

تحلیل فضایی بیزی زمان بقای بیماران مبتلا به کووید-۱۹

کیومرث مترجم

گروه آمار دانشگاه تربیت مدرس

چکیده: شیوع بیماری کووید ۱۹ از آغاز سال ۲۰۲۰ تأثیرات زیادی بر زندگی مردم در تمام دنیا داشته است. با توجه به مشکلات مربوط به ظرفیت سیستم‌های درمانی در کشورهای مختلف، بسیاری از تحقیقات انجام شده از زمان آغاز دنیاگیری کرونا بر مدل‌بندی زمان بقای بیماران بستری شده در بخش مراقبت‌های ویژه بیمارستان‌ها و مراکز درمانی متمرکز شده است. به نظر می‌رسد موقعیت مکانی افراد مبتلا به بیماری بر زمان بقا و در نتیجه مدت زمان بستری آنها موثر باشد. در این مقاله با در نظر گرفتن مدل‌های معمول و مدل بقای فضایی با رویکرد بیزی به تحلیل داده‌های مربوط به زمان بقای بیماران مبتلا به کووید ۱۹ در شهر تهران پرداخته شده است و پس از مقایسه عملکرد مدل‌های بقا، اثر متغیرهای تبیینی اندازه‌گیری شده بر میزان خطر فوت بیماران مورد بررسی قرار گرفته است.

واژه‌های کلیدی: داده‌های فضایی، مدل بقای فضایی، رویکرد بیزی، کووید ۱۹
کد موضوع بندی ریاضی (۲۰۱۰): 62F15, 62N86, 62H10.

۱ مقدمه

ویروس‌های کرونا یک خانواده بزرگ از ویروس‌ها هستند که می‌توانند باعث ایجاد یک سرماخوردگی ساده تا بیماری‌های حاد تنفسی مثل سارس شوند. ویروس کرونای جدید هم از این خانواده است که اولین بار در اواخر سال ۲۰۱۹ شهر ووهان چین مشاهده شد. در دسامبر ۲۰۱۹ مسؤلان سلامت عمومی چین به سازمان جهانی بهداشت اعلام کردند که ویروس جدید و ناشناخته‌ای باعث نوعی بیماری با علائمی شبیه به ذات‌الریه در شهر ووهان شده است. آن‌ها تشخیص دادند که این ویروس از خانواده‌ی ویروس کرونا است و به سرعت در حال تکثیر و راه‌یابی به بیرون از شهر ووهان است. میزان شیوع ویروس کرونا در انسان‌های مختلف متفاوت است و برخی از افراد مبتلا، این ویروس را به دیگران منتقل نمی‌کنند و عده‌ای هم بدون

* نام و آدرس پست الکترونیکی ارائه‌دهنده مقاله: کیومرث مترجم، k.motarjem@modares.ac.ir

داشتن علائم بیماری ناقل محسوب می‌شوند در انتقال ویروس از انسان به انسان نتایج مطالعات انجام شده توسط محققان نشان می‌دهد که هر فرد به طور متوسط قادر است بیماری را به $1/4$ تا $3/8$ نفر دیگر منتقل کند. (شنیرینگ، ۲۰۲۰) این اعداد که آنرا شاخص تکثیر پایه می‌گویند نشان می‌دهد چه شماری از افرادی که تازه آلوده شده‌اند ویروس را در جمعیت منتقل می‌کنند. نکته قابل توجه در انتشار ویروس کرونا از افراد بیمار به افراد سالم همواره فاصله و موقعیت قرارگیری افراد از یکدیگر بوده است. بنابراین همبستگی فضایی یکی از موضوعاتی است که لازم است در تحلیل بقا و مدت زمان بستری^۱ بیماران مبتلا به کووید ۱۹ مورد توجه قرار گیرد. در این مقاله و در بخش ۲ بطور مختصر به تاریخچه استفاده از مدل‌های بقا پرداخته شده است. در بخش ۳ مدل‌بندی بیزی داده‌های بقا معرفی می‌شود. در بخش ۴ نحوه کاربست مدل‌های بقای کلاسیک و فضایی با رویکرد بیزی برای تحلیل داده‌های زمان بقای بیماران مبتلا به کووید ۱۹ در شهر تهران ارائه می‌شود و در نهایت به بحث و نتیجه‌گیری در خصوص نتایج پرداخته می‌شود.

۲ مدل‌های بقا

مدل‌های کلاسیک بقا از اواسط قرن هفدهم میلادی در تحلیل داده‌های پزشکی مورد استفاده محققان بوده است (گلس و همکاران، ۱۹۶۳). در مدل‌های کلاسیک بقا فرض بر آن است که زمان رخداد حادثه برای افراد مختلف تحت بررسی مستقل از هم هستند اما در پاره‌ای از موارد ممکن فرض استقلال برقرار نباشد و زمان‌های رخداد بر حسب موقعیت مکانی دارای همبستگی فضایی باشند محققان زیادی مانند تیوریچک و مادن (۲۰۰۲) و رامزی و همکاران (۲۰۰۳) در مطالعات خود نشان دادند که در شرایط همبسته بوده داده‌های بقا استفاده از مدل‌های کلاسیک نتایج مطلوبی نخواهد داشت. معمولاً اثرات تصادفی به صورت یک مؤلفه پنهان در داده‌های بقا نهفته هستند که می‌توان با تشخیص وجود همبستگی فضایی و وارد کردن آن در مدل از طریق یک مدل بقای فضایی به نتایجی دقیق‌تر دست یافت. برای تحلیل اینگونه داده‌های بقا مدل‌های زیادی معرفی شده است که عمدتاً با رویکرد فراوانی‌گر به برآورد پارامترهای تابع بقا و تابع خطر می‌پردازند اما استفاده از این رویکرد بواسطه پیچیدگی زیاد توابع درستمایی موجب زمانبر شدن برآورد پارامترها می‌شود و از مسائل چالشی در تحلیل بقا است. به همین سبب محققانی چون بانرجی و همکاران (۲۰۰۴) و دارموفال (۲۰۰۹) در تحلیل داده‌های بقای فضایی از رویکرد بیزی استفاده کرده‌اند که در اینصورت حجم محاسبات به طور قابل توجهی کاهش پیدا می‌کند اما مدل‌های معرفی شده در شرایطی که داده‌های بقا بصورت شبکه‌ای بوده و همبستگی فضایی بین نواحی تعریف شده باشد کاربرد دارند. در این مقاله به تحلیل بیزی مدل‌های بقای فضایی زمین آماری پرداخته شده است. در صورت وجود همبستگی فضایی در داده‌های بقا مدل‌های معمول در تحلیل بقا نمی‌تواند تحلیل دقیقی از زمان بقا و تابع خطر ارائه دهند زیرا همبستگی فضایی سبب ایجاد عوامل خطر ناشناخته می‌شود. برای غلبه بر این مشکل در تحلیل داده‌های بقا، رده‌ای از مدل‌های شکنندگی ارائه شده است که به آنها مدل‌های بقای فضایی گفته می‌شود. در این مدل‌ها به طور معمول یک میدان تصادفی برای در نظر گرفتن همبستگی فضایی به مدل شکنندگی اضافه می‌شود. بنابراین تابع خطر فضایی (مترجم و همکاران، ۲۰۲۰) به صورت

$$h(t|X, Z) = h_0(t) \exp(\beta'X + Z(s)), \quad (1.2)$$

تعریف می‌شود که در آن $h_0(\cdot)$ تابع خطر پایه، β بردار پارامترهای رگرسیونی، $Z(\cdot)$ یک میدان تصادفی گاوسی و s موقعیت قرارگیری داده‌های زمان بقا را در $D \subseteq R^2$ نشان می‌دهد. استفاده از این مدل در صورت گاوسی بودن میدان تصادفی نتایج بسیار مطلوبی نسبت به مدل‌های کلاسیک در تحلیل بقا دارد.

¹Hospitalization

۳ تحلیل بیزی داده‌های بقا

در رهیافت فراوانی‌گرا عموماً برای برآورد پارامترهای مدل‌های بقا از روش ماکسیمم درست‌نمایی استفاده می‌شود که محاسبه آنها معمولاً با الگوریتم EM صورت می‌پذیرد. این روش برآورد پارامترها مشکلاتی را به همراه دارد که از جمله آنها می‌توان به مشکلات محاسباتی و امکان ارائه یک ماکسیمم موضعی به جای ماکسیمم کلی اشاره کرد. حتی در پاره‌ای از شرایط ممکن است همگرایی با این روش ایجاد نگردد. با توجه به خواص مجانبی برآوردهای ماکسیمم درست‌نمایی در صورت بزرگ بودن حجم نمونه برآوردهای حاصل تا حدود زیادی قابل اعتماد هستند اما برآورد پارامترها با این روش هنگام کوچک بودن حجم نمونه چالش استنباطی بزرگی را به دنبال خواهد داشت. برای اجتناب از برخورد با چنین مشکلاتی از رهیافت بیزی در برآورد پارامترها استفاده می‌شود.

۱.۳ مدل بیزی کاکس

یکی از اهداف اصلی در مدل‌بندی کردن داده‌های بقا، برآورد تابع خطر است. تابع خطر که با $h(t)$ نشان داده می‌شود پتانسیل لحظه‌ای وقوع شکست به شرط عدم رخداد شکست تا زمان t را برای هر مشاهده نشان می‌دهد و به صورت

$$h(t) = \lim_{\epsilon \rightarrow 0^+} \frac{P(t < T^* \leq t + \epsilon | T^* > t)}{\epsilon} = \frac{f(t)}{S(t)} \quad (1.3)$$

تعریف می‌شود. یکی از متداول‌ترین توابع خطر توسط کاکس (۱۹۷۲) تعریف شد. در این مدل با تغییر توزیع تابع خطر پایه، توابع خطر متنوعی قابل تعریف است. از آن جمله، تابع خطر پایه وایبول است که به واسطه داشتن دو پارامتر شکل و مقیاس از انعطاف پذیری زیادی در برازش به داده‌های بقا برخوردار است و به همین علت یکی از پرکاربردترین توابع خطر در تحلیل بقا است. مدل کاکس با تابع خطر پایه وایبول به صورت

$$h(t|X) = \nu t^{\nu-1} e^{\beta' X}$$

است، که در آن ν پارامتر شکل تابع خطر پایه و β بردار پارامترهای رگرسیونی است. در تحلیل بیزی با در نظر گرفتن توزیع‌های پیشین $p(\beta)$ و $p(\nu)$ برای پارامترها تابع چگالی پسینی به صورت

$$p(\beta, \nu; t, x, \delta) \propto L(\beta, \nu; t, x, \delta) p(\beta) p(\nu)$$

است و تابع درست‌نمایی آن نیز به صورت

$$\prod_{i=1}^n \{ \nu t_i^{\nu-1} \exp(\beta' x_i) \}^{\delta_i} \exp\{-t_i^{\nu} \exp(\beta' x_i)\}$$

حاصل می‌شود، که در آن δ تابع نشانگر سانسور است.

مقدار پارامتر شکل در تابع خطر پایه وایبول بر رفتار تابع خطر مؤثر است به قسمی که به ازای $\nu > 1$ تابع خطر افزایشی، $\nu < 1$ تابع خطر کاهش‌ی و به ازای $\nu = 1$ تابع خطر ثابت است.

۲.۳ مدل بیزی شکنندگی

در برخی مطالعات بقا بسیاری از عوامل تأثیرگذار بر زمان بقای واحدهای آزمایشی مورد مطالعه قابل مشاهده و یا اندازه‌گیری نیستند. در این حالات معمولاً از مدل‌های شکنندگی استفاده می‌شود. مدل شکنندگی تعمیمی از مدل کاکس با هدف ورود

یک متغیر به عنوان نماینده عوامل خطر ناشناخته است. در این مدل فرض بر این است که اثرات شکنندگی واحدهای مختلف مستقل از هم باشند. تابع خطر مدل شکنندگی به صورت

$$h(t|W, X) = Wh_*(t)\exp(\beta'X) \quad (2.3)$$

است، که در آن t زمان بقا، $h_*(\cdot)$ تابع خطر پایه، X بردار متغیرهای تبیینی، β بردار اثر متغیرهای تبیینی و W اثر شکنندگی است.

در مدل شکنندگی فرض می‌شود که مولفه شکنندگی مستقل از متغیرهای تبیینی است. در این مدل تغییرات کل تابع خطر و در نتیجه تغییرات تابع بقا به اثر متغیرهای معلوم و اثر عوامل ناشناخته از طریق مولفه تصادفی مدل تجزیه می‌شود. مدل شکنندگی با تخصیص اثر تصادفی، وجود تفاوت در آزمودنی‌ها را بر روی برآورد اثر متغیرهای درون مدل لحاظ می‌کند. با توجه به مثبت بودن تابع خطر، لازم است برای متغیر شکنندگی توزیعی با تکیه‌گاه مثبت در نظر گرفته شود. انتخاب توزیع مناسب برای متغیر شکنندگی از اهمیت زیادی برخوردار است. طیف گسترده‌ای از توزیع‌های آماری تکیه‌گاه مثبت در انتخاب مولفه شکنندگی پیش روی محققان قرار دارد. در استنباط بیزی، مدل شکنندگی با تابع خطر پایه وایبول، دارای چگالی پسینی به صورت

$$p(\beta, w, \nu; t, x, \delta) \propto L(\beta, w, \nu; t, x, \delta)p(w)p(\beta)p(\nu)$$

است و تابع درستنمایی آن نیز به صورت

$$\prod_{i=1}^n \{\nu t_i^{\nu-1} \exp(\beta'x_i + w_i)\}^{\delta_i} \exp\{-t_i^{\nu} \exp(\beta'x_i + w_i)\}$$

به دست می‌آید.

۳.۳ مدل بیزی بقای فضایی

برای مدل‌بندی داده‌های بقای فضایی از تابع خطر به صورت

$$h(s, t|X, Z) = h_*(t)\exp(\beta'X + Z(s)) \quad (3.3)$$

استفاده می‌شود، که در آن t زمان بقا، $h_*(\cdot)$ تابع خطر پایه، X بردار متغیرهای تبیینی، β بردار متغیرهای تبیینی، s موقعیت جغرافیایی و $Z(\cdot)$ میدان تصادفی گاوسی است.

با رهیافت بیزی، مدل بقای فضایی دارای چگالی پسینی به صورت

$$p(\beta, Z, \nu, \Theta; t, x, \delta) \propto L(\beta, Z, \nu; t, x, \delta)p(Z; \Theta)p(\beta)p(\nu)p(\Theta)$$

است و تابع درستنمایی با فرض تابع خطر پایه وایبول به صورت

$$L(\beta, Z, \nu, \Theta; t, x, \delta) \propto \prod_{i=1}^n \{\nu t_i^{\nu-1} \exp(\beta'x_i + Z_i)\}^{\delta_i} \exp\{-t_i^{\nu} \exp(\beta'x_i + Z_i)\}$$

حاصل می‌شود، که در آن Θ بردار پارامترهای میدان تصادفی و $Z_i = Z(s_i)$ است. در بخش بعد با استفاده از مدل‌های کاکس، شکنندگی و بقای فضایی به تحلیل داده‌های مربوط به زمان بقای بیماران مبتلا به کووید ۱۹ که در بخش مراقبت‌های ویژه مراکز درمانی شهر تهران بستری شده‌اند پرداخته می‌شود.

۴ مطالعه کاربردی: تحلیل زمان بقای بیماران مبتلا به کووید ۱۹

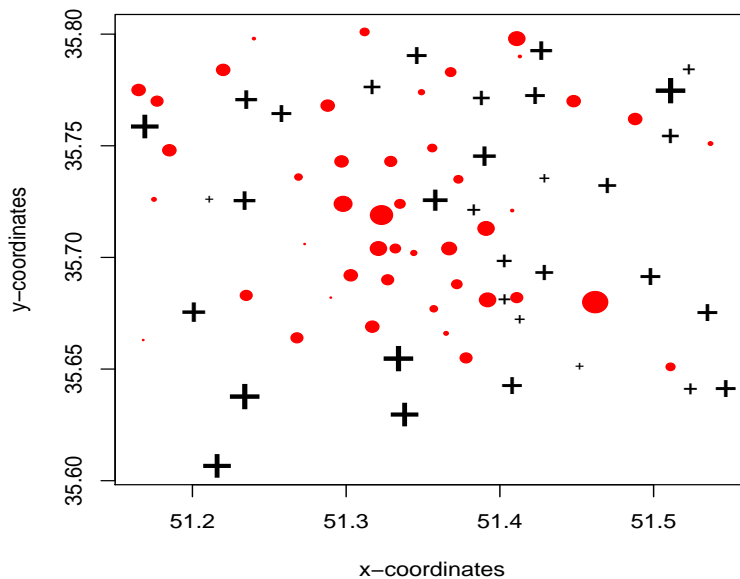
در این بخش با استفاده از رویکرد بیزی به تحلیل داده‌های مربوط به زمان بقای بیماران مبتلا به کووید ۱۹ در شهر تهران که در بخش مراقبت‌های ویژه بستری شده‌اند پرداخته می‌شود. نمونه شامل ۷۸ نفر از مبتلایان ساکن در مناطق مختلف شهر تهران است و اطلاعات مربوط به آنها توسط نویسنده با استفاده از پرسشنامه الکترونیکی و در بستر فضای مجازی در آذر ماه ۱۳۹۹ جمع آوری شده است. بر اساس اطلاعات جمع آوری شده در این بررسی، حداکثر زمان بستری بیماران در بخش مراقبت‌های ویژه ۱۸ روز بوده است. نمونه شامل ۱۴ مرد و ۳۷ زن بوده است که ۹۴ نفر از آنها دارای حداقل یک بیماری زمینه‌ای بوده‌اند. موقعیت مکانی بیماران در شکل ۱ نمایش داده شده است که در آن افرادی که در طول دوره انجام مطالعه فوت شده‌اند با دایره و کسانی که زنده مانده‌اند یا به عبارتی سانسور راست شده‌اند با علامت + نشان داده شده‌اند. مساحت دایره‌های قرمز رنگ نشان دهنده مدت زمان بستری تا فوت است به طوری که دایره‌های با مساحت کمتر نشان‌دهنده فوت بیمار در روزهای اول بستری است. لازم به ذکر است که در این مطالعه سانسور راست ۴۴ درصدی در داده‌ها وجود دارد. متغیرهایی تبیینی شامل جنسیت، سن، وضعیت تاهل و وجود یا عدم وجود بیماری زمینه‌ای در افراد متغیرهایی بوده‌اند که

جدول ۱: برآورد پارامتر مدل‌های برازنده شده به داده‌های بیماران مبتلا به کووید ۱۹

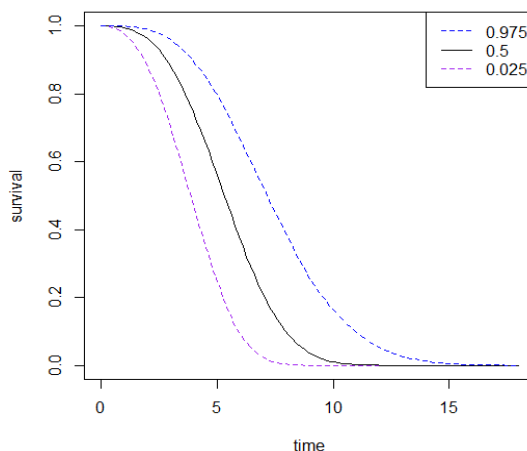
پارامتر	کاکس		شکندگی		بقای فضایی	
	برآورد	انحراف استاندارد	برآورد	انحراف استاندارد	برآورد	انحراف استاندارد
β_1	۰/۷۱۵	۰/۳۵۲	۰/۸۱۰	۰/۲۱۸	۰/۸۲۴	۰/۱۱۲
β_2	۰/۱۰۵	۰/۰۲۴	۰/۱۷۸	۰/۰۲۱	۰/۱۰۴	۰/۰۱۲
β_3	۱/۳۲۸	۰/۴۳۶	۱/۴۱۱	۰/۴۲۱	۱/۴۷۳	۰/۱۷۲
a	-	-	-	-	۰/۱۱۲	۰/۰۴۴
σ^2	-	-	-	-	۲/۴۰۳	۰/۱۸۵
<i>DIC</i>	۱۲۸۶/۳۱۲		۱۱۹۵/۲۹۸		۱۰۳۵/۱۱۸	

در این مطالعه مورد توجه قرار گرفته‌اند. در یک تحلیل اولیه از داده‌ها متغیرهای تبیینی جنسیت، سن و بیماری زمینه‌ای در مدت زمان بقای بیماران مبتلا به کووید ۱۹ معنی‌دار تشخیص داده شده‌اند که ضرایب این متغیرهای تبیینی در مدل‌های ارائه شده به ترتیب با β_1 ، β_2 و β_3 نمایش داده شده است. در مرحله بعد برای بررسی تأثیر هر یک از متغیرهای تبیینی و موقعیت مکانی بیماران بر زمان بقا، و برازش مدل‌های کاکس، شکندگی و بقای فضایی با توجه به مزدوج بودن توزیع‌های نرمال و گاما و همچنین ماهیت ضرایب رگرسیونی و واریانس‌های اثرهای تصادفی به ترتیب از این دو توزیع به‌عنوان توزیع‌های پیشین استفاده شده است. در برازش مدل‌ها با روش بیزی طول زنجیر ۱۰۰۰۰۰۰، زمان داغیدن ۲۰۰۰۰۰ و طول گام ۲۰ در نظر گرفته شده است.

همان‌طور که در جدول ۱ ملاحظه می‌شود، به علت وجود همبستگی فضایی بین داده‌ها مدل کاکس عملکرد مطلوبی نداشته و وجود اثر فضایی برآورد پارامترهای رگرسیونی را تحت تأثیر قرار داده است. لذا در صورت استفاده از این مدل، نتایج دقیق نخواهد بود. اما در مدل شکندگی به علت وجود مؤلفه شکندگی که نماینده عوامل خطر ناشناخته است، تا حدودی عملکرد مدل بهتر از مدل کاکس است و این موضوع با مقایسه ملاک *DIC* دو مدل به وضوح قابل مشاهده است.



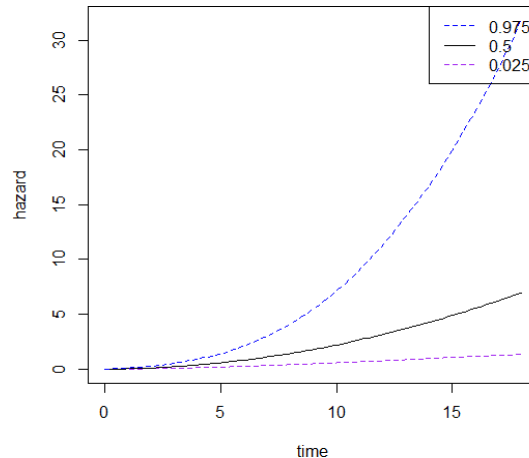
شکل ۱: موقعیت قرارگیری بیماران مبتلا به کووید ۱۹ در شهر تهران علامت + نشان دهنده داده‌های سانسور شده و دایره‌ها مشاهدات هستند



شکل ۲: برآورد تابع بقای بیماران بر حسب تعداد روزهای بستری

بحث و نتیجه‌گیری

در این مقاله پس از معرفی مدل بقای فضایی زمین آماری با رویکرد بیزی به تحلیل داده‌های زمان بستری بیماران مبتلا به کووید ۱۹ در شهر تهران پرداخته شد مقایسه نتایج حاصل از برازش مدل‌های مختلف با استفاده از ملاک DIC نشان داد که مدل فضایی نتایج مطلوب‌تری را ارائه نموده است. بر این اساس ضریب مثبت متغیر تبیینی جنسیت با در نظر گرفتن زنان به عنوان گروه مرجع نشان می‌دهد که نرخ خطر مرگ در مردان $2/280 = e^{0.824}$ یعنی بیش از دو برابر زنان است. ضریب مثبت متغیر سن حاکی از آن است که با افزایش هر سال در متغیر سن خطر مرگ بیمار $1/109 = e^{0.104}$ یعنی



شکل ۳: برآورد تابع خطر بیماران بر حسب تعداد روزهای بستری

حدود ده درصد افزایش می‌یابد. نتیجه بسیار مهم و قابل توجه در متغیر بیماری‌های زمینه‌ای نشان می‌دهد در صورت وجود بیماری زمینه‌ای در افراد خطر مرگ $e^{1/473} = 4/362$ بیش از چهار برابر افرادی است که بیماری زمینه‌ای ندارند. شکل شماره ۲ تابع زمان بقای بیماران بر حسب تعداد روزهای بستری را نشان می‌دهد بر این اساس شانس بقای بیمارانی که بیش از ۱۰ روز در بیمارستان بستری شده‌اند بسیار کم و به سمت صفر میل می‌کند. شکل شماره ۳ برآورد تابع خطر بیماران بر حسب تعداد روزهای بستری در بیمارستان را نشان می‌دهد. نکته قابل توجه این است که تابع خطر با افزایش تعداد روزهای بستری صعودی است و بعد از روز پانزدهم خطر فوت حدود ۵ برابر افزایش پیدا خواهد کرد. نتایج این مطالعه نشان می‌دهد که مهم‌ترین عامل فوت در بیماران مبتلا به کووید ۱۹ وجود بیماری‌های زمینه‌ای است که موجب افزایش بیش از چهار برابری خطر فوت در این دسته از بیماران می‌شود. این نتیجه لزوم رعایت هر چه بیشتر پروتکل‌های بهداشتی و مراقبتی توسط افرادی که داری بیماری زمینه‌ای هستند را توصیه می‌کند.

تقدیر و تشکر

نویسنده از حمایت قطب علمی تحلیل داده‌های وابسته فضایی و فضایی-زمانی دانشگاه تربیت مدرس کمال تشکر و قدردانی را دارد.

مراجع

Banerjee, S., Carlin, B. P. and Gelfand, A. E. (2004), *Hierarchical Modeling and Analysis for Spatial Data*, Chapman and Hall, Boca Raton.

Cox, D. R. (1972), Regression Models and Life-Tables, *Journal of the Royal Statistical Society, Series B*, **34**, 187-220.

Darmofal, D. (2009), Bayesian spatial survival models for political event processes, *Am. J. Polit. Sci.*, **53**, 241-257

Glass, D.; Ogborn, M.; Sutherland, I. (1963). John Graunt and His Natural and Political Observations [and Discussion]. *Proceedings of the Royal Society of London, Series B, Biological Sciences*, **159**, 2–37.

Motarjem, K., Mohammadzadeh, M. and Abyar, A. (2020), Geostatistical Survival Model with Gaussian Random Effect, *Statistical Papers*, **61**, 85-107.

Ramsay, T., Burnett, R. and Krewski, D. (2003), Exploring Bias in a Generalized Additive Model for Spatial Air Pollution Data, *Environmental Health Perspectives*, **111**, 1283-1288.

Schnirring, Lisa (2020). Doubts rise about China's ability to contain new coronavirus, *CIDRAP*, Retrieved 26 January 2020.

Turechek, W. W. and Madden, L. V. (2002), A Generalized Linear Modeling Approach for Characterizing Disease Incidence in Spatial Hierarchy, *Phytopathology*, **93**, 458-466.