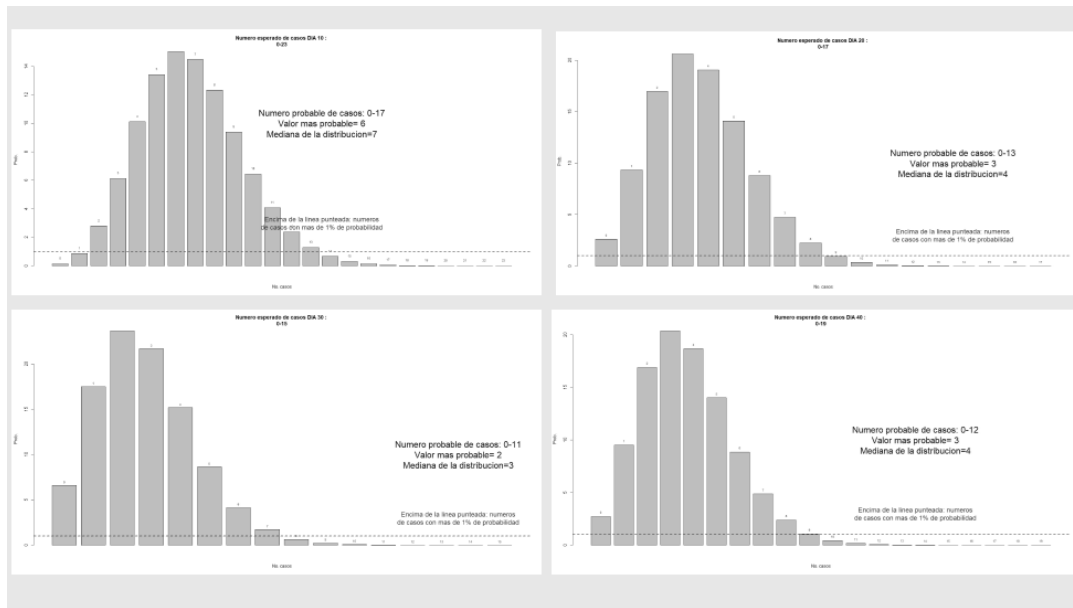
bayesiana que permite construir la distribución de probabilidad de incertidumbre de parámetros.

Y escogí arbitrariamente 8 días los días 10, 20, 30, 40, 50, 60, 70 y 80 de la epidemia para mostrarles la incertidumbre sobre R y después para evaluar cómo se comporta el modelo desde el punto de vista predictivo, cuando intentamos predecir los números de casos que actualmente ocurrieron en cada uno de esas ocho fechas.

EVOLUCION DE LA EPIDEMIA
URUGUAY-JUNIO 4
BASE DE DATOS: <https://www.ecdc.europa.eu>

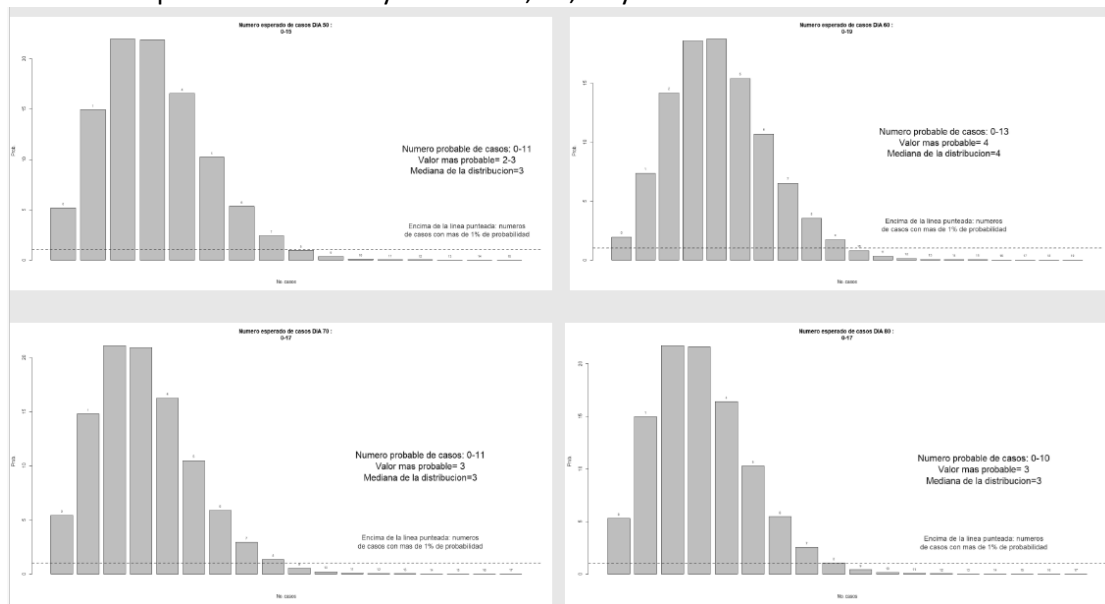


El día 10 de la epidemia son estas curvas de color gris, acá tenemos los valores en el eje de las abscisas. Como ven, la densidad de los valores mayores que 1 era apreciable, estamos en una situación de epidemia clara más allá de la duda razonable. Al día 20 pasamos a la curva de color rojo en que ahora al pasar de valores de R claramente mayores de 1, pasamos a valores en los cuales había una alta probabilidad de que fueran inferiores de uno, pero una probabilidad apreciable de que fueran mayores de 1. El día 30 la situación mejoró, el día 40 la situación empeoró. Noten que no solamente las curvas se desplazan, sino que “engordan o rebajan peso” por así decirlo. El día 50 tenemos la epidemia, en la distribución de color magenta, el día 60 de color negro, el día 70 la del color ciruela y el día 80 de color celeste. En la última curva todavía tenemos una probabilidad de 1.5% de que R es mayor que 1. Así que caminando pero por las piedras, por así decirlo.



Vamos a mostrar ahora cómo convertíamos el modelo a una máquina predictiva. Estas son las distribuciones predictivas para el día 10, para el día 20, para el día 30, para el día 40. Cualquier número de casos que entre dentro de esta distribución predictiva, uno técnicamente lo puede considerar como una buena predicción. Pero en cambio si el número de casos muy superior y entra fuera la distribución predictiva, deberíamos cuestionar por lo menos algunos aspectos de la máquina predictiva.

Acá tenemos para los otros días y los días 50, 60, 70 y 80.



¿Cómo evaluamos si las predicciones son buenas o malas?

En el área predictiva lo que generalmente hacemos son validaciones cruzadas: construimos modelos utilizando parte de los datos les llamamos la muestra de entrenamiento y después evaluamos cómo se predicen observaciones que han sido incluidas en otra etapa

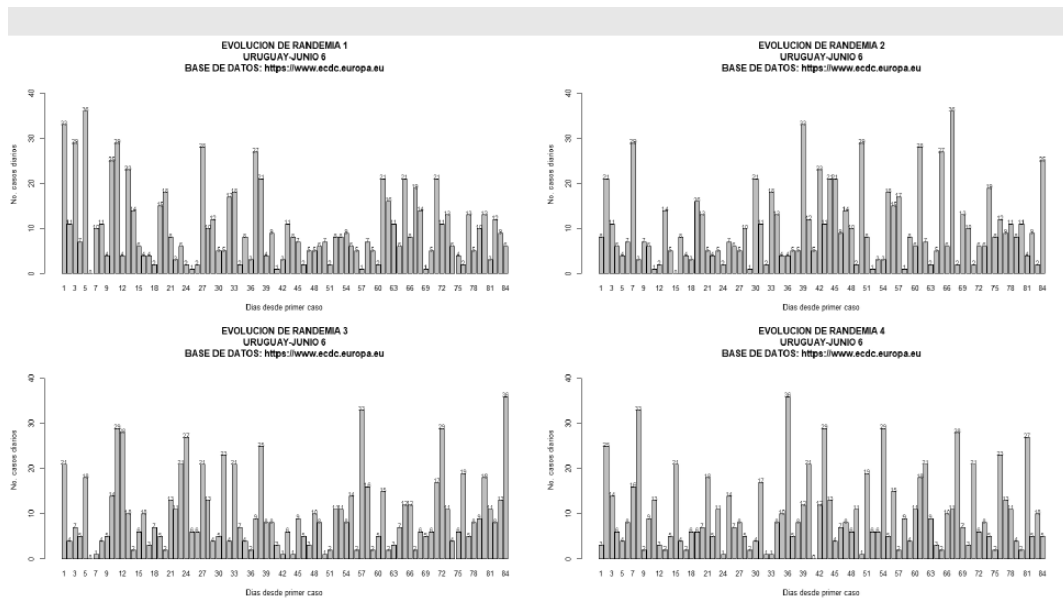
En este caso yo no dispongo de otra máquina predictiva. Entonces decidí comparar las predicciones con la peor máquina predictiva posible que es el caos, la entropía.

Seguramente vieron en Netflix una serie israelí muy buena llamada Fauda, que significa caos: ¿Cómo construimos el caos?

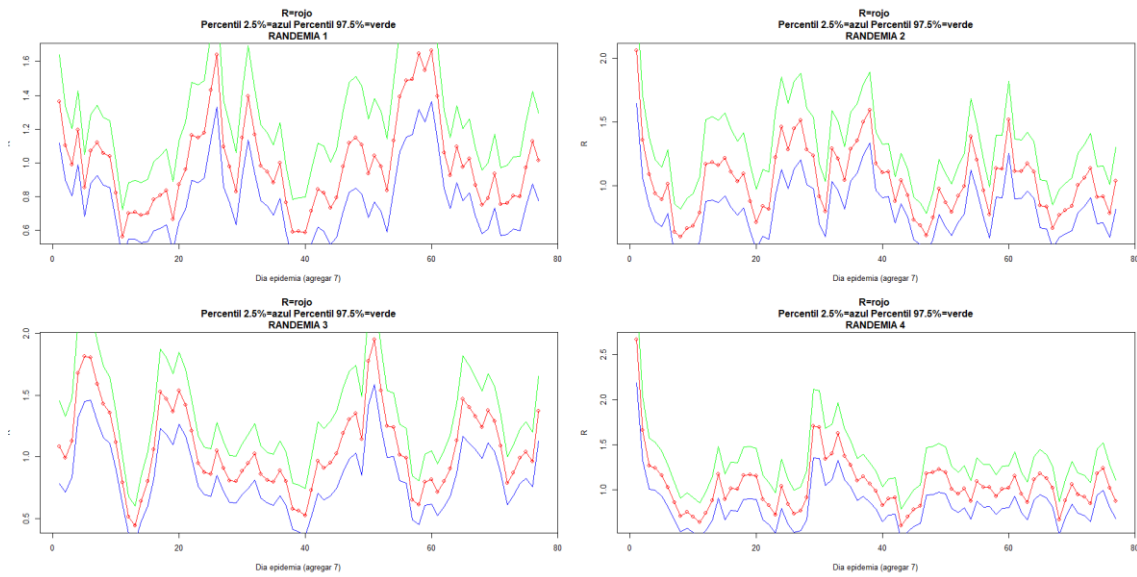
El caos, estamos hablando de pandemias -bueno acá tengo el copyright de esto-, vamos a hablar Randemias, que son epidemias aleatorias (random).

Tomamos todos los números de casos, por ejemplo, el día 83 de la epidemia Uruguay y los re-muestreamos al azar sin reemplazo. Si recuerdan matemáticas de preparatorios hay factorial 83 posibilidades que son 2.11×10 elevado a la 132 Randemias. Son epidemias totalmente al azar en el cual destruimos la asociación temporal entre el número de casos y los días en que ocurrieron en la epidemia verdadera, la epidemia real.

Bueno una Randemia es un poco equivalente como subir a un juego de un parque de diversiones de Montevideo, el Parque Rodó y los que tenemos unos cuantos años - yo no sé si existe todavía- recordarán que el divertimento consistía en subirse al Gusano Loco. Iba para atrás, iba para adelante, nunca se paraba, nunca se sabía cuándo iba a empezar la vuelta. Entonces de la idea es contar un número de virus en el gusano loco y hacerlos que se distribuirán al azar.



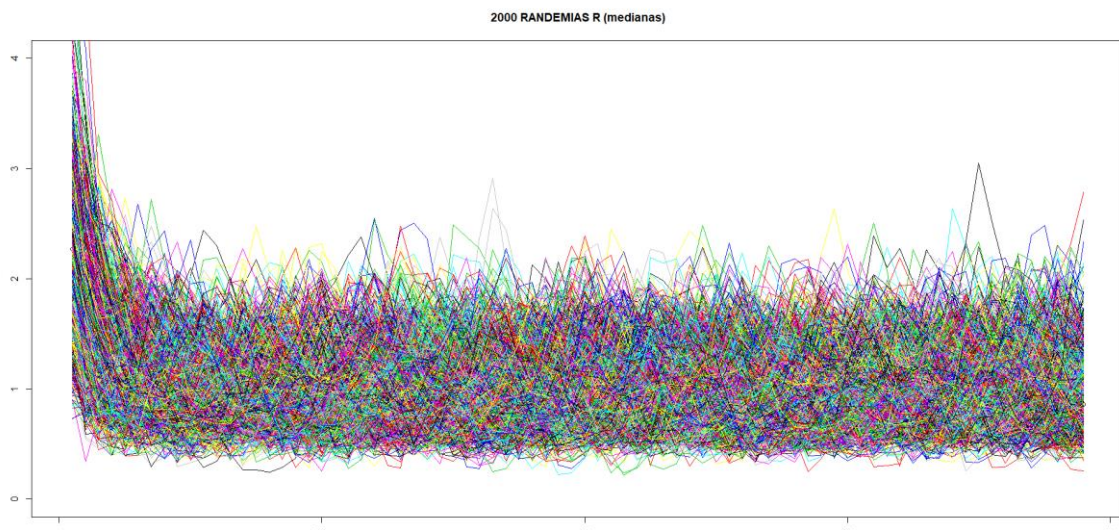
Acá tenemos 4 Randemias, esto se los muestro simplemente para que vean lo que ocurre. Tenemos el número mayor de casos que recordarán estuvo asociada a la *wedding*, número 36 acá en esta Randemia aparece en la posición 5, en esta aparece atrás, en un último día de la epidemia y aquí aparece en el número del día 36.



Para esas cuatro Randemias estimé el R, usando el programa Stem del Imperial College, las líneas rojas son las estimaciones del valor de R de esa epidemia aleatoria y las bandas verdes y azules son los percentiles de confianza, que dan entre 97.5 y 2.5% para cada una de las 4 Randemias.

Entonces estos son los valores de R que uno esperaría simplemente por variación aleatoria, una epidemia que carece de estructura. El virus está circulando como el gusano loco, pero la lógica de la inflexión es realmente aleatoria.

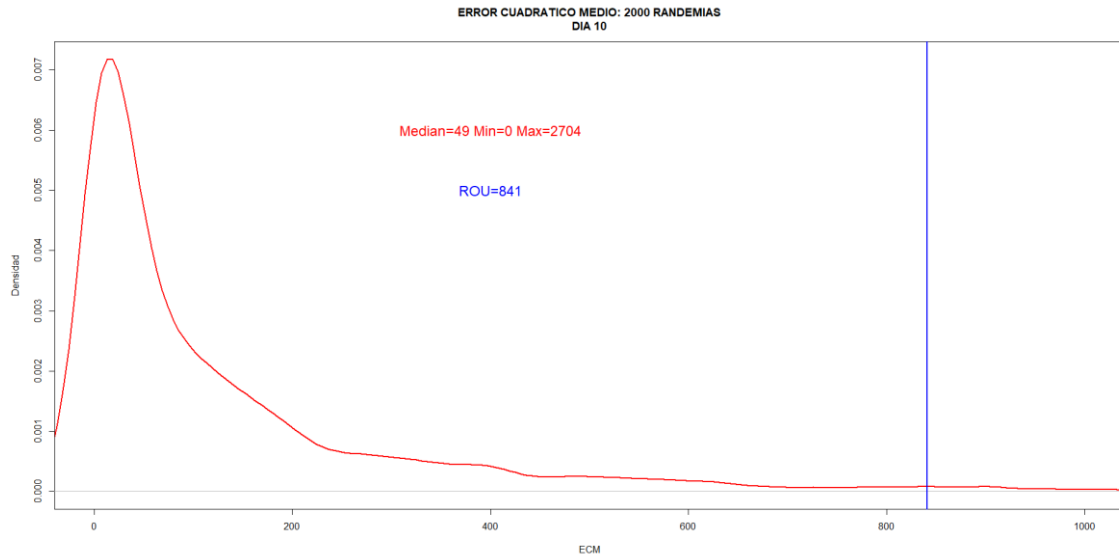
Por supuesto, no me voy a quedar con 4 Randemias y no pude hacer las 2.11 elevado por 10 a la 132 porque creo que no me alcanzarían todos los recursos de la Universidad de Wisconsin, pero acá tenemos 2000 Randemias.



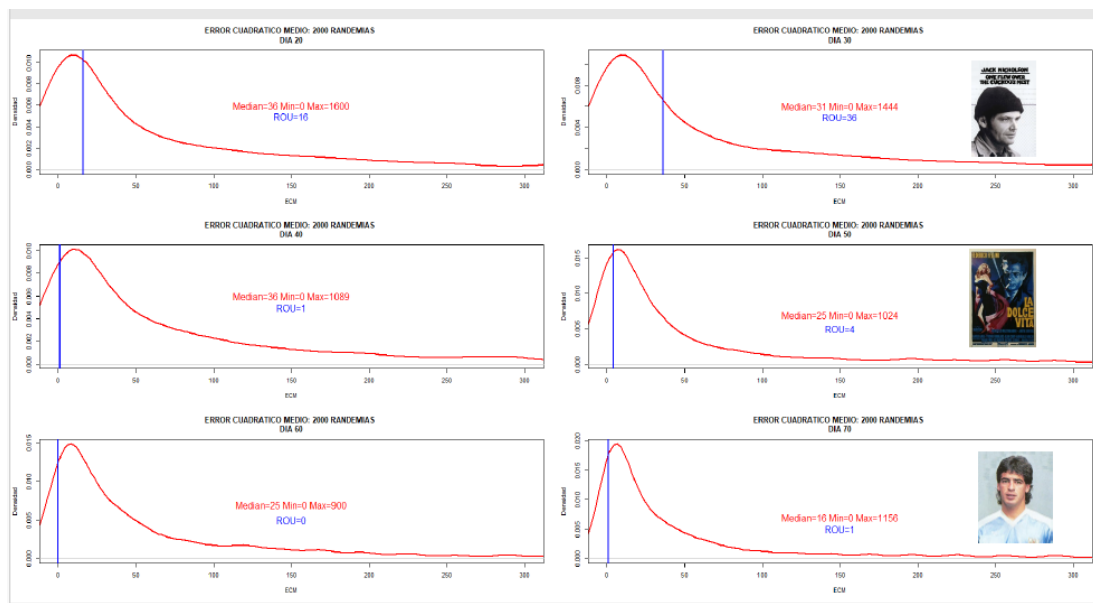
Observen que estas Randemias navegan por la situación estacionaria excepto al principio de la Randemia en las cuales todos comienzan con valores elevados de R y esto se debe a que el programa utiliza lo que llamo una distribución a priori en el cual a priori el valor de R es 5, pero con alta variabilidad, entonces en la medida que empiezan a ocurrir casos es a ese supuesto que es fuerte y lo equipara por ejemplo al virus de la polio, va siendo corregido

rápidamente por la evidencia y en la medida en que transcurre el tiempo, la evidencia pesa más, más, más, más, y más.

Usando esas 2000 Randemias, calculé lo que se llama la distancia euclidiana cuadrática entre lo que predeciría el modelo totalmente al azar y lo que verdaderamente ocurrió en el día 10.



Y esta distribución, es la distribución de esas distancias euclidianas cuadráticas y esta línea azul es la predicción hecha por el modelo. O sea, la primera la predicción para el día 10 del modelo fue espantosa porque tiene una distancia euclidiana cuadrática mucho más alta que las epidemias al azar. Y esto se debió a que el día 10, hubo 4 casos, pero en los días precedentes se dieron todos los casos que derivaron del *big fat greek wedding*, que andaban entre 20, 30, 25, 18, entonces, el modelo sobre predijo: cuando quería predecir 4, predijo 20 aproximadamente, eso produjo un error cuadrático medio muy alto.



En el día 20, la situación mejoró notoriamente, la mediana del error cuadrático medio en Randemia fue 36, el modelo fue de 16, o sea que redujimos la incertidumbre más del 50% con respecto al caos.

El día 30, el modelo prácticamente se comportó como una Randemia 36 versus 31, debido nuevamente a que esto fue después de los casos del hospital psiquiátrico, entonces el modelo sobre predijo porque hubo mucho menos casos de los que había habido en días precedentes.

Después se ha comportado bastante bien, el día 40 por ejemplo el error cuadrático medio del modelo es 1 dividido 36 del error cuadrático medio de la Randemia, el día 50 es más o menos la sexta parte, en el día 60 es infinito, en el día 70 el error cuadrático medio de la predicción fue un dieciseisavo del error cuadrático medio de la Randemia.

O sea este modelo que es ingenuo, carente de granularidad, simple, tiene capacidad predictiva, lo cual nos da muy buenas noticias. Quiere decir que si en Uruguay o en cualquier otro país, desarrollamos bases de datos con alto nivel de granularidad, si usáramos alta tecnología para medir desplazamientos, personas, eventos de contacto y agregar información basal sobre las características de los pacientes, podemos adelantarnos al virus y predecir cómo se va a comportar por ejemplo unos 5 ó 10 días antes. Es una situación a evaluar, son excelentes noticias.

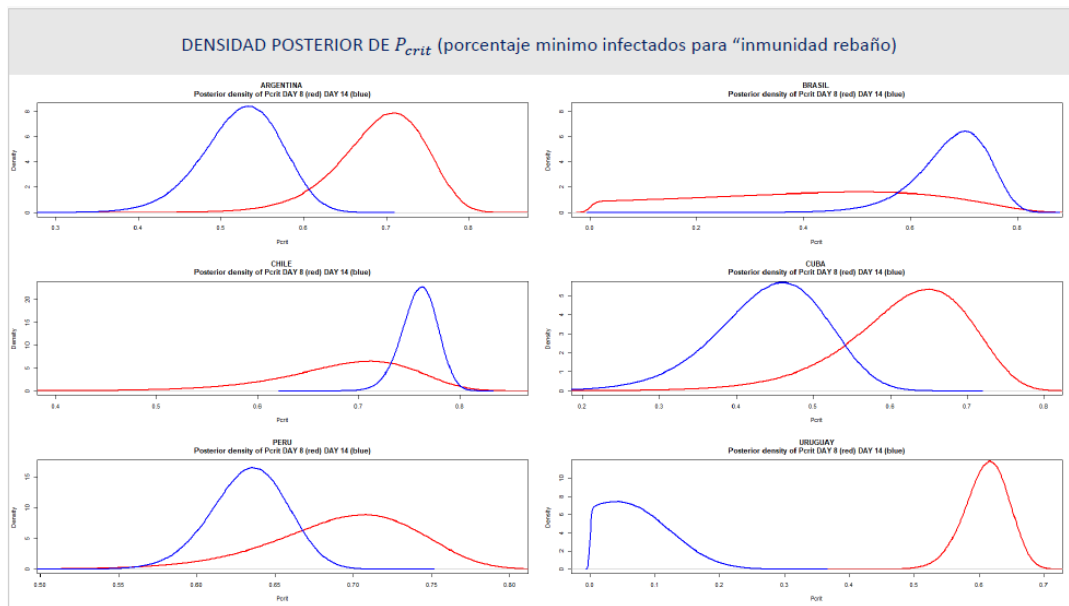
Si esta lógica la pudiéramos aplicar en el contexto uruguayo, no solamente a las epidemias humanas sino a las epidemias en animales, como el caso de la aftosa del año 2000, las epidemias vegetales, recordaran la famosa epidemia de los cítricos y también hay epidemias en cultivos, tendríamos la capacidad de adelantarnos al comportamiento de las epidemias y tomar medidas proactivas y orientadas a reducir los impactos.

Vamos a preguntarnos si fuéramos Suecia. En Suecia, el gobierno resolvió dejar que el virus se disemine por la sociedad, cuidar a los ancianos y así tener inmunidad. En la teoría epidemiológica hay un concepto que se llama "inmunidad de rodeo" y se da cuando cierta proporción de la población ha contraído el virus, suponiendo que los anticuerpos tengan valor protector, eso es equivalente prácticamente a tener la vacuna, porque el virus entonces no encuentra oportunidades de infección porque muchas personas que están inmunizadas, entonces la tasa de ataque es bastante baja.

Eso fue lo que optó Suecia.

Vamos ahora a preguntarnos qué hubiera ocurrido si Argentina, Brasil, Chile, Cuba, Perú y Uruguay hubieran adoptado la inmunidad de rebaño.

Para estimar el porcentaje mínimo infectado que tiene que tener la población para adquirir inmunidad rebaño, hay que tener conocimiento del valor del famoso R, no solamente del R sino de su incertidumbre y hay que medirlo rápidamente en la epidemia, que es cuando hay más incertidumbre por falta de número de casos.



Por ejemplo, vamos a ir directamente al caso de Uruguay. La línea roja son los porcentajes de inmunidad de rodeo que tendríamos que haber adquirido en Uruguay para proteger al resto de la población. Tendríamos que haber dejado que por lo mínimo se infectara el 50 % de la población. Si el gobierno hubiera optado por esa alternativa, tendríamos que haber esperado que el 50% de la población adquiriera inmunidad. ¿Cuánto tiempo eso lleva? No se sabe bien.

Al día 14, la situación mejoró sensiblemente con respecto a R, lo cual hizo que bajara el porcentaje mínimo de personas que se requiere que se infecten y pasamos a más o menos en un 10% o a lo sumo un 20 % de la población.








Para comparar, por ejemplo, vemos el caso de Argentina. En Argentina el día 8 hubieran necesitado que el 70% se contagiara, en Chile un 75% y así sucesivamente.

Vean este famoso numerito del porcentaje de la población, no puede dissociarse de la incertidumbre asociada.

¿Qué resultados dio la política sueca?

Una cosa que es muy importante y Alicia lo sabe mejor que yo -ella es una estadística profesional yo soy un estadístico del montón-, es que la ciencia se concibe no solamente de manera experimental, sino también de manera observacional. Y que cuando uno trata de extraer señales del ruido de datos observacionales, hay variables que llaman confounders, es decir cuando comparamos a Suecia con Uruguay hay varios confounders, por ejemplo, nosotros tenemos playas, los suecos tienen playas, pero yo no iría a una playa sueca. En Uruguay jugamos bastante bien al fútbol y los suecos más o menos. Uruguay es un país menos próspero que Suecia. En fin, todas las naciones difieren en un número enorme de aspectos, por los cuales las comparaciones entre naciones son sumamente delicadas.

Una forma de hacerlo es tratar de hacer comparaciones con naciones que son más o menos similares. Pero eso es verdad, más o menos, porque aún naciones que son aparentemente similares, difieren los aspectos. Por ejemplo, en Dinamarca la montaña más alta tiene 300 metros, es más o menos como el Cerro de Montevideo, en Noruega si van por ejemplo al Preikestolen, donde Tom Cruise hizo una película que seguramente han visto en el cine, hay montañas de miles de metros, o sea se parecen en estructuras sociales y en etnias, Dinamarca y Noruega, pero también difieren en otras cosas.

	PAIS	POBLACION (millones)	Tests/ millon	Casos/ millon	Muertes/ millon	PBI 2020(%)	
	SUECIA	10.1	27290	4431	462	-5.1	
	ALEMANIA	83.8	51917	2218	105	-6.1	
	DINAMARCA	5.8	121992	2063	102	-4.5	
	NORUEGA	5.4	47487	1577	44	-6.0	
	EEUU	330.9	63484	6038	339	-3.8	
	URUGUAY	3.5	13829	243	7	-----	
	REINO UNIDO	67.9	82241	4217	597	-8.7	

Las muertes por millón. Recuerden que Suecia optó por dejar que el virus se diseminara en la comunidad. Suecia tuvo por lo menos hasta anteayer, porque tres días atrás hice esta tabla, 462 muertos por millón. Alemania es bastante parecido a los países nórdicos en lo que se refiere al orden, a la disciplina, a las estructuras sociales, Alemania que es un éxito, especialmente comparado con los otros países de Europa, 105 muertes por millón. Dinamarca 102 muertes por millón, Noruega 44 muertos por millón, inclusive los Estados Unidos donde efectivamente tenemos una Randemia porque en este momento somos como un barco en llamas, a la deriva tenemos 339 muertos por millón.

Si tomamos esta diferencia y la expresamos en relación a lo que es la población en Estados Unidos, con más de 330 millones de habitantes, hoy los Estados Unidos tendrían 40 mil muertos adicionales. O sea, ya tenemos 35 mil muertos adicionales de acuerdo al estudio de Columbia University, 35 mil muertos y si a eso le agregamos 40 mil muertos, tendríamos 75 mil muertos en exceso en adición a los que tenemos.

Y finalmente el Reino Unido, comenzó con la estrategia sueca, pero cuando le tocó a Boris el COVID 19, echó para atrás y empezaron a tomar medidas. El número de muertes en el Reino Unido es dramático.

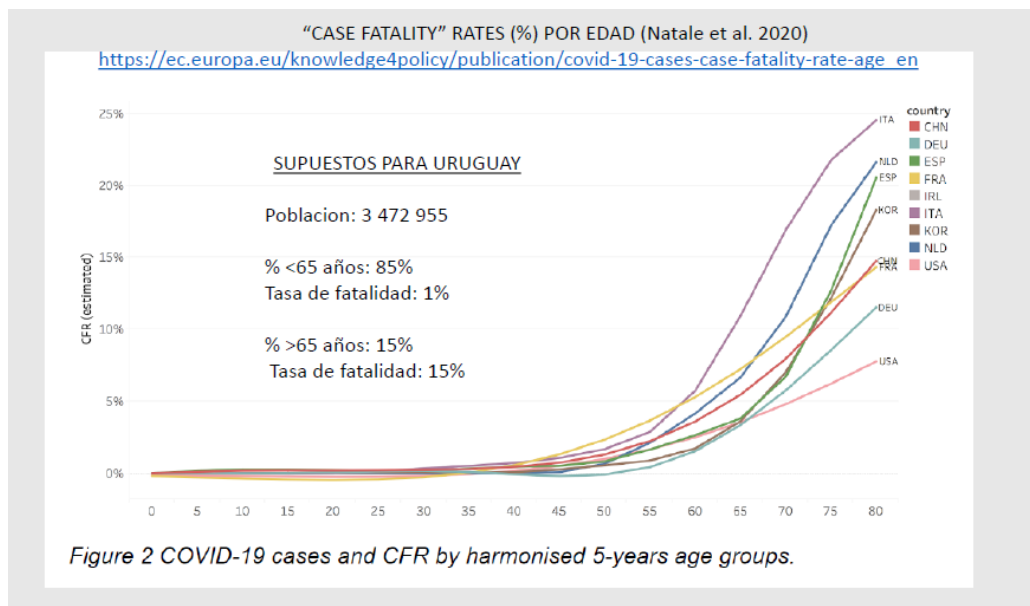
Lamento politizar esto, pero me siento como ciudadano del mundo en la obligación de decirlo. En Estados Unidos es un año de elecciones, estamos viviendo la atmósfera política muy especial, pero hasta hace dos o tres semanas, varios "opinólogos" de la cadena que ustedes reciben en cables en Uruguay, como Fox News, sostenían curiosamente que había que imitar a Suecia. Curiosamente lo digo porque punto de vista las opiniones de Fox News, son exactamente que Suecia es país despreciable, son todos alcohólicos, se suicidan, nadie trabaja, son todos haraganes, etc. Oh casualidad, como Suecia optó por dejar que todo siguiera como estaba para que la economía, supone Estados Unidos, siguiera como estaba, y dijeron "no hay que lavarse las manos, no hay que usar máscaras, no practique el distanciamiento social". Y esto lo vemos día a día por ejemplo cuando salimos a ejercitarnos en la calle si una persona tiene mascarilla, la probabilidad de que no simpaticen con Trump es de un 80 %.

En Suecia ha habido una enorme discusión. La estrategia del gobierno fue criticada por gran parte de la comunidad científica y el propio epidemiólogo jefe en Suecia, ha reconocido que cometieron un error y Suecia ha pagado caramente esa protección de la economía.

Déjenme retrotraerme. Esto lo saqué del último número de The Economist de la semana anterior. En Suecia, según The Economist el Producto Bruto Interno en 2020, va a caer un 5 %, en Alemania 6 %, en Dinamarca un 4.9%, en Estados Unidos un 3.8%. Esto se lo dejo a los economistas para que lo decidan en forma más afinada, medir las consecuencias de las políticas. Yo me pregunto ¿Valió la pena matar tanta gente para proteger la economía de una manera que aparentemente no fue muy efectiva?

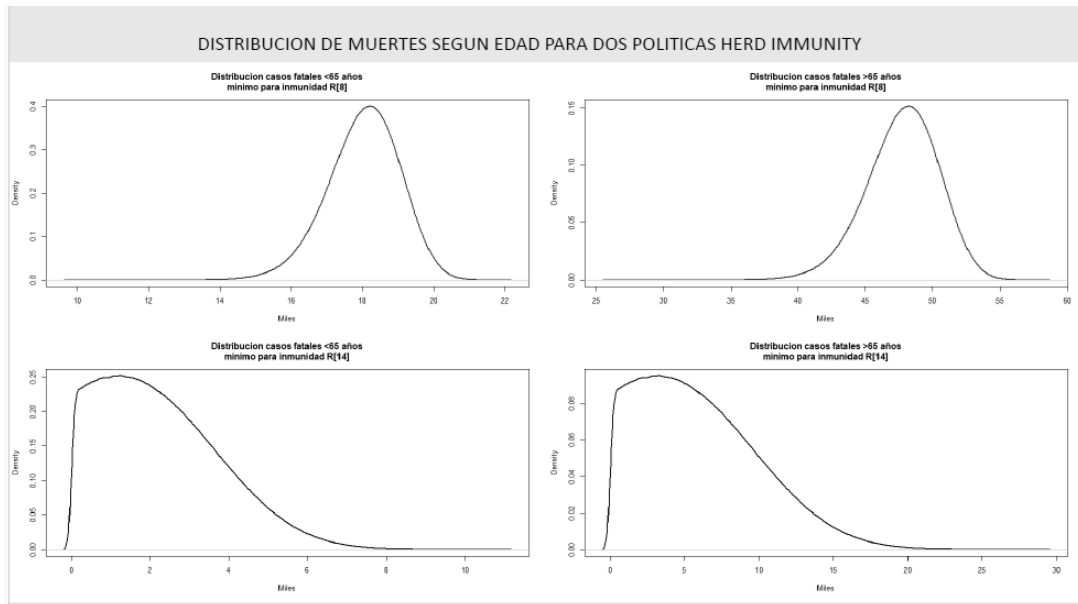
Mis pobres conocimientos de Economía me dicen que la mejor manera de proteger una economía es tratar de resolver la epidemia de manera drástica y poner todo en marcha, como aparentemente fue el caso de Wuhan en China. Perdonen por la digresión.

Recuerden lo que les dije: “si fuéramos Suecia”.



Vamos a ver qué se sabe hoy en día. Estas son las tasas de fatalidad, o sea cuál es la probabilidad de muerte si una persona contrae COVID 19 y eso depende de la edad. Para menores de 45 años es baja, no es nula, pero es baja. Pero en cambio, en la medida que nos internamos en capas etarias de más edad: 50, 60, 65, 70, estamos hablando de tasas de letalidad de un 15 a un 20%.

Entonces vamos a usar esta información y hacer con los siguientes supuestos: la demografía de Uruguay 85% es menor de 65 años, 15 % mayor de 15 años, la tasa de fatalidad 1% para menores de 65 años y 15% para mayores de 65 años.



Y esta es la distribución de probabilidad del número de muertes que hubieran ocurrido en Uruguay si hubiéramos adoptado una política de inmunidad de rodeo y si hubiéramos tomado las decisiones en el día 8 o en el día 14 de la epidemia.

Si todo funcionara de acuerdo a la lógica, en el día 8 hubiéramos tenido aproximadamente después de un año, 18.000 muertos menores de 65 años y aproximadamente 47.500 muertos mayores de 65 años.

Si hubiéramos basado nuestra decisión de inmunidad de rodeo en base al día 14, tendríamos entre 500 a 6 mil muertos menores de 65 años y entre 100 a 20.000 muertos mayores de 65 años.

Esto es simplemente una distribución de incertidumbre seguramente cambiando de los supuestos, introduciendo más granularidad en los cálculos, podríamos llegar a cifras más acertadas o más razonables. Pero simplemente ilustra qué tipo de juego se puede hacer con esta información que absolutamente crítica para tomar decisiones en tiempos de emergencia como son estos.

**“PUTTING THE RISK OF COVID-19 IN PERSPECTIVE” (New York City)
David Roberts, New York Times**

- “Micromorts” (Unidad, 1 en 1 millón chances de morir en un día)
- Día común y corriente (1 micromort)
- Anestesia general (5 micromorts por operación)
- Paracaidismo (7 micromorts por salto)
- Andar 70km diarios en motocicleta (11 micromorts)
- Combatir en Afganistan en 2010 (25 micromorts)
- Parto (210 micromorts)
- Trepas mas de 8 mil metros en el Himalaya (10 mil micromorts)
- Vivir en Michigan durante la epidemia 11 micromorts
- Vivir en NY durante la epidemia 50 micromorts adicionales
- Vivir en NY, ser “mayorcito, contraer COVID: ➔
ligeramente menos riesgoso que ser piloto de la RAF y bombardear Alemania en cuatro misiones durante la Segunda Guerra mundial

Ahora vamos a poner el riesgo en esta perspectiva. Esto lo saqué del New York Times de una persona que fue Scientific Advisor en una embajada de Estados Unidos, no recuerdo de que país, es físico, además. Usa una medida que llama micromorts, que es una chance en un millón, de morirse en un día y vamos a ver a qué micromorts estamos expuestos todos los días.

Cuando nos levantamos, a la hora del desayuno tenemos un micromort de chances de morirnos. Si es si ese día decidimos operarnos, tenemos que agregarle 5 micromorts, lo cual es un disuasivo para las cirugías plásticas si pueden evitar, en particular aquellas que requieran anestesia general.

Si resolvemos tirarnos en paracaídas tenemos que agregar los 7 micromorts. Imagínense: desayunamos, nos anestesiarnos y saltamos en paracaídas ya tenemos 13 micromorts, 13 en un millón chances de morirnos.

Si andamos 70 kilómetros diarios en una motocicleta le tenemos que agregar 11 micromorts y si vamos a defender la bandera de Estados Unidos y a combatir Afganistán agregamos de 25 micromorts.


Si decidimos dar a luz -yo simpatizo con "me too" pero no tanto como para alumbrar- agregamos 210 micromorts.

Si resolvemos irnos al Himalaya y trepar más de 8.000 metros, tenemos 10 mil micromorts adicionales.

Si vivimos en Michigan durante la epidemia 11 micromorts. Si vivimos en New York City, 50 micromorts. Pero si somos mayorcitos y contraemos el COVID, es ligeramente menos riesgoso que ser piloto del Royal Air Force y hacer cuatro misiones de bombardeo en Alemania durante la segunda guerra mundial.

Finalmente ¿Es que estamos en un Maracaná de la pandemia?

"Maracana de la pandemia"? (visto desde la distancia...)



- Se sabia que, tarde o temprano, SARS-COV-2 llegaria (equipamiento, camas CTI, etc.)
- Uruguay no es un "hub" (contrariamente a Milan, Frankfurt, NY City, etc.)
- Baja densidad demografica
- Decidida y rapida accion gubernamental. Alto nivel de testaje (regional). Control focos.
- Uso de evidencia y de asesoramiento cientifico ("Triunvirato", apoyado por un batallon de cientificos y medicos)
- Politica de switches, jugando el "partido en La Paz" (Radi et al.), sin demorar cambios (OWT)
- Consenso politico sobre las medidas (de "estado")
- Sistema de salud estructurado
- Excelente cobertura de atencion primaria
- Uruguayos: respondieron ante la emergencia
- NZ (islas) "erradico" el virus: 18 dias consecutivos sin ningun caso nuevo
- Jacinda Ardern no enfrenta el "riesgo Brasil"

Mi opinión personal y sé que este es un tema muy divisivo entre uruguayos, yo creo que lo de Maracaná fue fortuito. Si eso se hubiera repetido diez mil veces, creo que la mayor parte de las veces hubiera ganado Brasil. Me parece que es un punto de vista compartible es muy difícil de corroborar o refutar.

Afortunadamente no es así, no estamos ante un Maracaná de la pandemia, no es un milagro lo que está ocurriendo en Uruguay.

Primero, porque durante aproximadamente un mes estuvimos mirando por televisión y por la prensa que era prácticamente inevitable que el SARS COV 2 llegará a Uruguay, lo cual le dio al gobierno tiempo para prepararse: equipamiento, camas, CTI, etc. Tenemos la fortuna, nos quejamos de la falta de conectividad aérea, pero en este caso nos ayudó. Hay estudios estadísticos que han mostrado que lugares como Milán, Frankfurt, New York City, Los Ángeles, han sido focos de entradas del virus muy importantes. En Uruguay no tenemos muchos vuelos diarios y eso indudablemente, nos ayudó. Cuántas más ventanas abiertas más altas las probabilidades de que entre un ladrón.

Tenemos una baja densidad demográfica, hay departamentos de Uruguay en los cuales el COVID 19 no ha llegado, y en Estados Unidos finalmente llegó a todos los Estados. Si ustedes miran el caso de Montana, en Montana hay muy pocos casos de COVID 19, porque en Montana si uno va caminando por la calle es más fácil toparse con un bisonte que con una persona.

Creo que, mirado desde el exterior, que el gobierno actuó con decisión y con rapidez en un contexto regional donde tenemos un alto nivel de contagio, no tan alto como el de Perú o como el de Chile, afortunadamente con una situación mucho mejor, y creo que más allá de lo razonable, los focos que han ocurrido hasta ahora ha sido rápidamente y efectivamente controlados.

Una cosa que desde mi perspectiva sesgada como científico aplaudo a Uruguay, es por el uso de evidencia y de asesoramiento científico, hay un triunvirato muy visible de los Doctores: Radi, Cohen y Paganini que son los generales de una batalla que están siendo seguidos por un batallón de científicos y médicos que están atrás y de otros profesionales que apoyan esa acción.

Una cosa que ha sido creo muy interesante y muy inteligente de la acción uruguaya, es de no jugarse a una política decisiva con un *lock down*, sino que han manejado las perillas, los *switches* se mencionaron en algunas conferencias atrás y creo que la metáfora de Radi, el Dr. Radi de que vamos a tomar esto como si estuviéramos jugando el partido en La Paz a 4.000 metros de altura y pensando que vamos 0 a 0 y yo agregaría, pero sin demorar cambios, si es necesario hacerlos.

Saben que el presente es director técnico de la selección uruguaya, tema de discusión no es proclive a hacer los cambios en los momentos que se necesitan.

Una cosa muy importante y merece aplausos es el sistema político que en Uruguay ha funcionado formidablemente.

Y también vamos a no minimizar nuestras estructuras existentes, tenemos un sistema de salud formidable, nadie deja de ir al médico y tenemos excelente cobertura atención primaria.

Finalmente los uruguayos respondieron.

Ahora, hace dos o tres días la Primer Ministro, Jacinda Ardern de Nueva Zelanda anunció que el virus está erradicado. Tuvieron 18 días consecutivos sin tener ningún caso, Nueva Zelanda

ha vuelto a la vida normal pero no permiten que se ingrese ni que se egrese de la isla. Está considerando la construcción de un túnel aéreo con Australia.

Pero vamos a suponer que tengamos una secuencia de días sin casos, hay que recordar en Uruguay que la Primer Ministro Ardern es muy afortunada, porque no tiene a Brasil al lado.

Me voy a abstener de hacer comentarios de lo que pasa en Brasil.

Les agradezco mucho quiero terminar mi presentación con una pequeña celebración a la uruguaya.

Se expone un video de una cuerda de tambores uruguaya.

María Dolores Benavente.

Muchísimas gracias Daniel. Ahora le damos la palabra a Alicia.

Alicia Carriquiry

Voy a ser muy breve y voy a tratar de complementar la charla de Daniel, y hablar desde el punto de vista de cuáles son los desafíos.

Daniel presentó datos muy interesantes y buenas metodologías para hacer estimaciones epidemiológicas, pero voy a echar un poco de agua fría sobre lo que dijo Daniel en el sentido de que los sesgos, los errores, las incertidumbres que surgen de las fuentes de datos que tenemos para hacer estas estimaciones, pueden tener un impacto tremendo.

Antes de empezar, igual que Daniel quería reconocer que la gente en Uruguay está haciendo trabajo del primer nivel mundial: los Dres. Rafael Radi, Henry Cohen y Fernando Paganini, que Daniel ya nombró, y después un grupo de investigadores más que nada de la UdelaR, no sé sobre todos, pero sé sobre Daniel Herrera, Paola Bermolen, María Inés Fariello, que han producido un informe muy, muy interesante. A Ignacio, Nacho Álvarez que es un estadístico que se educó aquí y Juan Ignacio Sanguinetti, el hijo de un muy buen amigo mío que tendría que estar en Harvard haciendo un post doctorado, pero quedó varado en Uruguay por el problema del COVID.

Tanto sea en el ámbito científico, como en el ámbito político se ha hecho un muy buen trabajo y se nota en los resultados que tenemos.

Para discutir muy brevemente como les decía, los temas son estos:

- Las estimaciones que nosotros queremos hacer de los datos que tenemos disponibles
- Cuáles son las fuentes de información que se pueden utilizar,
- Algunos de los sesgos y otros errores e incertidumbres que se cuelan en estas fuentes de datos
- Impacto que estos sesgos y errores tienen y la mitigación de ese impacto.

Daniel habló del número básico de reproducción y eso es una de las estadísticas más básicas en epidemiología: R_0 , se calcula como la probabilidad de infección dado contacto. Primero que nada, el número básico de reproducción, el cálculo asume que hay una persona infectada y una población de personas todas susceptibles, así que no hay ninguna intervención. Entonces el cálculo de este número de reproducción consiste en el producto de tres cantidades, la probabilidad de infección dado contacto, el número promedio de contactos que la persona infectada tiene y por supuesto la duración del período de infeccioso de esa persona.

Y esencialmente lo que uno tiene es el número de casos secundarios, que se pueden atribuir a cada persona infectada.

También me voy a referir a este otro estadístico que se llama Case Fatality Rate (CFR) y el CFR el estimador más inocente digamos, es simplemente el número de muertes dividido por el número total de casos. Estos son los casos confirmados y estas son las muertes confirmadas de COVID.

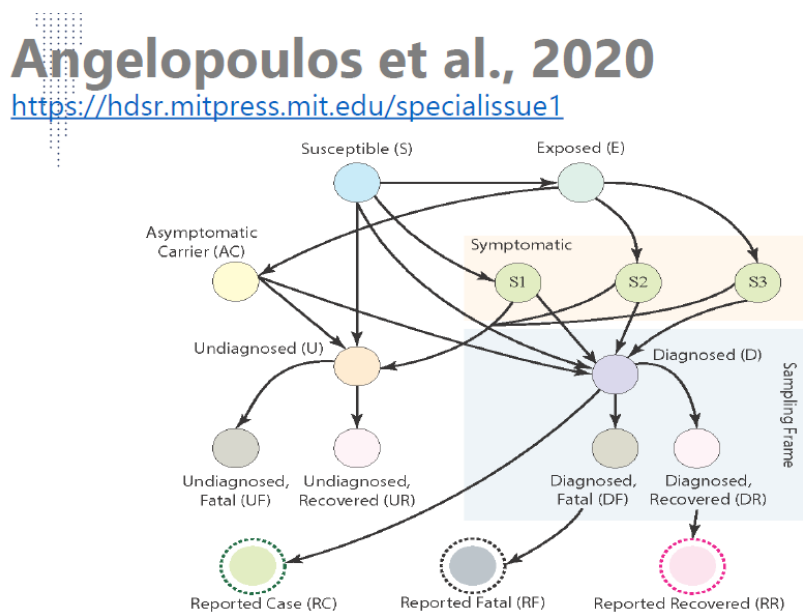
Por supuesto que hay otros, pero no voy a hablar de ellos hoy.

Esas cantidades las estimamos más que nada utilizando los “datos de vigilancia” y las fuentes más comunes de datos para los epidemiólogos son: el número de casos confirmados, el número de hospitalizaciones, los resultados de varios tipos de test, los test virológicos, los test serológicos para anticuerpos, las muertes atribuibles a COVID, lo que se llama las muertes excesivas relativas a períodos anteriores y luego las encuestas representativas de prevalencia, cuántos casos hay, cuántas personas infectadas hay verdaderamente en la población.

Estas son fuentes de datos que existen en Uruguay me imagino, que prácticamente se recogen en todos lados. El problema es que todas estas mediciones, todas estas fuentes de datos sufren de sesgos bastante importantes y de otros errores también, que no solamente son muy difíciles de estimar, sino que además son difíciles de mitigar.

Entonces la pregunta, la primera pregunta que uno se hace es ¿cuál es el impacto de esos sesgos y esos errores? Y en especial sobre estimadores de interés, como por ejemplo el número de base de reproducción y la tasa de mortalidad. El problema es que hay muchos de estos sesgos y a veces compiten.

Entonces es difícil saber por ejemplo si las estimaciones de estas cantidades terminan siendo demasiado altas o demasiado bajas y eso lo voy a discutir en un momentito, pero tampoco en mucho detalle.



Este es un esquema que pedí prestado de un de un *paper* que acaba de ser publicado creo que ayer, en esta publicación excelente que se llama el Harvard Data Science Review y que tiene un número especial dedicado a COVID, a la pandemia y que recomiendo muchísimo. Los *papers*

son muy accesibles, hay algunos que son poco accesibles, muy técnicos, pero este, por ejemplo, de Angelopoulos y colaboradores, que viene del laboratorio de Michael Jordan que capaz que Daniel sabe de quién hablo, en la Universidad de California, en Berkeley, es estupendo.

Aquí vemos un poco las complejidades del COVID famoso. Hay casos susceptibles en la población, algunos de ellos se exponen o son expuestos al virus. Los expuestos resultan en dos tipos de individuos sintomáticos, los S 2, que son las personas que demuestran síntomas característicos de COVID, pero también están, los S 3 que demuestran síntomas característicos de otras comorbilidades como por ejemplo problemas cardíacos o renales que, si bien están asociados al COVID, no aparecen como síntomas de COVID.

Después tenemos por supuesto personas S 1, que tienen síntomas como si fueran de COVID, pero en realidad no están afectados, tienen por ejemplo síntomas de influenza o de gripe, que pueden ser confundidos con COVID. El problema es la cantidad enorme de individuos que son asintomáticos. Este nodo en este modelo gráfico, de individuos asintomáticos, se compone de dos tipos individuos: las personas que nunca demuestran síntomas y las personas que tienen síntomas leves y que por lo tanto típicamente o muy a menudo no tienen exceso a tests.

En Estados Unidos el testeo ha sido muy restringido y sigue siendo restringido, en general, la gente que tiene acceso a tests y por supuesto a algún diagnóstico, es la gente en ciertas ocupaciones, como, por ejemplo, gente que trabaja en hospitales, o en lugares de faena de animales, donde los contagios han sido absolutamente galopantes. O personas que son fuertemente sintomáticas: para acceder a un test una persona tiene que estar muy enferma, lo cual inmediatamente sugiere que el número de casos diagnosticados, no son una muestra representativa de la gente que ha sido expuesta y que por supuesto ha sido infectada.

En todo caso, los errores y los sesgos a los que me voy a referir en un momento, afectan todos estos vértices, entre todos estos nodos y se van acumulando. Por ejemplo, el número de personas asintomáticas, que no reciben diagnóstico y que mueren sin diagnóstico, produce un sesgo hacia abajo del número de muertes debidas a COVID. Pero en otros casos, por ejemplo, los sesgos son hacia arriba porque como hay tantas personas que no han sido diagnosticadas, el número de casos confirmados se subestima, por lo que eso pone presión contraria, por ejemplo, a la estimación de la mortalidad.

¿Cuáles son las dificultades más típicas de COVID?

Hay un periodo asintomático o de síntomas leves, que es difícil de detectar. El virus es tremendamente infeccioso y fácil de transmitir de una persona a otra.

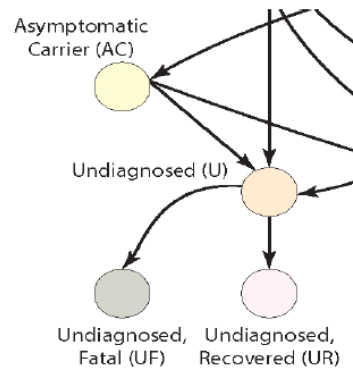
Los diagnósticos a veces son equivocados y se confunden los síntomas de influenza, gripe, otras comorbilidades con COVID. La muerte se atribuye a otras causas, por ejemplo, a gente que muere de problemas cardíacos, que están directamente relacionados al COVID, no necesariamente se clasifican como muertes debidas a COVID. Y Daniel habló un poco sobre la variabilidad en la severidad de los síntomas y de la enfermedad, dependiendo de las características de la población, las intervenciones y la adherencia a las intervenciones.

Así que los errores se cuelan en muchos lados, pero primeramente se cuelan porque tenemos muy, muy mala estimación sobre la proporción de personas que son infecciosas, pero asintomáticas. Porque este es uno de los nodos en la gráfica, la parte de izquierda que les mostraba, esa es gente que no se diagnostica y como no se diagnostica, las muertes tampoco

son atribuidas a COVID y la gente que se recupera tampoco es atribuida a COVID. En realidad, no sabemos, todo esto es desconocido.

Donde se cuean errores

- Casi completa ignorancia sobre la prevalencia de de infección (ausencia de tests en muestras aleatorias de la población).



forensicstats.org |

Esta ignorancia casi completa sobre la prevalencia de infección, se debe en gran parte al hecho de que en casi ningún país se han hecho esfuerzos para entender la prevalencia de la infección, haciendo tests en muestras aleatorias de la población.

En Estados Unidos, por cierto, no existe prácticamente en ningún lado un programa de test en muestras aleatorias de la población, que permitiría empezar a entender la proporción de personas asintomáticas pero expuestas.

Los errores también se cuean en los resultados de tests ya sea virológicos o de carga serológica. Los test, no sé en Uruguay, en Estados Unidos los test que se utilizan en general tienen lo que se llama buena especificidad: si la persona no está enferma, el test generalmente da negativo pero tienen baja sensibilidad, que quiere decir que cuando la persona está enferma a veces el test igual es negativo.

Los mejores tests en Estados Unidos en estos momentos tienen una sensibilidad del 80%, más o menos, quiere decir que hay un 20% de test negativos que son falsos.

Hay mucha demora entre el momento en que se hace el test y el momento en que se hace el diagnóstico, la confirmación del caso de COVID y el reporte de los casos de COVID, así que todos esos datos que estamos mirando, cuando vemos que hay 38 casos nuevos en un día tenemos que tener en cuenta que en realidad hay más que esos porque hay una demora bastante significativa a veces de semanas entre el momento en que una persona se testea y en el momento que se confirma ese caso.

Esto introduce sesgos en la estimación del número de casos confirmados, la estimación de número de casos diagnosticados que son fatales y el número de casos diagnosticados que se recuperan.

El acceso limitado de test ya lo mencioné, es una fuente de errores ya que obviamente cuanta menos gente se testea, menor es el número de casos confirmados, lo cual infla el número de fatalidades que se atribuyen a COVID.

Todos estos estadísticos, como, por ejemplo, la tasa de mortalidad, varían según las características del grupo. Así que cuando hablamos de tasa de mortalidad, o de número de reproducción, tenemos que tener en cuenta la edad del grupo al que nos estamos refiriendo, las comorbilidades, las otras cosas que hacen que la gente se enferme, la obesidad, si la gente fuma o no, presencia de diabetes, otras enfermedades.

Así que todas estas características tienen que ser tenidas en cuenta cuando uno toma una muestra aleatoria de la población para hacer el test y si la muestra de las personas testeadas no es representativa, en cuanto a edad, a obesidad, etc. en la población, entonces todos los estimadores, como por ejemplo R_0 ó CFR, van a estar sesgados en alguna dirección y no sabemos en qué dirección.

Finalmente, hay errores de clasificación y de atribución. ¿Cómo se define caso confirmado? ¿Cómo se define caso recuperado? ¿Cómo se define causa de muerte? En Wuhan por ejemplo, cuando cambiaron la definición de caso confirmado, el número de casos aumentó en un 600 % de un día para el otro, así que también hay que ser muy consecuente en cómo se utilizan esas definiciones.

Finalmente ¿cómo se mitigan esos impactos?

- Lo que se requiere es mejor información, por ejemplo, si quiero saber cuántas personas son expuestas en una población de susceptibles, tengo que tener información sobre la adherencia a las normas, a la ocupación, el tamaño del grupo familiar, los patrones del transporte, esos datos de los cuales Daniel ha estado hablando anteriormente, microdatos que no necesariamente son datos epidemiológicos, pero que sí son necesarios para entender la transmisión de la enfermedad.
- El número de personas asintomáticas: ya lo ya lo dije. La manera de entender un poco cuál es ese número, es únicamente a través de muestra aleatoria y haciendo por supuesto el *testing* mucho más accesible de lo que es, al menos en Estados Unidos, en Uruguay no estoy segura cuán accesibles son los tests.
- Y por cierto el uso de estimadores más sofisticados: los estimadores que yo le mostré son los más los obvios, donde se asume que los datos no tienen ningún error, pero hay estimadores mucho más sofisticados que permiten compensar estas limitaciones que tienen todos los datos de vigilancia que están disponibles.

El estimador de R por ejemplo que presentó Daniel, tiene algunas de estas características, no sé cómo toma en cuenta por ejemplo, los errores de medición en muchas de estas bases de datos.

En realidad, es como medio deprimente saber que los datos son tan deficientes. Pero no quiero que la gente se deprima mucho, simplemente quiero enfatizar el hecho de que hay que ser muy cauteloso en la interpretación de los datos, hay que entender las limitaciones de la información y de los estimadores de las cantidades de interés y utilizar las metodologías más sofisticadas a las que se tenga acceso.

Por suerte hay muchos recursos bibliográficos últimamente; vuelvo a recomendar este número especial de Harvard Data Science Review tiene artículos de todo nivel y muy interesantes.²

² <https://hdsr.mitpress.mit.edu/specialissue1>

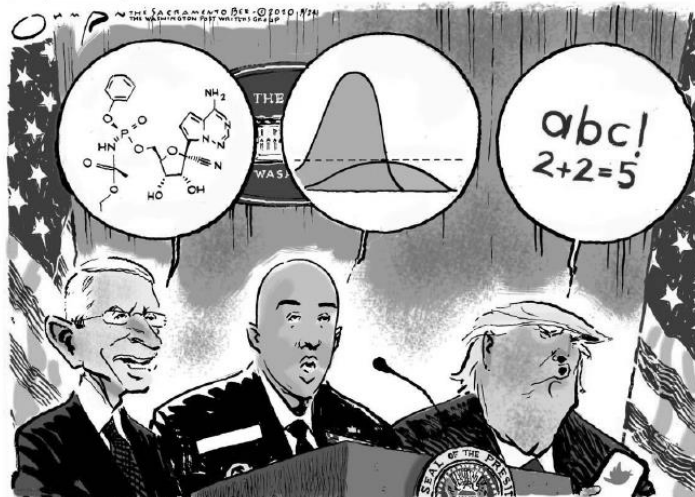
Daniel es más políticamente cauteloso, yo no. En Estados Unidos tenemos el peor gobierno del Universo, quizás le estamos haciendo competencia a Brasil. Pero se ha politizado muchísimo la pandemia en Estados Unidos, están los Demócratas, por un lado, los Republicanos por el otro. los Republicanos que siguen a Trump, Trump que todavía piensa que la pandemia se va a terminar por milagro, no sé por qué. Hay controversias hasta por el uso de las máscaras.

Máscaras!



Me divierte mucho este *cartoon* y el último: más ciencia menos política por favor.

Más ciencia, menos política



De nuevo en Estados Unidos hay gente espectacularmente competente, nosotros tenemos aquí centros como el CDC (Centers for Disease, Control and Prevention), que solía ser el modelo mundial para este tipo de problemas, pero gracias al liderazgo de la administración de Trump, el CDC ha sido dejado de lado y en estos momentos estamos a la deriva, como bien dijo Daniel. Veremos qué pasa en este país.

Si alguien desea hacerme más preguntas, esta es mi dirección: Alicia@iastate.edu

Muchas gracias

Isidoro Hodara

Abrimos un espacio de preguntas.

Pregunta del público

Si fuéramos una isla... pero somos vecinos de Brasil, por lo tanto es muy difícil de predecir. Qué predice tu modelo si es posible y en cuánto tiempo se podría lograr los resultados como Nueva Zelanda, de eliminar el virus.

Daniel Gianola

El modelo considera una población homogénea y cerrada. En principio si observáramos un número suficientemente largo de días, sin ningún tipo de casos, el modelo empezaría a enflaquecer la distribución predictiva sobre cero, pero probablemente siempre va a haber algo de fluctuación aleatoria, a menos que el número de días se convierta en infinito y que cada uno de esos días tenga cero casos, en ese caso se colapsaría a cero. De cualquier manera, podemos calcular las probabilidades de tener tantos o más días sin casos y tomar una decisión sobre la base de esa regla. Por supuesto con un nivel de granularidad más alto, pero sí que nos puede guiar a tomar una decisión, por ejemplo, si decidimos que la probabilidad de tener un caso después de 20 días sin casos, es menos de 1 en 50 mil, por ejemplo, lo digo por decir, tal vez podríamos decir "sí, vamos a abrir la economía y volver a la normalidad".

En principio este modelo y otros muchos más sofisticados pueden ser utilizados para tomar una decisión racional, pero siempre va a haber un poco de incertidumbre.

Alicia Carriquiry

Pienso que lo que Nueva Zelanda ha hecho no es eliminar el virus, el virus existe siempre, va a existir para el resto de nuestra vida: si llega a aparecer un neozelandés que estuvo de viaje en Europa, capaz que vuelve a reintroducir el virus en Nueva Zelanda. La cuestión es cómo compensar riesgos con beneficios, hasta que tengamos una vacuna o alguna otra solución.

Daniel Gianola

De acuerdo. Yo usé el término de erradicación en el sentido literal, que lo usó el Primer Ministro de Nueva Zelanda, pero siempre hay riesgos latentes de reintroducción del virus y es posible también que convivamos con ese virus como hemos convivido con muchos otros, sin que importe a la gran y feliz familia, siempre que R sea lo suficientemente bajo, que tengamos una vacuna y que haya tratamientos porque todavía no los hay.

Un comentario que quería hacer sobre lo que decía Alicia: es muy interesante que, en estas epidemias, los organismos oficiales tienen que actuar sobre la base de la información disponible. Y la información disponible a la vez sugiere sus limitaciones y sugiere cosas que se deberían hacer. Y muchas veces, lamentablemente, esas cosas que se deberían hacer se relegan exclusivamente al ámbito académico, con lo cual se desarrollan modelos, necesidades, corrección de los sesgos, etc. Y en ese sentido es muy importante lo que decía Alicia: "más ciencia menos política", porque la ciencia puede guiar a la acción pública y a la acción pública de buena calidad.

Pregunta del público

Gracias Daniel por el panorama de modelaje y a Alicia por el llamado a la cautela con el tema de los datos porque al final de cuentas, por más sofisticadas que sean las herramientas, las herramientas son tan buenas como los datos que introducimos en ellas y sin ninguna duda los datos tienen limitaciones y lo más desconcertante es que no sabemos en qué sentido son estas limitaciones, si estamos sobre estimando o sub estimando casos y fatalidades.

Me pregunto Daniel, tú has usado los datos como están presentados para estimar el R y el Case fatality rate. Los datos son lo que son y no sabemos qué rango de sesgo y de errores estamos mirando. ¿Valdría la pena hacer un *worst case scenario* y un *best case scenario*, para tener una idea de qué predicciones esperar si asumimos un peor caso de sub estimación o sobre estimación? Dado que realmente no sabemos dónde estamos parados con los datos.

Daniel Gianola

La persona que acaba de hablar es la Dra. Nora M. Bello, que, si notaron el acento cordobés, es profesora de estadística en Kansas State University, un gusto de verte. Nora como bien tú sabes, no soy epidemiólogo y me limité simplemente a utilizar datos básicos. Yo he bajado las bases de datos a casos diarios y casos de muerte, así nada más. Tú me estas planteando considerar escenarios y naturalmente sí que se podría hacer, para examinar la sensibilidad.

Personalmente y esto lo hemos discutido muchas veces, si me permiten, voy a hacer una pequeña digresión: en estadística, en general en ciencias, hay dos problemas: un problema inductivo y un problema predictivo.

El problema inductivo no tiene solución, porque queremos estimar cosas que no observamos, intangibles, que en realidad no existen, son entelequias, los parámetros son construcciones intelectuales que no existen. La forma en que abordamos esto es que construimos un modelo que le damos un valor existencial a los parámetros, y después usamos criterios condicionalmente a que el modelo sea correcto, que el modelo tenga un grado de plausibilidad, cuán cerca estamos de ese valor que no conocemos pero que tenemos que suponer cuál es. Entonces, eso nos lleva a simular escenarios y ver qué pasa con las cosas si uno varía para aquí o para allá. Lo que no sabemos es cuál es el grado de plausibilidad de ese escenario.

El segundo enfoque que ha tenido un enorme éxito recientemente, que es el enfoque para áreas que tienen incapacidad de generar experimentos randomizados, es el de hacer un enfoque clasificatorio predictivo. Uno siempre puede calibrar la calidad predictiva del modelo o la calidad clasificatoria. Después ve si las predicciones fueron buenas o malas.

Eso, naturalmente, no nos permite elaborar con respecto a una teoría porque en la teoría generalmente esos métodos no tienen la estructura paramétrica, pero sí nos permite especular sobre qué información adicional podríamos obtener para conseguir mejores predicciones o clasificaciones.

El enfoque predictivo sí tiene una respuesta. Naturalmente con las epidemias uno no puede esperar que se muera demasiada gente para ver si las cosas anduvieron bien, pero ese es el tipo de respuesta que en muchos ámbitos intentamos hacer, o ver, en problemas complejos.

Porque recordemos que tanto los modelos que usé como algunos de los refinamientos que Alicia sugiere, son cosas que tienen una estructura sumamente simple, pero que requieren supuestos sumamente fuertes. Ante la complejidad, hacer supuestos fuertes a utilizar, es un acto que tiene sumo riesgo.

En fin, no tengo una respuesta, es sensato, es factible, es prudente hacer estudios de sensibilidad pero cuál es la plausibilidad, es muy difícil decir a menos que lo cotejemos con su capacidad predictiva y si esa capacidad predictiva es buena, creo que el modelo funcionalmente e interpretativamente es útil.

Alicia Carriquiry

Estoy de acuerdo con lo que dijiste Daniel. Yo uso muchos modelos predictivos también, pero hay que tener un poco de ojo con los modelos predictivos. Porque uno de los supuestos de los modelos predictivos, es que uno parte de datos de los cuales uno conoce la verdad. Todos los modelos predictivos tienen que ser “entrenados” sobre datos para los cuales donde tú conoces, por ejemplo, que este caso pertenece a esta categoría y este otro caso pertenece a esta otra categoría. Y en el caso de una pandemia como el COVID, por ejemplo, donde tenemos una cantidad enorme de información faltante, casos que no observamos, como por ejemplo los casos asintomáticos o infecciosos, es muy difícil, creo yo, hacer predicción con confiabilidad. Pienso que una de las cosas que verdaderamente faltan en esta pandemia, en Uruguay y en otros lados y que yo haría como actividad número uno, es tratar de estimar la prevalencia de base, tratar de hacer un muestreo aleatorio representativo de la población y administrar tests y ver exactamente cuál es la proporción de personas infectadas. Con esos datos bien estimados, creo que el resto se puede corregir.

Daniel Gianola

Los que seguimos al fútbol, conocemos un comentarista que se llama Kesman, que lo conocen todos y que tiene una famosa frase: “es lo que hay valor”.

Entonces, en un momento determinado, tenemos que tomar decisiones, tenemos que hacer predicciones. Esas predicciones nos ilustran sobre la insuficiencia de los datos existentes y nos hacen reflexionar sobre qué información adicional deberíamos tener y por qué falta información, cuáles son los sesgos cuando hay datos faltantes.

Como tú muy bien sabes es el pan nuestro de cada día, en selección animal. Y es verdaderamente un monstruo difícil de manejar, pero a la vez, para incorporar esos faltantes, hay que proponer un modelo de por qué faltan los datos y esos modelos típicamente deben de contener errores, que introducen nuevos errores.

De manera que, en este momento, lo que me planteé como epidemiólogo casero es decir ¿Qué puedo hacer con la información que tengo? Y otros como tú o como el batallón de personas que están trabajando sobre el COVID 19 de la UdelaR, está haciéndose preguntas más finas como, por ejemplo, ¿Estamos subestimando los casos? ¿Cuál es el porcentaje de contagiados en la población? ¿A qué ritmo aumenta el contagio? Curiosamente el nivel de contagios en New York City y en Milán es bajísimo, se ha estimado en 10, 15% de las personas a pesar de haber tenido miles y miles de casos y muertes. Aparentemente el virus contagia rápidamente en un sentido, pero cuando se distribuye al azar en la población, lo hace a un ritmo más lento.

Yo no tengo la base epidemiológica como para responder a eso.

Pregunta del público

Acá en Kansas State, como me imagino en Iowa o en Wisconsin, acá nos ha tapado todo el tema del *teaching on line* y uno quisiera tener más tiempo para meterse más en la parte cuantitativa de esto.

Me pregunto, ustedes que están en el tema, si han escuchado o si se está implementando el tema del testeo en grupo como ha sido propuesto.

Me refiero a *Group testing*. Distintos tipos de metodologías estadísticas que han salido en los últimos dos o tres años de University of Nebraska, que tienen que ver con aumentar el testeo y en forma ideal hacer un testeo de prevalencia de base en la población, usando sistemas de testeo en grupos, para que no sea una cosa tan onerosa, tan cara y disminuir la resistencia para lograr este estadístico de base, que hace falta para corregir o interpretar todo el resto de los datos.

Daniel Gianola

En Uruguay se está haciendo, se llama *pull testing*, hace 19 años cuando se empezó la selección genómica en animales, las pruebas de ADN eran sumamente caras y entonces lo que hacíamos era tomar una batería de toros y hacíamos una sola prueba y eso tenía la ventaja de que se perdía variabilidad, pero se ganaba en precisión. Porque al tomar la muestra de un grupo es como extraer un promedio.

De manera que eso se está haciendo en Uruguay y probablemente se esté haciendo en otros lados también.

Alicia Carriquiry

La pregunta es ¿Cuál es la parte más cara del test? Si es la parte del laboratorio o si es la parte de tomar la muestra de la persona y no estoy segura para COVID cuál es la parte más cara.

En los *pull testing* que se hacen en problemas ambientales, donde querés sacar muestras de un lago y en vez de analizar 400 muestras diferentes, sacas muchas muestras las mezclas y después haces un solo test sobre la muestra. Pero no estoy segura que eso sea aplicable a COVID donde la parte cara es tomar a la persona, sacarle la muestra, preservar esa muestra y llevarla al laboratorio.

Isidoro Hodara

Lo primero que tenemos que hacer es agradecer a Daniel y a Alicia, por la claridad con la que nos transmitieron un tema que ciertamente es complejo y por la generosidad que tuvieron en compartir con nosotros sus conocimientos.

También y eso es más personal, agradecer por la caricia a nuestro orgullo como uruguayos que ustedes nos han hecho, por el nivel con que han hecho sus presentaciones.

Para todos quiero recordarles que tanto las presentaciones como el grabado de esta sesión van a estar disponibles en la página de la Academia Nacional de Economía.

Muchas gracias.

Daniel Gianola

Muchas gracias.

El mensaje final es que tengan en cuenta que acá hay dos ingenieros agrónomos que están hablando de epidemiología, así que tomen todo esto con pinzas.