

پیش‌بینی‌هایی در ارتباط با ویروس جدید کرونا (کووید-۱۹)

وحید رضائی‌تبار^۱

تاریخ دریافت: ۱۳۹۹/۰۸/۱۴

تاریخ پذیرش: ۱۴۰۰/۱۲/۲۶

چکیده:

در پایان دسامبر ۲۰۱۹، گسترش یک بیماری عفونی جدید در شهر ووهان چین گزارش شد که توسط یک کروناویروس جدید ایجاد شده و به‌طور رسمی توسط سازمان بهداشت جهانی به‌عنوان کووید-۱۹ نام‌گذاری گردید. با عبور تعداد قربانیان این ویروس از مرز ۱۰۰۰ نفر، سازمان بهداشت جهانی برای بیماری ناشی از آن نام رسمی کووید-۱۹ را انتخاب کرد که اشاره‌ای به «کرونا»، «ویروس»، «بیماری» و سال ۲۰۱۹ دارد. پیش‌بینی‌هایی در ارتباط با کووید-۱۹ می‌تواند به دولت در تصمیم‌گیری‌های بهتر کمک کند. در این مقاله، از یک رویکرد عینی برای پیش‌بینی‌هایی در ارتباط با کووید-۱۹ مبتنی بر روش‌های آماری استفاده شده است. تحلیل‌ها با این فرض انجام شده‌اند که داده‌های ارائه شده قابل اعتماد هستند. مهم‌ترین هدف، پیش‌بینی روند شیوع ویروس کرونا برای تعداد موارد مبتلای قطعی، فوتی و بهبود یافته و تخمین مدت‌زمان مدیریت این ویروس است که از روش هموارسازی نمایی برای پیش‌بینی استفاده شده است. این خانواده از مدل‌ها، به‌خصوص برای سری‌های زمانی کوتاه‌مدت بسیار مناسب است. هموارسازی نمایی به زبان ساده نوعی میانگین متحرک است که خود را اصلاح می‌کند. در نهایت با توجه به پیش‌بینی آماری صورت گرفته، به ارائه راهکارهایی می‌پردازیم.

واژه‌های کلیدی: کووید-۱۹، پیش‌بینی، رویکرد عینی، هموارسازی نمایی.

۱ مقدمه

که ما را به احتمال خیلی قوی، مخزن کووید-۱۹ هستند. مطالعه بنوناتو و همکاران (۲۰۲۰) نشان داد که کووید-۱۹ خیلی با کروناویروس جدا شده از نوع خاصی از خفاش‌های چینی مرتبط و نزدیک هستند [۲]. تحقیقات آن‌ها از این نظریه پشتیبانی می‌کند که زنجیره انتقال از خفاش‌ها به سمت انسان شروع شده است. چان و همکاران (۲۰۱۹) تأیید کردند که کووید-۱۹ یک کروناویروس جدید بوده که با کروناویروس سارس خفاش بسیار مرتبط است [۳]. اخیراً ژو و همکاران (۲۰۲۰) دریافتند که همسانی توالی بین کووید-۱۹ و کروناویروس سارس برابر ۵/۷۹٪ است [۴]. همچنین آن‌ها دریافتند که کووید-۱۹ با کروناویروس خفاش دارای هومولوژی بالایی است. بنابراین شواهد فعلی به‌شدت تأیید می‌کنند که کووید-۱۹ از خفاش‌ها به‌دست آمده است، اگرچه میزبان‌های میانی یا واسطه‌ای کووید-۱۹ هنوز مشخص نیستند.

در دسامبر ۲۰۱۹ برای اولین بار در شهر ووهان استان هوئی چین، پس از این‌که مردم بدون علت مشخصی دچار سینه‌پهلو شدند و واکسن‌ها درمان‌های موجود مؤثر نبودند، نوع جدیدی از کروناویروس با همه‌گیری در انسان شناسایی شد. با عبور تعداد قربانیان این ویروس از مرز ۱۰۰۰ نفر، سازمان بهداشت جهانی برای بیماری ناشی از آن نام رسمی کووید-۱۹ را انتخاب کرد که اشاره‌ای به «کرونا»، «ویروس»، «بیماری» و سال ۲۰۱۹ دارد. ویروس‌های کرونا گروه بزرگی از ویروس‌ها هستند که می‌توانند حیوانات و انسان‌ها را آلوده کنند و باعث بروز ناراحتی‌های تنفسی شوند. این ناراحتی‌ها ممکن است به‌اندازه سرماخوردگی خفیف یا به‌اندازه ذات‌الریه شدید باشند. مطالعه‌ای نشان داده است که کووید-۱۹ یک ویروس نوترکیب بین کروناویروس خفاش و کروناویروس با منشأ ناشناخته است [۱]. آن‌ها با مقایسه با حیوانات دیگر، دریافتند

۲ پیش‌بینی موارد قطعی

برای پیش‌بینی موارد ابتلای قطعی به کووید-۱۹، از روش‌های پیش‌بینی در سری‌های زمانی استفاده می‌شود. برای این منظور پیش‌بینی با استفاده از مدل‌هایی از خانواده هموارسازی نمایی انجام می‌شود [۵]. این خانواده از مدل‌ها، به‌خصوص برای سری‌های زمانی کوتاه‌مدت بسیار مناسب است. هموارسازی نمایی به زبان ساده نوعی میانگین متحرک است که خود را اصلاح می‌کند. امروزه تقریباً نمی‌توان حوزه‌ای را یافت که از پیش‌بینی آینده بی‌نیاز باشد و حتی در بسیاری از موارد نیز پیش‌بینی اهمیت حیاتی دارد. در بسیاری از حوزه‌ها مانند صنعت، اقتصاد، سیاست، محیط‌زیست، سلامت، فرهنگ و غیره تلاش می‌گردد تا به ابداع و استفاده از روش‌هایی مناسب و همچنین استفاده از داده‌های مربوط به حوادث و تجربه‌های گذشته، مدلی تهیه گردد که بتواند آینده نزدیک یا دور را پیش‌بینی نماید. به عبارت دیگر، رفتار یک پدیده در گذشته به کمک مدل‌های آماری و ریاضی، در آینده برون‌یابی می‌شود و بر اساس معیارهایی تعریف‌شده از خطا، ارزیابی می‌گردد تا مشخص شود تا چه اندازه این روش می‌تواند مؤثر و مفید واقع گردد. روش‌های کیفی و کمی بسیار متنوعی در علم پیش‌بینی وجود دارد که یک گروه از روش‌های مهم در پیش‌بینی، روش‌های هموارسازی نمایی است که از مزیت‌های این روش می‌توان به سادگی و سرعت انجام محاسبات در آن اشاره کرد.

۳ مدل هموارسازی نمایی

هموارسازی به‌منظور کمک به شناخت بهتر الگو و روند در سری‌های زمانی انجام می‌شود و موجب حذف نوفه (اثرات تصادفی) می‌گردد. در یک تقسیم‌بندی کلی، روش‌های هموارسازی را می‌توان به دو گروه روش‌های میانگین متحرک و روش‌های هموارسازی نمایی دسته‌بندی کرد. هموارسازی نمایی از جمله روش‌های آماری پرکاربرد برای پیش‌بینی بر مبنای سری‌های زمانی می‌باشد و ایده‌ی آن بر این اساس است که مشاهدات اخیر معمولاً بهترین راهنمایی را برای آینده فراهم خواهد کرد. بنابراین وزن بیش‌تری به مشاهدات اخیر می‌دهد و همچنان که مشاهدات دورتر می‌شود وزن‌ها به‌صورت نمایی کاهش می‌یابد. این روش کلاسیک برای هر دو مورد سری‌های زمانی واریانس همسان و واریانس ناهمسان به کار می‌رود. در حالت واریانس همسان تخمین مدل برابر مدل خودهمبسته میانگین متحرک جمع بسته می‌باشد اما در حالت واریانس ناهمسان این‌گونه نیست. بنابراین هموارسازی نمایی

می‌تواند در رده‌ی گسترده‌تری از مدل‌ها نسبت به مدل آریما مورد استفاده قرار گیرد. انواع مختلفی از مدل‌های هموارسازی نمایی برای حالت‌ها و شکل‌های مشخصی از سری‌های زمانی استفاده می‌گردد که در ادامه به توضیح هر یک از آن‌ها می‌پردازیم.

۱.۳ هموارسازی نمایی تکی

این روش برای پیش‌بینی سری‌های زمانی استفاده می‌شود که روند و همچنین الگوهای فصلی در آن وجود نداشته باشد و معادله‌ی آن به‌صورت معادله‌ی ۱ می‌باشد:

$$S_t = \alpha Y_t + (1 - \alpha)S_{t-1}, 0 \leq \alpha \leq 1$$

$$\hat{Y}_t = S_{t-1} \quad (1)$$

که در آن:

S_t : مقدار هموارشده در زمان t ،
 α : پارامتر هموارسازی (مشخص‌کننده نرخ تغییر وزن)،
 Y_t : مقدار واقعی سری در زمان t ،
 $\hat{Y}_t$: مقدار برازش شده در زمان t است.

انتخاب وزن‌های کمتر (نزدیک به صفر) موجب هموارسازی بیش‌تر و از بین بردن اجزای دوره‌ای و نامنظم می‌گردد. از طرف دیگر وزن‌های بزرگ‌تر (نزدیک به یک) سبب هموارسازی کمتر شده و باعث می‌شود خروجی مدل واکنش بیش‌تری نسبت به تغییرات نشان دهد. بنابراین بهترین راه برای انتخاب دقیق آن این است که مقدار α با استفاده از آزمون و خطا طوری انتخاب شود که میانگین توان دوم کمینه گردد.

همچنین برای محاسبه‌ی تخمین اولیه از مقدار هموارشده‌ی سری در زمان $t = 0$ (یعنی S_0) داریم:

$$S_0 = \bar{Y} = \frac{\sum_{t=1}^n Y_t}{n}, n \leq 6 \quad (2)$$

علت به کار بردن کلمه‌ی نمایی در مدل‌های هموارسازی نمایی نیز بدین شرح است که طبق رابطه‌ی ۱ داریم:

$$S_t = \alpha Y_t + (1 - \alpha)S_{t-1}$$

$$= \alpha Y_t + (1 - \alpha)[\alpha Y_{t-1} + (1 - \alpha)S_{t-2}]$$

$$S_t = \alpha Y_t + (1 - \alpha)\alpha Y_{t-1} + (1 - \alpha)^2 S_{t-2}$$

$$= \alpha Y_t + (1 - \alpha)\alpha Y_{t-1} + (1 - \alpha)^2 \alpha Y_{t-2} + \dots +$$

$$(1 - \alpha)^{t-1} \alpha Y_1 + (1 - \alpha)^t S_0 \quad (3)$$

در پیش‌بینی دارد. حالت $\gamma = 0$ نیز نشان‌دهنده‌ی ثابت بودن جزء روند در بین تمام مقادیرهای سری زمانی و همچنین پیش‌بینی بوده و در حالت $\gamma = 1$ جزء روند به‌طور حداکثری از مشاهده‌ای به مشاهده‌ی دیگر تغییر می‌کند. در اینجا نیز بهترین روش به‌منظور انتخاب دقیق دو پارامتر هموارسازی (γ, α) این است که ترکیب‌های مقادیرهای مختلف آن‌ها با استفاده از آزمون و خطا طوری انتخاب شود که میانگین توان خطا کمینه گردد. محاسبه سطح و روند اولیه در زمان $t = 0$ (ت. و ل. ۱) نیازمند برآورد مدل رگرسیون خطی برای داده‌های سری زمانی موردنظر (متغیر Y) در برابر زمان (متغیر X) است. در این حالت مقدار ثابت (عرض از مبدأ) این رگرسیون تخمین اولیه (آغازین) از جزء سطح (L_0) بوده و شیب نیز تخمین اولیه از جزء روند (T_0) می‌باشد.

۳.۳ هموارسازی نمایی سه‌گانه

این روش هنگامی استفاده می‌گردد که سری زمانی داده‌ها با روند و یا بدون روند اما شامل الگوهای فصلی باشد. اگر نمودار سری نشان‌دهنده‌ی الگوهای تکرارشونده‌ای در طول دوره‌های زمانی یک‌ساله به سبب تأثیرات فصلی باشد، الگوی فصلی نامیده می‌شود. این حالت نیاز به سومین پارامتر هموارسازی (δ) برای محاسبه‌ی جزء فصلی دارد. مدل هموارسازی نمایی سه‌گانه دارای دو شکل است: الف) شکل حاصل ضربی، در این حالت سری زمانی با تغییرات فصلی افزایشی استفاده می‌شود و معادله‌ی آن به‌صورت ۵ در زیر است:

$$\begin{aligned} L_t &= \alpha \left(\frac{Y_t}{S_t - p} \right) + (1 - \alpha)[L_{t-1} + T_{t-1}], 0 \leq \alpha \leq 1 \\ T_t &= \gamma[L_t - L_{t-1}] + (1 - \gamma)T_{t-1}, 0 \leq \gamma \leq 1 \\ S_t &= \delta \left(\frac{Y_t}{L_t} \right) + (1 - \delta)S_{t-p}, 0 \leq \delta \leq 1 \\ \hat{Y}_t &= (L_{t-1} + T_{t-1})S_{t-p} \end{aligned} \quad (5)$$

که در آن:

L_t : سطح سری در زمان t ,

α : پارامتر وزن مربوط به سطح،

T_t : روند در زمان t ,

γ : پارامتر وزن مربوط به روند،

S_t : جزء فصلی در زمان t ,

δ : پارامتر وزن مربوط به جزء فصلی،

p : تعداد فصل‌ها در یک سال (دوره فصلی) $[p = 12]$ برای داده‌های

ماهانه و $p = 4$ برای داده‌های سه‌ماهه (فصلی)]،

$\hat{Y}_t$: مقدار واقعی سری در زمان t ,

مشاهده می‌شود مقادیر سری وزن‌های ایجادشده برای پیش‌بینی در زمان t یعنی مقادیر $\alpha, (1 - \alpha)\alpha, (1 - \alpha)^2\alpha, \dots$ هنگامی که مشاهدات به عقب می‌رود به شیوه‌ی نمایی به سمت صفر کاهش می‌یابد.

۲.۳ هموارسازی نمایی دوگانه

هموارسازی نمایی دوگانه با عنوان هموارسازی نمایی با روند نیز شناخته می‌شود. این روش هنگامی به کار می‌رود که مقادیر سری زمانی داده‌ها دارای روند ثابت باشند اما الگوهای فصلی در آن مشاهده نگردد. به بیان دیگر این روش وقتی کارا است که داده‌ها تنها دارای روند هستند. اگرچه سری‌های زمانی به‌طور معمول شامل نوسانات تصادفی است، اما ممکن است مقادیر سری نشان‌دهنده‌ی حرکات و انتقال‌های تدریجی به مقادیری نسبتاً بالاتر یا پایین‌تر در طول یک مدت‌زمان طولانی باشد و در حالتی که نمودار سری چنین رفتاری را نشان دهد نشان از وجود الگوی روند در آن است. اگر نمودار سری نشان‌دهنده‌ی روند باشد می‌توان انتظار داشت که این روش بهتر از هموارسازی نمایی تکی خواهد بود. بنابراین در صورت وجود روند، هموارسازی نمایی ساده نیاز به برخی تعدیلات و اضافه کردن دومین پارامتر هموارسازی و متعاقباً معادله‌ی هموارسازی دیگری با پارامتر γ برای محاسبه روند دارد. این روش دارای یک جزء سطح و یک جزء روند در هر مرحله و در نتیجه دو پارامتر هموارسازی (α, γ) است و معادله‌ی آن به‌صورت معادله ۴ می‌باشد:

$$\begin{aligned} L_t &= \alpha Y_t + (1 - \alpha)[L_{t-1} + T_{t-1}], 0 \leq \alpha \leq 1 \\ T_t &= \gamma[L_t - L_{t-1}] + (1 - \gamma)T_{t-1}, 0 \leq \gamma \leq 1 \\ \hat{Y}_t &= L_{t-1} + T_{t-1} \end{aligned} \quad (4)$$

که در آن:

L_t : سطح در زمان t ,

α : پارامتر وزن مربوط به سطح،

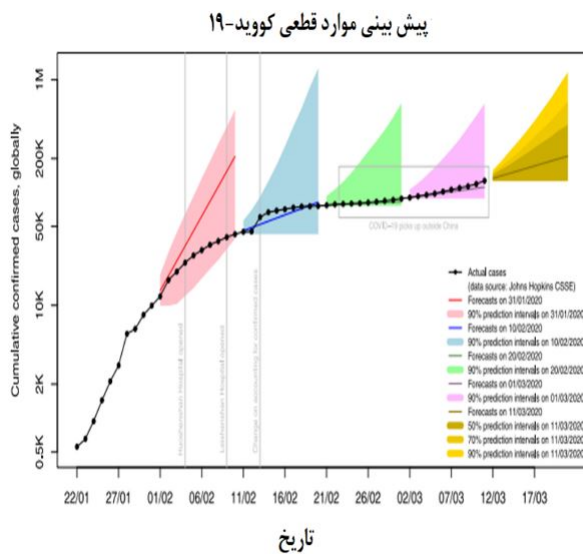
γ : پارامتر وزن مربوط به روند،

Y_t : مقدار واقعی سری در زمان t ,

$\hat{Y}_t$: مقدار برازش شده در زمان t (یا پیش‌بینی یک گام جلوتر در زمان t) هستند. رابطه‌ی سطح نشان‌دهنده‌ی تخمین هموارشده (یکنواخت شده)

از مقادیرهای داده‌ها در پایان هر دوره است و معادله‌ی روند نیز تخمین هموارشده‌ای از میانگین رشد در پایان هر مرحله می‌باشد. انتخاب

مقادیر کوچک α بدین مفهوم است که مشاهدات آخر نقش کم‌تری در پیش‌بینی نسبت به سایر مقادیر دارد و بالعکس، انتخاب مقادیرهای نسبتاً بزرگ α بدین معنی است که آخرین مشاهدات نقش مهم‌تری



تاریخ

شکل ۱. فراوانی تجمعی تعداد افراد مبتلای واقعی و پیش‌بینی شده

• دوره اول: پیش‌بینی‌ها از ۱ فوریه ۲۰۲۰ تا ۱۰ فوریه ۲۰۲۰

برای پیش‌بینی دوره اول، داده‌ها از تاریخ ۲۲ ژانویه ۲۰۲۰ (۲ بهمن ۱۳۹۸) به مدت ۱۰ روز یعنی تا ۳۱ ژانویه ۲۰۲۰ (۱۱ بهمن ۱۳۹۸) در نظر گرفته شده و پیش‌بینی برای ۱ فوریه ۲۰۲۰ (۱۲ بهمن ۱۳۹۸) تا ۱۰ فوریه ۲۰۲۰ (۲۱ بهمن ۱۳۹۸) انجام می‌شود. پیش‌بینی انجام شده با رنگ قرمز در شکل شماره ۱ نشان داده شده است (توجه کنید که محور عمودی لگاریتم مقیاس است). میانگین پیش‌بینی شده (پیش‌بینی نقطه‌ای) برای ۱۰ روز تعداد ۲۰۹ هزار مبتلا و فاصله اطمینان ۹۰ درصدی برای این میانگین از ۳۸ هزار تا ۵۰۴ هزار مبتلا برآورد شده است. این پیش‌بینی در مقایسه با مقدار واقعی مبتلایان در تاریخ ۱۰ فوریه ۲۰۰ (۲۱ بهمن ۱۳۹۸) که فقط ۴۳ هزار را نشان داد، بسیار فاصله دارد و خطای قابل ملاحظه‌ای در حدود ۱۶۶ هزار مبتلا را دارد. این پیش‌بینی بسیار نادقیق است. ولی نکته قابل ملاحظه این است که تعداد واقعی مبتلایان در فاصله اطمینان پیش‌بینی شده وجود دارد. طبیعی است که حجم کم داده در این پیش‌بینی و خطای زیاد مؤثر است. البته دلیل دیگر برای نادقیق بودن پیش‌بینی که بسیار مؤثر است، این است که از تاریخ ۳ فوریه ۲۰۲۰ تا تاریخ ۸ فوریه ۲۰۲۰ کشور چین دو بیمارستان را تأسیس کرد و همین موضوع در کاهش تعداد مبتلایان اثر بسیاری داشت (توجه داشته باشید که تا این زمان فقط کشور چین درگیر این بیماری به صورت رسمی شده بود).

• دوره دوم: پیش‌بینی‌ها از ۱۱ فوریه ۲۰۲۰ تا ۲۰ فوریه ۲۰۲۰

برای پیش‌بینی ۱۰ روز بعدی یعنی از تاریخ ۱۱ فوریه ۲۰۲۰ (۲۲ بهمن ۱۳۹۸) تا ۲۰ فوریه ۲۰۲۰ (۱ اسفند ۱۳۹۸)، طبیعتاً داده واقعی برای پیش‌بینی افزایش پیدا کرده و امید می‌رود که نتایج بهتر شود. بنابراین

$\hat{Y}_t$: مقدار برازش شده در زمان t .

(ب) شکل جمعی، در این حالت برای سری‌های زمانی با تغییرات فصلی ثابت به کار برده می‌شود و معادله‌ی آن به صورت ۶ در زیر است:

$$L_t = \alpha(Y_t - S_{t-p}) + (1 - \alpha)[L_{t-1} + T_{t-1}], 0 \leq \alpha \leq 1$$

$$T_t = \gamma[L_t - L_{t-1}] + (1 - \gamma)T_{t-1}, 0 \leq \gamma \leq 1$$

$$S_t = \delta(Y_t - L_t) + (1 - \delta)S_{t-p}, 0 \leq \delta \leq 1$$

$$\hat{Y}_t = L_{t-1} + T_{t-1} + S_{t-p} \quad (۶)$$

که در آن:

L_t : سطح سری در زمان t ,

α : پارامتر وزن مربوط به سطح،

T_t : روند در زمان t ,

γ : پارامتر وزن مربوط به روند،

S_t : جزء فصلی در زمان t ,

δ : پارامتر وزن مربوط به جزء فصلی،

Y_t : مقدار واقعی سری در زمان t ,

$\hat{Y}_t$: مقدار برازش شده در زمان t ,

p : تعداد فصل‌ها در یک سال (دوره فصلی) $[p = ۱۲]$ برای داده‌های ماهیانه و $p = ۴$ برای داده‌های سه ماهه (فصلی)]. همان‌طور که مشاهده شد، روش هموارسازی نمایی سه‌گانه در هر مرحله دارای یک جزء سطح، روند و فصلی است و بنابراین شامل سه وزن (پارامتر هموارسازی) برای به هنگام کردن این اجزاء در هر دوره می‌باشد. مقدارهای اولیه برای اجزاء سطح و روند همچون حالت روش هموارسازی نمایی دوگانه توسط یک رگرسیون خطی روی زمان و مقدارهای اولیه برای جزء فصلی نیز توسط رگرسیون متغیر موهومی با استفاده از داده‌های روندزایی شده به دست می‌آید.

۴ پیش‌بینی موارد قطعی کووید-۱۹

در این مقاله برای پیش‌بینی از هموارسازی نمایی سه‌گانه استفاده شده است. برای این منظور، اطلاعات هر ۱۰ روز را گرفته و پیش‌بینی برای ۱۰ روز آینده به صورت نقطه‌ای و فاصله‌ای استفاده شده است. به شکل شماره ۱ و توضیحات مربوط به آن توجه فرمایید.

شیب ثابت نزدیک است و این بدان معنی است که اگر کشورهایی مثل ایران، کره جنوبی و ایتالیا درگیر این ویروس نبودند، روند مبتلایان به بیماری به زودی متوقف می‌شد. اما متأسفانه درگیری کشورهایی مثل ایران، ایتالیا و کره جنوبی این روند را برهم زد. بنابراین از دوره بعدی احتمال افزایش خطا در پیش‌بینی نیز وجود دارد.

• **دوره چهارم: پیش‌بینی‌ها از ۲ مارس ۲۰۲۰ تا ۱۱ مارس ۲۰۲۰**
پیش‌بینی دوره چهارم از تاریخ ۲ مارس ۲۰۲۰ (۱۲ اسفند ۲۰۲۰) تا تاریخ ۱۱ مارس ۲۰۲۰ (۲۱ اسفند ۱۳۹۸) با رنگ بنفش در شکل شماره ۱ نشان داده شد است. میانگین برآورد شده تا تاریخ ۱۱ مارس ۲۰۲۰ به عدد ۱۱۲ هزار رسید و مقدار واقعی ۱۲۷ هزار مبتلا را نشان می‌دهد و این موضوع نشان‌دهنده خطای ۱۰۱۲ درصدی است. البته نکته مهم این است که همین میزان واقعی همچنان در فاصله اطمینان پیش‌بینی شده وجود دارد. این برای دومین بار است که مقدار پیش‌بینی از مقدار واقعی کمتر است. این موضوع بدین دلیل است که رشد نمایی در آمار مبتلایان به این بیماری در کشورهای مثل ایران، کشورهای اروپایی و آمریکا اتفاق افتاده است. نکته قابل توجه این است که آمار مبتلایان در کشور کره جنوبی به‌طور قابل ملاحظه‌ای در این دوره با اعمال محدودیت‌های شدید کاهش یافت. نکته عجیب در کشور کره جنوبی این است که مجموعه اقدامات انجام شده برای مدیریت بحران در این کشور، منجر به آن شده که کنترل آن نسبت به چین با سرعت بیش‌تری انجام شود، به‌طوری‌که از ۵ مارس (۱۵ اسفند ۱۳۹۸) یعنی حدود ۲ هفته پس از ورود کرونا به این کشور، روند تعداد مبتلایان نزولی شد.

• **دوره پنجم: پیش‌بینی‌ها از ۱۲ مارس ۲۰۲۰ تا ۲۱ مارس ۲۰۲۰**
پیش‌بینی آخرین دوره از تاریخ ۱۲ مارس ۲۰۲۰ (۲۲ اسفند ۱۳۹۸) تا ۲۱ مارس ۲۰۲۰ (۲ فروردین ۱۳۹۹) می‌باشد که در شکل شماره ۱ با رنگ زرد و طلایی نشان داده شده است. پیش‌بینی‌ها حاکی از ۸۳ هزار مورد جدید ابتلای قطعی است. در این دوره پیش‌بینی فاصله‌ای با احتمال‌های ۵۰ درصد، ۷۰ درصد و ۹۰ درصد نیز انجام پذیرفته است و رنگ‌های زرد و طلایی حاکی از فاصله پیش‌بینی شده بر اساس این درصد از اطمینان‌ها است. در این دوره پیش‌بینی موارد جدید مبتلا به بیماری فقط برای کشور چین با در نظر گرفتن داده‌های مربوط به این کشور، تک‌رقمی شد. اما پیش‌بینی موارد مبتلا برای کل داده‌ها (شامل کشورهای درگیر ویروس) روند صعودی در پیش گرفت.

تعداد داده واقعی ۲۰ عدد شده است که از تاریخ ۲۲ ژانویه ۲۰۲۰ تا ۱۰ فوریه ۲۰۲۰ است. با توجه به شکل شماره ۱، این پیش‌بینی با خط آبی نمایش داده شده است. همان‌طور که مشاهده می‌شود تعداد مبتلایان واقعی از تاریخ ۱۱ فوریه ۲۰۲۰ (۲۲ بهمن ۱۳۹۸) تا ۲۰ فوریه ۲۰۲۰ (۱ اسفند ۱۳۹۸)، به میانگین برآورد شده نزدیک است. خطای پیش‌بینی در این مدت فقط ۵ هزار و ۸۰۰ مورد (حدود ۷۰۷ درصد خطا) بوده است که در مقایسه با دوره اول از خطای بسیار کم‌تری برخوردار است. بنابراین مدل با افزایش حجم داده در طول زمان از کارایی بالاتری برخوردار می‌شود. در پیش‌بینی این دوره، یک نکته مهم این است که این پیش‌بینی دقیق‌تر با کاهش قابل توجهی در شیب در مقایسه با پیش‌بینی ده روز گذشته همراه بوده است. به عبارت دیگر بین ۲۰ تا ۳۰ روز کافی است تا روند مبتلایان به این بیماری روند نزولی در پیش بگیرد (همان‌طور که قبلاً هم ذکر شد، ۲۳ روز پس از اعلام رسمی اولین مورد ابتلا در چین، روند نزولی تعداد مبتلایان در این کشور با اعمال محدودیت‌های لازم و رعایت نکات بهداشتی شروع شد). نکته دیگر این است که تمام مقادیر واقعی در فاصله‌های پیش‌بینی قرار گرفته‌اند.

• **دوره سوم: پیش‌بینی‌ها از ۲۱ فوریه ۲۰۲۰ تا ۱ مارس ۲۰۲۰**
برای پیش‌بینی دوره سوم، تعداد داده‌های واقعی افزایش پیدا کرده و امیدواری به پیش‌بینی نیز افزایش پیدا کرده است. بنابراین ۳۰ داده واقعی در نظر گرفته شده و پیش‌بینی از تاریخ ۲۱ فوریه ۲۰۲۰ (۲ اسفند ۱۳۹۸) تا ۱ مارس ۲۰۲۰ (۱۱ اسفند ۱۳۹۸) انجام می‌شود. پیش‌بینی نقطه‌ای و فاصله‌ای در شکل شماره ۱ و با رنگ سبز نمایش داده شده است. میانگین پیش‌بینی شده برای ۱۰ روز پیش‌رو عدد ۸۳ هزار بوده است. همان‌طور که ملاحظه می‌شود شیب پیش‌بینی در مقایسه با شیب پیش‌بینی شده در دو دوره اول و دوم کمتر است و این بدان مفهوم است که تعداد مبتلایان جدید روند بسیار کندتری در جهان پیش گرفته است. همچنین یک کاهش قابل ملاحظه در عدم قطعیت نیز مشاهده می‌شود. زیرا فاصله اطمینان پیش‌بینی شده بهتر انجام شده و به میانگین پیش‌بینی شده نزدیک‌تر است. تعداد مبتلایان قطعی واقعی تا تاریخ ۱ مارس ۲۰۲۰ عدد ۸۸ هزار بوده است که با مقدار پیش‌بینی شده فاصله نزدیکی دارد و میزان خطا فقط ۲۰۶ درصد (به عبارتی ۵ هزار ۵۰۰ نفر) بوده است. نکته قابل توجه این است که دوره سوم، اولین دوره‌ای است که مقدار پیش‌بینی شده از مقدار واقعی کمتر است. این موضوع به این دلیل است که این ویروس در این زمان در کشورهایی مثل کره جنوبی، ایران و ایتالیا شیوع پیدا کرده و مقدار واقعی بیش از مقدار برآورد شده است. این دوره بسیار دوره مهمی است، زیرا همان‌طور که در شکل شماره ۱ نمایش داده شده روند این بیماری با پیش‌بینی انجام شده به

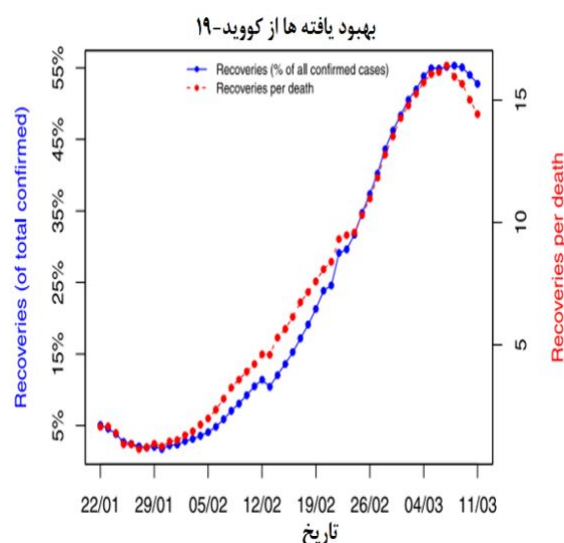
۵ تحلیل داده‌های مربوط به بیماران ۶ بحث و نتیجه‌گیری

بهبودیافته

همان‌طور که ملاحظه می‌شود تحلیل داده‌ها برای شکل شماره ۱ برای همه کشورهای دنیا از جمله چین انجام شده است که توانست ظرف مدت ۵۰ روز به کنترل این ویروس دست یابد. بنابراین دوره‌های ۱۰ روزه از آغاز شروع این ویروس در چین در نظر گرفته شده و تحلیل انجام شده است. بنابراین الگوی ارائه شده در شکل شماره ۱، الگوی بسیار موفق برای همه کشورهاست که سازمان بهداشت جهانی نیز برای پیش‌بینی و اظهار نظر در مورد این ویروس به این الگو متوسل شده است. چون در الگوی ارائه شده در شکل شماره ۱، آمار همه کشورها دخیل است، بنابراین الگوی بسیار مناسبی برای کشور عزیزمان ایران است تا روند پیش‌بینی طبق آن صورت پذیرد. نکته بسیار مهم این است که این الگو برای پیش‌بینی در هر بازه زمانی مناسب است. بنابراین در کشور ایران، طبق آمار رسمی وزارت بهداشت می‌توان تحلیل‌های زیر را ارائه کرد. طبق آمار رسمی وزارت بهداشت در کشور ایران، تا تاریخ ۱۵ فروردین ۱۳۹۹ آمار مبتلایان قطعی عدد ۵۳۱۸۳ نفر، بهبودیافتگان ۱۷۹۳۵ نفر و فوتی‌ها تعداد ۳۲۹۴ نفر می‌باشد که نسبت بهبودیافتگان به کل موارد قطعی حدود ۳۳ درصد و نسبت بهبودیافتگان به ازای هر فوتی حدود ۵ نفر است که با نسبت‌های بیان شده در مهار ویروس کرونا فاصله داشت. با توجه به محدودیت بیشتر دولت در این بازه زمانی روند مبتلایان نزولی بود و همچنین:

- به‌طور متوسط ۲۵ روز (میانگین دوره دوم پیش‌بینی) لازم بود تا روند مبتلایان جدید به این ویروس، نزولی شود. برای مثال، اگر شیوع این بیماری را به‌طور جدی از تاریخ ۲۰ اسفند ۹۸ در ایران نظر بگیریم، روند نزولی تعداد مبتلایان جدید از تاریخ حدود ۱۵ فروردین با اعمال سیاست‌های لازم شروع شد که در واقعیت هم همین اتفاق افتاد، زیرا دولت محدودیت‌های زیادی برای مردم تعیین کرد.
- به‌طور متوسط ۳۵ روز (میانگین دوره سوم) طول می‌کشد تا روند ابتلای جدید به این بیماری از شیب ثابت نسبی برخوردار باشد. البته همان‌طور که از شکل شماره ۱ پیدا است، در دوره چهارم نیز شیب ثابت نسبی وجود دارد، بنابراین می‌توان رسیدن به یک شیب ثابت نسبی را به عبارتی ۴۵ روز نیز در نظر گرفت. بازهم به‌عنوان یک مثال، اگر شیوع این بیماری را از تاریخ ۲۰ اسفند در نظر بگیریم، تا تاریخ حدود ۵ اردیبهشت ۱۳۹۹ بسته به سیاست‌های اعمال شده در راستای کنترل رفت‌وآمد و رعایت فاصله‌گذاری اجتماعی، تعداد مبتلایان جدید از شیب ثابت برخوردار می‌شود. در واقعیت هم همین مسئله رخ داد و تا اوایل اردیبهشت روند ابتلایان به این بیماری در ایران نزولی شد.

حال توجه خود را به موارد بهبودیافته معطوف می‌کنیم. برای این منظور درصد موارد بهبودیافته از کل موارد قطعی مبتلا به این بیماری و همچنین نسبت موارد بهبودیافته به ازای هر فوتی مدنظر قرار می‌گیرد. به شکل شماره ۲ دقت کنید. همان‌طور که در شکل شماره ۲ مشاهده می‌شود، یک رابطه قوی بین دو نمودار مشاهده می‌شود و نشان از همبستگی بالای نسبت بهبودیافته به کل موارد قطعی و نسبت بهبودیافته به ازای هر فوتی است. اگرچه تعداد بهبودیافتگان تا آخر ژانویه بسیار کم بوده است (کم‌تر از ۲ درصد)، اما بعدازآن این روند صعودی شده و به عدد ۸۰۵۲ درصد نیز رسیده است. همچنین نسبت بهبودیافته‌ها به ازای هر فوتی نیز به عدد ۱۵ رسیده است. همان‌طور که مشاهده می‌شود، در تاریخ ۸ مارس ۲۰۲۰ شاهد روند نزولی تعداد بهبودیافتگان هستیم و این به دلیل افزایش موارد قطعی و فوتی سایر کشورها به‌جز چین است. با توجه به این‌که در تاریخ ۱۰ مارس ۲۰۲۰ کشور چین توانست ویروس کرونا را مهار کند، این موضوع در شرایطی اتفاق افتاد که نسبت بهبودیافتگان به کل مبتلایان قطعی در جهان به بالای ۵۰ درصد رسید و نسبت بهبودیافتگان به ازای هر فوتی نیز به بیش از ۱۵ رسید. باید توجه داشت که چون بیش‌تر آمار (به‌خصوص بهبودیافتگان) تا این تاریخ مربوط به کشور چین است، بنابراین با الگوبرداری از کشور چین، اگر در کشوری این نسبت رخ دهد، امیدوار به مهار کرونا خواهد بود.



شکل ۲. بهبودیافته‌ها به‌عنوان درصدی از کل موارد قطعی (سمت چپ) و بهبودیافته‌ها به ازای هر فوتی (سمت راست)

ایام تحت فشار بیش‌تری هستند، می‌تواند مشکلات را تا حدی مرتفع سازد. البته این امر مستلزم این است که فرایند اداری این کمک‌های دولتی تسهیل شود و مخصوصاً رسته‌های شغلی برای دریافت حمایت دولت، زمان زیادی صرف نکنند و واقع دولت نقش حمایت‌کننده واقعی را داشته باشد. نکته قابل ذکر در خصوص اقدامات دولت این است که جمهوری اسلامی ایران با وجود تحریم و همه فشارهای اقتصادی سعی کرده است همه خدمات مربوط به بیماران را رایگان انجام بدهد. از پیشنهادها اثرگذار به دولت می‌توان موارد زیر را نیز ذکر کرد:

- اول اجماع و یک‌صدایی خود دولت در رابطه با مسائل مربوط به کرونا مهم و تأثیرگذار است.

- سلامت جامعه اولویت بر دیگر مسائل و حتی مسائل اقتصادی داشته باشد.

- هزینه مبارزه با بهداشت جامعه را از قبیل لباس، دستکش و ماسک و ... را دولت تأمین کرده و با حداقل قیمت به دست مصرف‌کننده برساند.

- از زمان شروع تا پایان زمان بحران ناشی از ویروس، بخشودگی و یا حداقل زمان لازم تنفس برای هزینه‌های بیمه، مالیات و بازپرداخت تسهیلات و امهال وام‌ها را بپذیرند و بانک‌ها نرخ سود تسهیلات را حتی المقدور برای بدهکاران شاغل در فضای کسب‌وکار کرونایی را در دوران کرونا کاهش دهند.

- هزینه بیمه بیکاری را در طی دوره بحران تأمین و بپردازند.

- از صدور بخشنامه‌های خلق الساعه که محل کسب‌وکار است، بپرهیزند و فضای کسب‌وکار را سخت‌تر از شرایط موجود نکنند.

- در سال جدید که سال جهش تولید نام‌گذاری شده است، ضمن کاهش تعرفه‌های موردنیاز تولید در حوزه واردات، خصوصاً برای تولید صادراتی، هزینه‌های تولید و قیمت تمام‌شده را کاهش دهند.

- شرایط را برای بازگرداندن بازارهای ازدست‌رفته صادرات به علت بسته شدن مرزها فراهم نمایند و سیاست منطقی و درست در پیش گیرد.

- از نظرات اساتید محترم دانشگاه که متخصص در امور اقتصادی هستند بهره بیش‌تری گرفته شود و از دانشگاه خواسته شود تا راهکار خود را ارائه دهد.

لازم به ذکر است که در این گزارش فقط از متغیرهای تعداد مبتلایان قطعی، بهبودیافته و فوتی استفاده شده و در اصطلاح از تحلیل یک متغیری استفاده شده است. در صورتی که اگر متغیرهای پیشگوی دیگری مانند تعداد بیمارستان‌ها، تخت‌های بیمارستانی و ... و حتی در نظر گرفتن تحریم به‌عنوان یک متغیر مهم در نظر گرفته شود می‌توان این

- به‌طور متوسط بعد از حدود ۵۰ روز از شیوع این بیماری، نسبت بهبودیافتگان به کل موارد قطعی به بیش از ۵۰ درصد افزایش می‌یابد و یا بهبودیافتگان به ازای هر فوتی نیز به بیش از ۱۵ نفر می‌رسد که نشان از مهار این ویروس می‌باشد. البته متأسفانه در ایران این اتفاق رخ نداد و دولت محدودیت‌های لازم را بعد از ۵۰ روز برداشت و مجدد از حدود ۱۰ اردیبهشت با برداشتن محدودیت‌ها که سیاست غلطی بود، تعداد مبتلایان افزایش یافت.

- نکته مهم این است که به‌طور متوسط ۵۰ روز برای مهار ویروس کرونا زمانی است که دولت و مردم در راستای فاصله‌گذاری اجتماعی و رعایت نکات بهداشتی باهم هم‌قدم شوند. در صورتی که متأسفانه در اردیبهشت‌ماه این اتفاق برای ایران رخ نداد و هم دولت محدودیت‌ها را به‌سرعت برداشت و هم مردم رعایت نکات بهداشتی را از دستور کار خارج کردند و همین موضوع کافی بود تا مجدد به وضعیت دوره اول برگردیم. به عبارت دیگر اواسط اردیبهشت‌ماه ۱۳۹۹ در واقع دوره اول شکل شماره ۱ در نظر گرفته شد. در هر زمان که محدودیت‌ها کم‌تر شود و یا فاصله‌گذاری اجتماعی رعایت نشود، ممکن است مجدد به دوره اول با شیب زیاد برگردیم و روند مهار این ویروس نیز بسیار زمان‌بر باشد. عملاً آنچه در ایران تا مردادماه ۱۳۹۹ رخ داد همین وضعیت است و متأسفانه استفاده از ماسک به‌صورت اجباری بسیار دیر در ایران عملی شد و همین باعث شد عملاً به دوره اول شکل شماره ۱ برگردیم و منتظر ۵۰ روز بعدی برای مهار این ویروس باشیم. دولت در چند ماه اخیر تلاش گسترده‌ای برای مقابله با کرونا اجرا کرده است که البته در این امر کادر درمان پرستاران و پزشکان نقش اصلی را داشته‌اند. از سوی دیگر تولیدکنندگان هم نقش بی‌بدیلی در تأمین تجهیزات و اقلام موردنیاز مقابله با کرونا انجام داده‌اند. دستگاه دیپلماسی هم با تبیین شرایط ایران هم سعی کرد منابع لازم و تجهیزات موردنیاز را از کشورهای خارجی تأمین کند و هم شرایط ایران در مواجهه با این بیماری با عنایت به تحریم‌های ظالمانه آمریکا برای افکار عمومی جهانی تبیین کند. اما متأسفانه کاهش محدودیت‌های لازم از اواسط اردیبهشت در ایران توسط دولت، باعث اوج‌گیری مجدد این ویروس شد و تعداد بیماران و فوتی‌ها افزایش پیدا کرد. هرچند مدتی بعد دولت استفاده از ماسک را اجباری کرد، اما می‌توانست این اقدامات از اردیبهشت‌ماه با همان قدرت قبلی صورت پذیرد. بنابراین اعمال محدودیت‌های لازم از طرف دولت و تشویق مردم به رعایت گذاری فاصله‌گذاری اجتماعی می‌تواند در مهار این ویروس بسیار مؤثر باشد. در این راستا ارائه وام و تسهیلات درکنار حمایت‌های معیشتی از مردم، بخشی از بسته مقابله با بحران کرونایی دولت می‌تواند باشد. همچنین حمایت مالی از رسته‌های شغلی که در این

پیش‌بینی را دقیق‌تر انجام داد. برای متغیری مانند تحریم می‌توان از طیف لیکرت بین ۱ تا ۵ استفاده کرد و روزانه تحریم را رتبه‌بندی را کرد. برای مثال، زمانی که اشخاص و سازمان‌های بیش‌تری تحریم می‌شوند عدد ۵ و به همین ترتیب روزها را امتیازدهی کرد.

مراجع

- [1] Li, X. (2020). Homologous recombination within the spike glycoprotein of the newly identified coronavirus may boost cross_species transmission from snake to human. *Journal of Medical Virology*, online.
- [2] Benvenuto, D., Giovannetti, M., Ciccozzi, A., Spoto, S., Angeletti, S., and Ciccozzi, M. (2020). the 2019-new coronavirus epidemic: evidence for virus evolution. *Journal of Medical virology*, **92(4)**, 455-459.
- [3] Chan, J.F., Kok, K.H., Zhu, Z., Chu, H., To, K.K., Yuan, S., and Yuen, K.Y. (2020). Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. *Emerging Microbes and Infections*, **9(1)**, 221-236.
- [4] Zhao, Y., Zhao, Z., Wang, Y., Zhou, Y., Ma, Y., and Zuo, W. (2020). Single-cell RNA expression profiling of ACE2, the Receptor of SARS-CoV-2. *Am J Respir Crit Care Med.*, **202(5)**, 756-759.
- [5] Hyndman, R.J., Koehler, A.B., Snyder, R.D., and Grose, S. (2002). A state space framework for automatic forecasting using exponential smoothing methods. *International Journal of Forecasting*, **18(3)**, 439-454.

Forecasting the novel coronavirus COVID-19

Vahid Rezaei Tabar¹

Abstract:

At the end of December 2019, the spread of a new infectious disease was reported in Wuhan, China, caused by a new coronavirus and officially named Covid-19 by the World Health Organization. As the number of victims of the virus exceeded 1,000, the World Health Organization chose the official name Covid-19 for the disease, which refers to "corona", "virus", "disease" and the year 2019. The forecasting about Covid-19, can help the government make better decisions. In this paper, an objective approach is used for forecasting Covid-19 based on the statistical methods. The most important goal in this paper is to forecast the prevalence of coronavirus for confirmed, dead and improved cases and to estimate the duration of the management of this virus using the exponential smoothing method. The exponential smoothing family model is used for short time-series data. This model is a kind of moving average model that modifies itself. In other words, the exponential smoothing is one of the most widely used statistical methods for time series forecasting, and the idea is that recent observations will usually provide the best guidance for the future. Finally, according to the exponential smoothing, we will provide some suggestions.

Keywords: Covid-19, Forecasting, Objective approach, Exponential Smoothing.

¹Assistant Professor, Department of Statistics, Allameh Tabataba'i University, Tehran, Iran