Journal of Statistical Sciences, Spring and Summer, 2024
Vol. 18, No. 1, pp 1-17
DOI: 10.52547/jss.18.1.01

Geostatistical Survival Data Analysis with Bayesian Approach

Pakdel, M., Motarjem, K.
Department of Statistics, Tarbiat Modares University, Tehran, Iran.

Corresponding author: K. Motarjem, k.motarjem@modares.ac.ir

Received: 27/12/2023 **Revised:** 8/3/2024 **Accepted and Published Online:** 11/3/2024.

Introduction

Survival analysis examines the time until an event occurs, such as death or disease. Typically, independence among observations is assumed. However, survival times depend on spatial survival data due to the spatial arrangement of the data. We employ spatial survival models that consider spatial dependence to analyze such data. These models allow for a more accurate analysis of spatial survival data and are especially useful when the spatial location influences survival time. The goal is to explore the relationship between event occurrence time and explanatory variables in spatial survival data.

Material and Methods

A Bayesian approach was employed to analyze and implement the spatial survival model due to the computational complexity and challenges in parameter estimation. This approach allows for fast and accurate model parameter estimation, enabling large-scale dataset analysis. The framework provides flexible and robust parameter estimation by utilizing Bayesian inference, which incorporates prior information and updates it with observed data. The Bayesian approach also facilitates uncertainty quantification through posterior distributions, offering more comprehensive insights than point estimates from frequentist methods. Additionally, it enables the analysis of high-dimensional datasets and provides a stronger statistical framework for addressing the complexities and challenges in parameter estimation of spatial survival models.

Results and Discussion
The spatial survival model's advantage stems from its ability to account for spatial dependence among observations, allowing for a more accurate representation of the underlying survival process. By comparing the Deviance Information Criterion (DIC) and Bayesian Information Criterion (BIC) values, it was evident that the spatial survival model demonstrated a better fit to the data and offered a more parsimonious representation compared to the Cox proportional hazards and accelerated failure time models. The spatial survival model's lower DIC and BIC values indicate its superior performance in capturing the complex spatial patterns present in the leukemia data, particularly when censoring is involved. The results of the leukemia data analysis showed that the spatial survival model consistently outperformed traditional models, such as Cox proportional hazards, Frailty, and accelerated failure time models, across all censoring scenarios.

Conclusion
The spatial survival model outperformed traditional models (Cox proportional hazards and accelerated failure time) in all censoring scenarios. The model accurately captured spatial patterns by incorporating spatial dependence and exhibited superior performance with lower DIC and BIC values. Considering spatial correlation is crucial in survival analysis, especially for high-dimensional and large-scale datasets. The Bayesian approach provided a robust framework for efficient parameter estimation and allowed for accurate inference in spatial survival modeling. The spatial survival model, implemented using Bayesian methods, is a powerful tool for analyzing spatially dependent survival data. It offers precise parameter estimation and facilitates the exploration of complex spatial patterns, contributing to improved decision-making and insights in fields such as epidemiology, environmental studies, and public health.

Keywords: Survival Analysis, Spatial Correlation, Spatial Survival Model, Bayesian Approach.

Mathematics Subject Classification (2010): 62N01, 62M30.

©The Author(s). The Publisher is Iranian Statistical Society.
This is an open access article distributed under the terms and conditions of (CC BY-NC 4.0)

مجله علوم آماری، بهار و تابستان ۱۴۰۳
جلد ۱۸، شماره ۱، ص ۱ -- ۱۷
DOI: 10.52547/jss.18.1.01
مقاله پژوهشی

تحلیل داده‌های بقای زمین آماری با رویکرد بیزی

میلاد پاکدل، کیومرث مترجم
گروه آمار، دانشگاه تربیت مدرس

نویسنده مسئول : کیومرث مترجم، k.motarjem@modares.ac.ir

تاریخ دریافت: ۱۴۰۲/۱۰/۶ تاریخ بازنگری: ۱۴۰۲/۱۲/۱۸ تاریخ پذیرش و انتشار: ۱۴۰۲/۱۲/۲۱

چکیده: گاهی در عمل زمان تا وقوع یک رویداد می‌تواند متاثر از مکان باشد که این نوع از مشاهدات داده بقای فضایی نامیده می‌شوند. برآورد سریع و دقیق پارامترها در مدل بقای فضایی بواسطه پیچیده بودن تابع درستنمایی یکی از چالش‌های استفاده از رویکرد فراوانی‌گرا است که همین امر استفاده از رویکرد بیزی در تحلیل بقا را پررنگ نموده است. در یک مدل بقای فضایی بیزی، همبستگی فضایی بین زمان‌های رویداد با استفاده از یک مدل زمین‌آماری تبیین می‌شود. در این مقاله در قالب یک مطالعه شبیه‌سازی به برآورد پارامترهای مدل‌های کلاسیک و فضایی بقا پرداخته می‌شود و عملکرد هر کدام از مدل‌ها در برازش به داده‌های بقای شبیه‌سازی شده مورد ارزیابی قرار می‌گیرد. در نهایت نشان داده می‌شود که مدل بقای فضایی در تحلیل داده‌های سرطان خون کارایی بهتری نسبت به مدل‌های مرسوم دارد.

واژه‌های کلیدی: تحلیل بقا، همبستگی فضایی، مدل بقای فضایی، رویکرد بیزی.
کد موضوع‌بندی ریاضی (۲۰۱۰): 62N01، 62M30.

۱ مقدمه

یکی از شاخه‌های کاربردی علم آمار تحلیل بقا[1] است، که به مطالعه مدت زمان بقای یک واحد یا فرد در شرایط خاص گفته می‌شود. تحلیل داده‌های بقا یا زمان تا وقوع رویداد[2] نقش مهمی در پزشکی، همه‌گیرشناسی، زیست شناسی، جمعیت شناسی، اقتصاد، مهندسی و سایر زمینه‌ها دارد. یکی از مهمترین اهداف اصلی تحلیل بقا، تعیین رابطه بین

©نویسند(گان). ناشر انجمن آمار ایران است.
این مقاله با دسترسی آزاد تحت شرایط و ضوابط (CC BY-NC 4.0) توزیع شده است.

[1] Survival analysis
[2] Time to event

زمان وقوع رویداد و متغیرهای تبیینی است (ثمرین و همکاران، ۲۰۱۷). در تحلیل بقا عموماً فرض بر استقلال مشاهدات است، اما در برخی مسائل کاربردی بقا با داده‌هایی مواجه می‌شویم که مستقل نبوده و وابستگی بین زمان‌های بقا ناشی از موقعیت و مکان قرارگرفتن داده‌ها در فضای مورد بررسی است. تحلیل این گونه مشاهدات که داده‌های بقای فضایی نامیده می‌شوند با روش‌های معمول از دقت کافی برخوردار نبوده و لازم است به نحوی ساختار همبستگی داده‌ها در تحلیل آماری لحاظ شود (محمدزاده، ۱۳۹۸). در عمل تحلیل این نوع از داده‌ها با یک رده جدید از مدل‌های بقا تحت عنوان مدل‌های بقای فضایی امکان‌پذیر خواهد بود. این مدل‌ها تأثیر وابستگی فضایی داده‌ها را در مدل‌سازی منظور می‌کنند. بکارگیری این مدل‌ها در شرایطی مفید است که زمان بقای افراد یا واحدهای آزمایشی تحت تأثیر موقعیت مکانی آنها باشد (گلفند و همکاران، ۲۰۱۷). اجرا و پیاده‌سازی مدل‌های بقای فضایی زمین آماری در مسائل کاربردی به ویژه در حالات ناگاوسی بودن میدان تصادفی امکان‌پذیر است (مترجم، ۱۴۰۰).

با توجه به آنکه داده‌های فضایی را می‌توان به دو دسته زمین آماری و مشبکه‌ای تقسیم‌بندی نمود بر همین اساس مدل‌های بقای فضایی نیز به دو دسته اصلی برای تحلیل داده‌های زمین آماری و مشبکه‌ای تقسیم می‌شوند مدل بقای فضایی مورد بررسی در این مطالعه برای تحلیل داده‌های فضایی از نوع زمین آماری مناسب است. روش‌های زیادی وجود دارد که می‌توان از طریق رویکرد بیزی اثرات فضایی را در مدل بقا لحاظ کرد. بنا بر مطالعات بانرجی وهمکاران (۲۰۰۳) و همچنین پژوهش‌های ژو و همکاران (۲۰۱۷) رویکرد بیزی برای وارد نمودن وابستگی‌های فضایی بر روی اثرات تصادفی، از همبستگی مناطق مجاور آنها بهره می‌برد و به همین سبب رویکرد بیزی فاقد پیچیدگی‌های محاسباتی موجود در رویکرد فراوانی‌گر است و نگرانی دستیابی به یک ماکسیمم محلی به جای ماکسیمم کلی که در رویکرد فروانی‌گرا و در هنگام رویارویی با توابع درست‌نمایی پیچیده وجود دارد، در این رویکرد وجود نخواهد داشت. به همین علت در مطالعه حاضر چارچوب استنباطی بیزی مورد استفاده قرارگرفته است. ادامه مقاله به‌صورت زیر تنظیم شده است که در بخش ۲ به معرفی مفاهیم اساسی در تحلیل بقا پرداخته می‌شود. در بخش ۳، به ترتیب مهمترین مدل‌های کلاسیک بقا، شامل مدل‌های کاکس، شکنندگی و زمان شکست شتابیده مورد بررسی قرار می‌گیرد در بخش ۴ مدل بقای فضایی زمین آماری با رویکرد بیزی معرفی می‌شود. در بخش ۵ به منظور ارزیابی عملکرد مدل‌های معرفی شده یک مطالعه شبیه‌سازی انجام شده است و در نهایت کاربست روش‌های معرفی شده در تحلیل داده‌های سرطان خون مورد بحث و بررسی قرارگرفته است.

۲ مدل‌های کلاسیک در تحلیل بقا

در تحلیل بقا، مدل‌های کلاسیک به عنوان ابزارهای مهمی برای بررسی و پیش‌بینی زمان بقای افراد یا واحدهای آزمایشی مورد استفاده قرار می‌گیرند. این مدل‌ها عمدتاً بر اساس فرضیاتی درباره توزیع زمان بقا و عوامل مؤثر بر آنها ساخته می‌شوند. در این بخش به معرفی پرکاربردترین مدل‌های مرسوم مورد استفاده محققین در تحلیل داده‌های بقا پرداخته می‌شود (لیو و همکاران، ۲۰۲۳).

یکی از مهمترین مدل‌ها در تحلیل بقا، مدل مخاطرات متناسب کاکس است که توسط کاکس (۱۹۷۲) پیشنهاد

شده است. یکی از فرض‌های اساسی مدل کاکس این است که متغیرهای تبیینی با زمان مرتبط نبوده و در طول زمان تغییر نمی‌کنند. به عبارت دیگر، فرض می‌شود که نسبت خطر برای هر دو واحدآزمایشی دلخواه مستقل از زمان است. مدل کاکس به‌صورت $h(t|X) = h_\circ(t)\exp(\beta' X)$ تعریف می‌شود، که در آن $X = (X_۱, \ldots, X_k)$ بردار متغیرهای تبیینی، $\beta' = (\beta_۱, \ldots, \beta_k)$ بردار اثرهای ثابت و $h_\circ(t)$ تابع خطر پایه است. برای برآورد پارامترهای مدل مخاطرات متناسب کاکس با روش ماکسیمم درست‌نمایی لازم است، تابع درستنمایی مدل محاسبه گردد. بر اساس مشاهدات (t_i, δ_i, X)، $i = ۱, \ldots, n$، تابع خطر پایه $h_\circ(t)$ و تابع بقای مدل مخاطرات متناسب کاکس به‌صورت $S(t) = \exp(-H(t))$ است.، لذا تابع چگالی زمان‌های بقا به‌صورت $f(t) = h(t)S(t)$ حاصل می‌شود. بنابراین تابع درست‌نمایی مدل مخاطرات متناسب به‌صورت

$$L(\beta) = \prod_{i=۱}^{n} \left(h_\circ(t_i)\exp(\beta' X_i)\right)^{\delta_i} \exp(-H(t_i))^{\exp(\beta' X_i)} \tag{۱}$$

است، که در آن

$$\delta_i = \begin{cases} ۱ & T_i \leq C_i \\ \circ & T_i > C_i \end{cases}$$

متغیر تصادفی سانسور و C_i زمان سانسور واحد iام است. درصورت معلوم بودن شکل تابع خطر پایه با ماکسیمم کردن تابع درست‌نمایی (۱) نسبت به پارامترهای رگرسیونی، برآورد ضرایب β' به راحتی حاصل می‌شود. بکارگیری مدل کاکس با رویکرد بیزی بجای رویکرد فراوانی‌گرا موجب می‌شود که یک مبادله ساده بین انعطاف‌پذیری مدل کاکس بیزی نسبت به اشکال مختلف تابع خطر پایه و مدل کاکس فراوانی‌گرا اتفاق بیفتد، که بر مبنای آن به‌توان رویکرد بیزی را بکار برد در این رویکرد تابع درست‌نمایی مدل کاکس بیزی به‌صورت

$$L(\beta; t, x, \delta) \propto \prod_{i=۱}^{I} \left(h_\circ(t_i; x_i)\right)^{\delta_i} \exp(-H(t_i))^{\exp(\beta' X_i)}$$

است.، که با در نظر گرفتن توزیع‌های پیشین مناسب برای پارامترهای رگرسیونی و همچنین پارامترهای تابع خطر پایه، فرم پسینی این مدل در حالت کلی به‌صورت

$$\pi(\beta, \lambda | t, x, \delta) \propto L(\beta; t, x, y)\pi(\beta)\pi(\lambda) \tag{۲}$$

است، که در آن t مجموعه‌ای از زمان‌های وقوع رویداد، λ پارامترهای تابع خطر پایه است. در رابطه (۲) به دلیل عدم وجود فرم بسته برای توزیع پسینی، جهت انجام استنباط آماری و برآورد پارامترهای مدل، از روش‌های زنجیره

مارکوف مونت کارلو و نمونه‌گیری از توزیع پسینی می‌توان پارامترهای مدل را برآورد نمود (ژانگ و همکاران، ۲۰۲۳).

۲.۱ مدل شکنندگی

مدل شکنندگی[1]، یک رده دیگر از مدل‌های پرکاربرد در تحلیل بقا است که به وسیله افزودن یک اثر تصادفی به‌صورت ضربی، اثر عوامل خطر ناشناخته در زمان‌های بقا را کنترل می‌کند و به‌صورت

$$h(t|W_i, h_\circ, X) = W_i h_\circ(t) \exp(\beta' X_i), i = ۱, \ldots, n \tag{۳}$$

تعریف می‌شود که به آن مدل کاکس اصلاح‌شده یا همان مدل شکنندگی بقا گفته می‌شود و W_i به عنوان مولفه شکنندگی در نظر گرفته می‌شود. در مدل (۳) فرض بر این است که شکنندگی برای هر واحد آزمایشی به‌صورت مجزا و مستقل از بقیه افراد است. البته مدل فوق برای حالتی که بین مولفه‌های شکنندگی همبستگی وجود داشته باشد توسط ژانگ و همکاران (۲۰۲۳) تعریف شده است.

برای برآورد پارامترها در مدل شکنندگی تابع درست‌نمایی به‌صورت

$$L(\beta|W) = \prod_{i=۱}^{n} \left(W_i h_\circ(t_i) e^{\beta' X_i}\right)^{\delta_i} \exp\left(-W_i H(t_i)\right)^{\exp(\beta' X_i)}$$

است. تابع بقا به‌صورت

$$S(t|W) = \exp(\int_\circ^t h(s|W)\mathbf{ds} = \exp(-W \int_\circ^t h_\circ(s)\mathbf{ds}) = \exp(-W H_\circ(t))$$

به دست خواهد آمد. بنابراین می‌توان تابع بقا را به‌صورت

$$S(t) = E(S(t|W)) = E(\exp(-W H_\circ(t))) = LA(H_\circ(t))$$

[1] Frailty model

نوشت، که در آن LA تبدیل لاپلاس است. بنابراین توابع به‌صورت

$$\begin{aligned} f(t) &= -h_\circ(t)LA'(H_\circ(t)) \\ h(t) &= h_\circ(t)\frac{LA'(H_\circ(t))}{LA(H_\circ(t))} \\ \mathrm{E}(W) &= -LA'(\circ), \\ \mathrm{Var}(W) &= LA''(\circ) - (LA'(\circ))^{۲} \end{aligned}$$

به‌دست می‌آیند، که در آن LA' و LA'' به ترتیب مشتق اول و دوم تبدیل لاپلاس هستند. در رویکرد بیزی تابع درست‌نمایی مدل شکنندگی بیزی به‌صورت

$$L(\beta, W; t, x, \delta) \propto \prod_{i=۱}^{I} h_\circ(t_i; x_i)^{\delta_i} \exp(-H_\circ(t_i))^{\exp(\beta' X_i + W_i)}$$

است که در آن $H_\circ$ تابع خطر تجمعی است.. در اینجا با در نظر گرفتن توزیع‌های پیشین مناسب برای پارامترهای رگرسیونی، مولفه شکنندگی و همچنین پارامترهای تابع خطر پایه فرم پسینی این مدل در حالت کلی به‌صورت

$$\pi(\beta, W, \lambda | t, x, \delta) \propto L(\beta, W; t, x, y)\pi(W)\pi(\beta)\pi(\lambda)$$

حاصل می‌شود که در آن t مجموعه‌ای از زمان‌های وقوع رویداد، λ پارامترهای تابع خطر پایه و W مولفه شکنندگی است. مشابه آنچه که در خصوص مدل مخاطرات متناسب کاکس بیان شد، به دلیل عدم وجود فرم بسته برای توزیع پیشین مدل شکنندگی تک متغیره جهت انجام استنباط آماری از روش‌های زنجیره مارکوف مونت کارلو و نمونه‌گیری از توزیع پسینی می‌توان پارامترهای مدل را برآورد نمود.

۲.۲ مدل زمان شکست شتابیده

اگر چه روش‌های نیمه‌پارامتری مانند مدل مخاطرات متناسب کاکس برای تحلیل داده‌های بقا بسیار کاربردی هستند، اما یکی از فرضیات اساسی استفاده از این مدل‌ها برقراری فرضیه مخاطرات متناسب است که در عمل و در بسیاری از مواقع این فرض محقق نمی‌شود، لذا در این شرایط، مدل زمان شکست شتابیده[۱] AFT جایگزینی مناسب برای مدل مخاطرات متناسب کاکس در تحلیل داده‌های بقا است. مدل زمان شکست شتابیده به‌صورت

$$\log(T_i) = X_i\beta' + \sigma\epsilon_i, i = ۱, \ldots, n$$

[1] Accelerated failure time model

تعریف می‌شود، که در آن T_iها زمان‌های وقوع رویداد، X_iها بردارهای متغیرهای تبیینی، β' بردار ضرایب رگرسیونی، σ پارامتر مقیاس و ϵ_iها خطای تصادفی مستقل وهم توزیع هستند. فرض کنید $f(t_L)$ تابع چگالی T_L باشد و $f_\circ$ تابع چگالی ϵ_L را نشان می‌دهد. داده‌های بقای سانسور شده را در نظر بگیرید (t_i, δ_i, x_i) که در آن t_i زمان مشاهده شده تصادفی است. سپس تابع درستنمایی را می‌توان به‌صورت

$$L = \prod_{i=1}^{n} f_\circ(z_i)^{\delta_i} S_\circ(z_i)^{1-\delta_i}$$

نوشت، که در آن $L(\beta, \sigma|t_i, \delta_i) = (\frac{1}{\sigma t_i} f_\circ(z_i))^{\delta_i} S_\circ(z_i)^{1-\delta_i}$ و $z_i = \frac{\log(t_i) - X_i\beta'}{\sigma}$ است. در رویکرد بیزی تابع درستنمایی مدل زمان شکست شتابیده بیزی به‌صورت

$$L(\beta, \sigma; t_i, z_i, \delta_i) \propto (\frac{1}{\sigma t_i} f_\circ(z_i))^{\delta_i} S_\circ(z_i)^{1-\delta_i}$$

است. بنابراین با توجه به تابع درست‌نمایی و در نظر گرفتن توزیع‌های پیشین مناسب برای پارامترهای رگرسیونی، فرم پسینی مدل در حالت کلی به‌صورت

$$\pi(\beta, \sigma|X) \propto L(\beta, \sigma; t_i, z_i, \delta_i)\pi(\beta)\pi(\sigma)$$

به‌دست می‌آید که در آن t مجموعه‌ای از زمان‌های بقا، σ پارامتر مقیاس و δ_i نشانگر سانسور شدگی در داده‌ها است (تانگ و همکاران، ۲۰۲۳). در این حالت نیز برای برآورد پارامترها، می‌توان از روش‌های نمونه‌گیری مانند مارکوف–مونت کارلو استفاده کرد. این روش‌ها این امکان را فراهم می‌آورد تا نمونه‌هایی از تابع چگالی احتمال پسین گرفته شوند و با استفاده از آنها، برآوردهای بهتری از پارامترها به‌دست آید.

۳ مدل بقای فضایی

مدل خطر بقای فضایی برای برازش به داده‌ها به‌صورت

$$h(t_i|\psi, Y_i) = h_\circ(t_i; w) \exp(\beta' X_i + Y_i) \tag{۴}$$

تعریف می‌شود، که در آن t_i زمان مشاهده شده برای فرد iام، β بردار پارامترهای رگرسیونی، X_i بردار متغیرهای کمکی برای فرد iام، $h_\circ(t; w)$ تابع خطر پایه (که در آن t زمان بقا و w پارامترهای تابع خطر پایه $h_\circ$ هستند) در

این مدل Y معرف یک میدان تصادفی است که به عنوان مولفه شکنندگی همبسته فضایی[1] عمل می‌کند. به جهت سهولت در نمادگذاری، پارامترها بصورت $\psi = (\beta, w, \eta)$ تعریف شده است، که به ترتیب اثرات متغیرهای تبیینی، پارامترهای تابع خطر پایه و پارامترهای تابع کوواریانس فضایی را نشان می‌دهند در معادله (۴) به طور معمول یک میدان تصادفی گاوسی برای در نظر گرفتن همبستگی فضایی در نظر گرفته می‌شود. برای سادگی در تفسیر، تنظیمات باید به گونه‌ای انجام می‌شود که $\mathrm{E}[\exp Y] = ۱$ و همچنین واریانس حاشیه‌ای آن $\sigma^۲$ است (ثمرین و همکاران، ۲۰۲۱). از توابع کوواریانس معتبر فضایی که در مدل بقای فضایی مورد استفاده قرار می‌گیرند می‌توان به مدل‌های نمایی اشاره نمود، در این ساختار، کوواریانس بین هر دو نمونه با فاصله d از هم به‌صورت $K(d) = \sigma^۲ e^{-d/\phi}$ مشخص می‌کند، که در آن $\sigma^۲$ واریانس و ϕ پارامتر دامنه میدان است. واضح است که هرچه ϕ بزرگتر باشد، دامنه وابستگی فضایی بیشتری در میدان وجود دارد. با فرض آنکه $\xi_i = (\psi, Y_i) = (\beta, w, \eta, Y_i)$، تابع چگالی برای یک مدل مخاطرات متناسب فضایی به‌صورت

$$\begin{aligned} f(t;\xi_i) &= h(t;\xi_i)S(t;\xi_i), \\ &= h(t;\xi_i)\exp -H(t;\xi_i), \\ &= \exp[X_i\beta + Y_i]h_\circ(t;w)\exp\left[-\exp[X_i\beta + Y_i]\int_\circ^t h_\circ(s;w)\mathbf{d}s\right], \\ &= \exp[X_i\beta + Y_i]h_\circ(t;w)\exp[-\exp[X_i\beta + Y_i]H_\circ(t;w)]; \end{aligned}$$

نوشته می‌شود، که در آن تابع خطر، بقا و خطر تجمعی همگی به پارامترهای ξ وابسته هستند، بنابراین از نمادگذاری $f(t;\xi_i)$، $h(t;\xi_i)$، $S(t;\xi_i)$ و $H(t;\xi_i)$ برای نمایش این توابع استفاده شده است. به راحتی می‌توان تابع توزیع تجمعی مدل مخاطرات متناسب بقای فضایی را به‌صورت

$$F(t;\xi_i) = ۱ - \exp[-\exp[X_i\beta + Y_i]H_\circ(t;w)]. \tag{۵}$$

بدست آورد. با در دست داشتن تابع توزیع زمان‌های بقای (۵)، می‌توان با استفاده از رویکرد بیزی به استنباط آماری در خصوص پارامترها در مدل بقا پرداخت. در رویکرد بیزی فرم پسینی به‌صورت

$$\pi(\xi|\text{data}) \propto \pi(\text{data}|\xi)\pi(\xi)$$

برای برآورد پارامترها در مرحله بعد، پس از اجرای شبیه‌سازی $MCMC$، نمونه‌های به دست آمده از توزیع پسین مورد بررسی قرار می‌گیرند همچنین در پسا $MCMC$ می‌توان آماره‌های مانند میانگین، واریانس و توزیع تجمعی را نیز برآورد کرد که از آنها برای تحلیل و بررسی ویژگی‌های میدان تصادفی استفاده می‌شود. برای اجرای روش

[1] Spatially correlated frailties

پسا زنجیره مارکوفی به جهت برازش یک مدل بقا فضایی و در نهایت پیش‌گویی مقدار میدان تصادفی در نقاط فاقد مشاهده $\tilde{Y}_1, \ldots, \tilde{Y}_m$ می‌توان از انتگرال

$$\pi(\tilde{Y}_{1:m}|\text{data}) = \int \pi(\tilde{Y}_{1:m}|Y_{1:n}, \eta, \text{data})\pi(Y_{1:n}, \eta|\text{data})dY_{1:n}d\eta$$

استفاده کرد. با فرض اینکه $\tilde{Y}_{1:m}$ به طور مشروط مستقل از داده‌های داده شده $Y_{1:n}$ است، می‌توان یک برآورد نااریب برای عبارت فوق به‌صورت

$$\pi(\tilde{Y}_{1:m}|\text{data}) \approx \frac{1}{n}\sum_{i=1}^{n} \pi(\tilde{Y}_{1:m}|Y^i_{1:n}, \eta^i) \qquad (۶)$$

به‌دست آورد، که برای مثال η^i، iمین نمونه پارامتر η از زنجیره مارکوف مونت کارلو است. به دلیل اینکه توزیع $\pi(\tilde{Y}_{1:m}, Y^i_{1:n}|\eta^i)$ گاوسی است توزیع شرطی $\pi(\tilde{Y}_{1:m}|Y^i_{1:n}, \eta^i)$ نیز به‌صورت دقیق در دسترس است (روو و هلد، ۲۰۰۵). بنابراین تقریب مونت کارلو استفاده شده در رابطه (۶) می‌تواند از یک توزیع گاوسی آمیخته تبعیت کند. بر این اساس برای به دست آوردن برآورد نااریب چندک‌های تابع چگالی حاشیه‌ای از رابطه

$$\int_{-\infty}^{y} \pi(\tilde{Y}_j|\text{data})d\tilde{Y}_j \approx \frac{1}{n}\sum_{i=1}^{n}\int_{-\infty}^{y} \pi(\tilde{Y}_j|Y^i_{1:n}, \eta^i)d\tilde{Y}_j$$

استفاده می‌شود، که انتگرال در سمت راست عبارت بالا را می‌توان با روش‌های عددی محاسبه کرد. علی‌رغم آنکه محاسبات فوق قابل انجام است، اما در حالت کلی و برای حجم نمونه‌های نسبتاً بزرگ از هزینه محاسباتی بالایی دارند. این هزینه برای هر توزیع شرطی $\pi(\tilde{Y}_{1:m}|Y^i_{1:n}, \eta^i)$، از مرتبه $O(n^3)$ است. در این راستا، روش محاسبه مستقیم معرفی شده توسط دیگل و همکاران (۱۹۹۸)، به طور اصولی ساده‌تر است و شامل اجرای زنجیره $MCMC$ برای نمونه‌برداری از توزیع $\pi(Y, \beta, w, \eta|\text{data})$ تا رسیدن به همگرایی است. در این مرحله بردار پارامتر به وسیله $\tilde{Y}$ها و بکارگیری طرح گیبز برای نمونه‌گیری از توزیع $\pi(\tilde{Y}, Y, \beta, w, \eta|\text{data})$ با استفاده از تجزیه

$$\pi(\tilde{Y}, Y, \beta, w, \eta|\text{data}) = \pi(\tilde{Y}|, Y, \beta, w, \eta, data)\pi(Y, \beta, w, \eta|\text{data})$$

حاصل می‌شود. با توجه به این مدل، ویژگی استقلال شرطی به

$$\pi(Y, \beta, w, \eta|\text{data}) = \pi(Y, \beta, w, \eta|\tilde{Y}, \text{data})$$

اشاره دارد. لذا نمونه‌گیری $MCMC$ که در قبل استفاده شد را می‌توان همچنان برای نمونه‌گیری از چگالی $\pi(Y, \beta, w, \eta|\text{data})$ استفاده کرد و در هر یک از تکرارها باقی‌مانده نمونه‌گیری $\pi(\tilde{Y}|, Y, \beta, w, \eta, \text{data})$ یک کاربرد مستقیم از توزیع چند متغیره شرطی گاوسی است. (لوپیس و همکاران، ۲۰۲۳).

۴ شبیه‌سازی و مثال کاربردی

در این بخش با استفاده از یک مطالعه شبیه‌سازی به مقایسه کارایی مدل بقای فضایی با رویکرد بیزی و مدل‌های کاکس، شکنندگی و زمان شکست شتابیده در حضور همبستگی فضایی در داده‌های بقا پرداخته می‌شود. برای تولید داده بقا در این مطالعه از دو متغیر تبیینی $X_۱$ و $X_۲$ که به ترتیب دارای توزیع یکنواخت پیوسته بر بازه $[۵, ۵۰]$ و توزیع دوجمله‌ای با شانس موفقیت ۰٫۵ استفاده شده است که ضرایب این دو متغیر تبیینی $\beta_۱ = ۱٫۰۳$ و $\beta_۲ = ۱٫۰۲$ در نظر گفته شده است. برای تولید داده بقا براساس دو متغیر تبیینی تابع خطر به‌صورت

$$h(t) = h_۰(t) \exp\left(۱٫۰۳X_۱ + ۱٫۰۲X_۲\right)$$

در نظر گرفته شده است. در این مقاله از روش معرفی شده توسط مترجم و همکاران (۲۰۲۰) برای تولید داده‌های بقای فضایی استفاده شده است و همچنین داده‌های بقا براساس متغیرهای تبیینی و از طریق تابع خطر پایه وایبول با پارامترهای $k = ۱٫۳$ و $\lambda = ۰٫۰۳$، تولیدشده‌اند و مقادیر $\sigma = ۰٫۷$ و $\phi = ۰٫۱$ به ترتیب واریانس و دامنه تابع کوواریانس فضایی نمایی در نظر گرفته شده است. با توجه به این که در این مطالعه از رویکرد بیزی استفاده شده است و بدلیل آنکه که Y تحققی از یک میدان تصادفی گاوسی پنهان است لذا مستقیماً با آن کار نمی‌شود و نمی‌توان پیشینی برای آن در نظر گرفت، بلکه به جای آن با Γ که در آن $Y = -\sigma^۲/۲ + \sum_{\sigma,\phi}^{۱/۲} \Gamma$ و دارای توزیع پیشین $\Gamma \sim N(۰, \mathbf{1})$ است استفاده شده است، که در آن ۱ ماتریس یکه[1] و $\sum_{\sigma,\phi}^{۱/۲} +$ تجزیه چولسکی ماتریس کوواریانس است. بر این مبنا توزیع پیشین پارامترهای رگرسیونی برابر با $N \sim (۰, ۱۰)$، توزیع پیشین پارامترهای تابع خطر $N \sim (۰, ۱۰)$ و توزیع پیشین پارامترهای تابع کوواریانس فضایی $N \sim (۰, ۰٫۵)$ و $N \sim (-۲, ۰٫۵)$، در نظر گرفته می‌شود. در برازش مدل‌ها با روش بیزی طول زنجیر ۲,۰۰۰,۰۰۰، دوره داغیدن ۵۰۰,۰۰۰ و طول گام زنجیره ۲۰ در نظر گرفته شده است، به این منظور نمونه‌های از داده‌های بقای فضایی با حجم‌های ۲۰۰، ۴۰۰ و ۶۰۰ تولیده شده است. بواسطه اهمیت وجود سانسور به عنوان یک جز جدایی‌ناپذیر در تحلیل بقا در این شبیه‌سازی از سانسور راست ۲۰ و ۸۰ درصدی، استفاده شده است و برای دقت بیشتر در نتایج عملیات شبیه‌سازی به ترتیب ۵۰، ۱۰۰ و ۲۰۰بار تکرار شده‌اند. پس از تولید داده بقای فضایی با شرایط ذکر شده مدل‌های بقای معرفی شده به داده‌ها برازش شده است و نتایج آن در جداول ۱ تا ۳ آورده شده است. درجدول ۱ نتایج برازش مدل‌های بقا به داده‌های شبیه‌سازی شده در حجم ۲۰۰ و در سه تکرار ۵۰، ۱۰۰ و ۲۰۰ و سانسور راست ۲۰ و ۸۰ درصدی نشان

[1] Identity matrix

جدول ۱. نتایج برازش مدل‌های بقا در حجم نمونه ۲۰۰ و سه تکرار با سانسور ۲۰ و ۸۰ درصد

سانسور	تکرار	مدل	$(\beta_1, Std.)$	$(\beta_2, Std.)$	DIC	BIC
	۵۰	بقای فضایی	(۱٫۰۱۳۹, ۰٫۰۱۱۱)	(۱٫۰۱۶۸, ۰٫۰۱۲۲)	۲۸۸٫۳۳۷۷	۲۹۲٫۳۲۱۵
		کاکس	(۱٫۰۰۹۵, ۰٫۰۱۴۹)	(۱٫۰۱۰۱, ۰٫۰۵۷۰)	۲۹۲٫۹۵۱۹	۲۹۷٫۵۲۱۴
		AFT	(۱٫۰۰۹۹, ۰٫۰۱۳۰)	(۱٫۰۱۲۱, ۰٫۰۵۵۰)	۲۹۱٫۵۵۵۵	۲۹۴٫۲۱۸۷
		شکنندگی	(۱٫۰۱۴۲, ۰٫۰۱۲۱)	(۱٫۰۱۳۹, ۰٫۰۵۰۵)	۲۹۰٫۶۷۷۸	۲۹۳٫۷۰۳۲
%۲۰	۱۰۰	بقای فضایی	(۱٫۰۱۴۸, ۰٫۰۰۹۵)	(۱٫۰۱۸۳, ۰٫۰۱۱۳)	۳۱۷٫۵۳۶۶	۳۲۱٫۳۲۵۷
		کاکس	(۱٫۰۱۲۰, ۰٫۰۱۴۲)	(۱٫۰۱۲۸, ۰٫۰۴۸۰)	۳۲۱٫۰۱۶۷	۳۲۸٫۹۹۷۱
		AFT	(۱٫۰۱۳۴, ۰٫۰۱۲۶)	(۱٫۰۱۴۰, ۰٫۰۴۶۳)	۳۲۰٫۵۳۶۹	۳۲۵٫۲۷۸۴
		شکنندگی	(۱٫۰۱۳۸, ۰٫۰۱۱۵)	(۱٫۰۱۵۳, ۰٫۰۴۳۵)	۳۱۹٫۳۵۴۴	۳۲۳٫۶۶۴۵
	۲۰۰	مدل بقای فضایی	(۱٫۰۱۷۵, ۰٫۰۰۹۰)	(۱٫۰۱۹۲, ۰٫۰۰۹۸)	۳۷۵٫۵۷۵۹	۳۸۱٫۲۹۸۹
		کاکس	(۱٫۰۱۰۰, ۰٫۰۱۳۶)	(۱٫۰۱۳۹, ۰٫۰۴۱۴)	۳۸۹٫۵۰۴۶	۳۹۳٫۹۹۸۰
		AFT	(۱٫۰۱۲۹, ۰٫۰۱۲۰)	(۱٫۰۱۵۷, ۰٫۰۳۷۹)	۳۸۸٫۲۱۲۰	۳۹۱٫۳۹۹۸
		شکنندگی	(۱٫۰۱۳۵, ۰٫۰۱۱۰)	(۱٫۰۱۶۷, ۰٫۰۳۴۲)	۳۸۷٫۶۶۸۲	۳۹۰٫۹۰۱۵
	۵۰	بقای فضایی	(۱٫۰۴۶۵, ۰٫۰۱۸۶)	(۱٫۰۴۰۱, ۰٫۰۲۳۰)	۱۰۴۷٫۳۱۲۰	۱۰۵۳٫۲۷۸۸
		کاکس	(۱٫۰۵۲۴, ۰٫۰۳۶۷)	(۱٫۰۵۱۴, ۰٫۰۸۵۱)	۱۰۷۸٫۶۹۹۹	۱۰۸۱٫۳۸۷۴
		AFT	(۱٫۰۵۱۷, ۰٫۰۳۶۴)	(۱٫۰۴۵۸, ۰٫۰۸۴۰)	۱۰۷۷٫۰۰۵۴	۱۰۷۹٫۷۴۱۲
		شکنندگی	(۱٫۰۵۰۹, ۰٫۰۳۵۵)	(۱٫۰۴۳۹, ۰٫۰۸۲۶)	۱۰۶۷٫۳۶۵۸	۱۰۷۱٫۶۹۸۷
%۸۰	۱۰۰	مدل بقای فضایی	(۱٫۰۴۳۰, ۰٫۰۱۳۵)	(۱٫۰۳۹۰, ۰٫۰۲۲۱)	۱۱۶۳٫۸۳۰۱	۱۱۷۱٫۸۷۹۵
		کاکس	(۱٫۰۴۸۰, ۰٫۰۳۶۲)	(۱٫۰۴۹۳, ۰٫۰۸۲۸)	۱۲۰۳٫۸۷۵۱	۱۲۰۶٫۹۵۱۴
		AFT	(۱٫۰۴۷۲, ۰٫۰۳۵۷)	(۱٫۰۴۲۷, ۰٫۰۸۰۳)	۱۲۰۲٫۴۹۲۵	۱۲۰۴٫۹۲۱۶
		شکنندگی	(۱٫۰۴۵۵, ۰٫۰۳۴۹)	(۱٫۰۴۰۰, ۰٫۰۷۷۹)	۱۱۹۷٫۴۲۹۸	۱۲۰۰٫۲۰۲۸
	۲۰۰	مدل بقای فضایی	(۱٫۰۴۱۳, ۰٫۰۱۲۳)	(۱٫۰۳۵۸, ۰٫۰۱۹۸)	۱۰۳۱٫۱۷۵۲	۱۰۴۰٫۵۲۱۸
		کاکس	(۱٫۰۴۵۳, ۰٫۰۳۵۴)	(۱٫۰۴۳۲, ۰٫۰۷۹۸)	۱۰۷۶٫۳۴۶۲	۱۰۸۱٫۹۸۲۱
		AFT	(۱٫۰۴۴۸, ۰٫۰۳۵۰)	(۱٫۰۴۲۰, ۰٫۰۷۴۵)	۱۰۷۳٫۹۸۸۸	۱۰۷۶٫۸۴۱۲
		شکنندگی	(۱٫۰۴۴۴, ۰٫۰۳۴۴)	(۱٫۰۳۹۲, ۰٫۰۷۲۴)	۱۰۷۲٫۸۲۳۴	۱۰۷۵٫۶۹۸۷

داده شده است، همانطور که در جدول مشاهده می‌شود طبق ملاک‌های DIC و BIC مدل بقای فضایی در همه حالات، عملکرد بهتری نسبت به سه مدل دیگر دارد. نتایج این جدول نشان می‌دهد که با افزایش سانسور مدل بقای از عملکرد بهتری نسبت به سایر مدل‌ها دارد. درجدول ۲ نتایج برازش مدل‌های بقا به داده‌های شبیه‌سازی شده در حجم ۴۰۰ و در سه تکرار ۵۰، ۱۰۰ و ۲۰۰ و سانسور راست ۲۰ و ۸۰ درصدی نشان داده شده است، همانطور که در جدول مشاهده می‌شود طبق ملاک‌های DIC و BIC در این حجم نمونه نیز مدل بقای فضایی در هر همه حالات عملکرد بهتری نسبت به سه مدل دیگر دارد. همچنین با افزایش حجم نمونه دقت نتایج در همه حالات از نتایج جدول ۱ بالاتر است. در جدول ۳ نتایج برازش مدل‌های بقا به داده‌های شبیه‌سازی شده در حجم ۶۰۰ و در سه تکرار ۵۰، ۱۰۰ و ۲۰۰ و سانسور راست ۲۰ و ۸۰ درصدی نشان داده شده است، همانطور که در جدول مشاهده می‌شود طبق ملاک‌های DIC و BIC در این حجم نمونه نیز مدل بقای فضایی در هر همه حالات عملکرد بهتری نسبت به سه مدل دیگر دارد. همچنین با افزایش حجم نمونه دقت نتایج در همه حالات از نتایج جدول ۱ و ۳ بالاتر است که این موضوع اهمیت حجم نمونه در مطالعات بقا به ویژه در حالاتی که درصد بالایی از سانسور رخ می‌دهد را تبیین می‌کند.

جدول ۲. نتایج برازش مدل‌های بقا در حجم نمونه ۴۰۰ و سه تکرار با سانسور ۲۰ و ۸۰ درصد

سانسور	تکرار	مدل	$(\beta_1, Std.)$	$(\beta_2, Std.)$	DIC	BIC
	۵۰	بقای فضایی	(۱/۰۱۷۹, ۰/۰۰۶۷)	(۱/۰۱۸۲, ۰/۰۰۹۲)	۸۳۹/۶۱۵۶	۸۴۵/۹۸۵۴
		کاکس	(۱/۰۱۱۸, ۰/۰۰۹۵)	(۱/۰۱۲۳, ۰/۰۳۸۶)	۸۴۴/۸۳۵۰	۸۵۰/۹۹۸۲
		AFT	(۱/۰۱۶۵, ۰/۰۰۹۰)	(۱/۰۱۵۱, ۰/۰۳۶۴)	۸۴۲/۴۹۵۵	۸۴۸/۳۲۶۵
		شکنندگی	(۱/۰۱۶۹, ۰/۰۰۸۲)	(۱/۰۱۷۰, ۰/۰۳۴۰)	۸۴۰/۶۳۶۸	۸۴۶/۹۹۸۲
%۲۰	۱۰۰	بقای فضایی	(۱/۰۱۸۹, ۰/۰۰۵۲)	(۱/۰۱۹۷, ۰/۰۰۸۴)	۷۷۱/۷۵۴۰	۷۷۷/۳۲۶۸
		کاکس	(۱/۰۰۹۰, ۰/۰۰۸۵)	(۱/۰۱۴۲, ۰/۰۳۱۵)	۷۷۹/۱۵۳۹	۷۸۶/۳۶۹۰
		AFT	(۱/۰۱۱۸, ۰/۰۰۸۳)	(۱/۰۱۷۴, ۰/۰۲۹۸)	۷۷۷/۷۵۳۹	۷۸۳/۳۶۹۸
		شکنندگی	(۱/۰۱۲۹, ۰/۰۰۷۷)	(۱/۰۱۸۸, ۰/۰۲۸۰)	۷۷۷/۲۱۲۷	۷۸۲/۲۰۳۰
	۲۰۰	بقای فضایی	(۱/۰۲۰۱, ۰/۰۰۳۴)	(۱/۰۲۰۹, ۰/۰۰۷۲)	۸۳۳/۱۴۹۲	۸۴۰/۹۸۵۲
		کاکس	(۱/۰۱۲۴, ۰/۰۰۸۰)	(۱/۰۱۶۵, ۰/۰۲۶۱)	۸۴۴/۹۱۷۹	۸۴۹/۳۹۸۸
		AFT	(۱/۰۱۲۸, ۰/۰۰۷۹)	(۱/۰۱۸۵, ۰/۰۲۴۳)	۸۴۴/۷۸۰۸	۸۴۸/۹۹۶۳
		شکنندگی	(۱/۰۱۳۸, ۰/۰۰۵۴)	(۱/۰۱۹۶, ۰/۰۲۳۵)	۸۴۲/۵۱۴۰	۸۴۶/۳۲۸۰
	۵۰	بقای فضایی	(۱/۰۳۹۹, ۰/۰۱۱۴)	(۱/۰۳۸۶, ۰/۰۱۷۲)	۱۹۹۰/۵۴۱۷	۱۹۹۶/۹۹۸۳
		کاکس	(۱/۰۱۸۹, ۰/۰۳۵۱)	(۱/۰۴۳۹, ۰/۰۷۱۹)	۲۱۰۵/۵۶۵۴	۲۱۱۱/۹۸۵۲
		AFT	(۱/۰۱۸۱, ۰/۰۳۴۲)	(۱/۰۴۲۶, ۰/۰۷۰۹)	۲۱۰۱/۹۲۷۱	۲۱۰۵/۲۰۳۶
		شکنندگی	(۱/۰۲۰۰, ۰/۰۳۳۸)	(۱/۰۴۱۷, ۰/۰۶۹۲)	۲۱۰۰/۹۲۳۸	۲۱۰۴/۳۲۰۵
%۸۰	۱۰۰	بقای فضایی	(۱/۰۳۸۲, ۰/۰۱۰۵)	(۱/۰۳۶۲, ۰/۰۱۵۹)	۲۱۶۷/۶۳۳۳	۲۱۷۱/۳۶۹۹
		کاکس	(۱/۰۲۱۴, ۰/۰۳۴۹)	(۱/۰۴۲۱, ۰/۰۶۷۶)	۲۲۲۷/۴۶۸۱	۲۲۳۲/۰۶۹۲
		AFT	(۱/۰۲۲۶, ۰/۰۳۳۸)	(۱/۰۴۱۸, ۰/۰۶۵۱)	۲۲۲۴/۴۳۷۴	۲۲۲۷/۱۴۷۸
		مدل شکنندگی	(۱/۰۲۴۴, ۰/۰۳۳۳)	(۱/۰۴۰۶, ۰/۰۶۳۲)	۲۲۱۵/۴۵۶۲	۲۲۲۱/۲۰۵۸
	۲۰۰	بقای فضایی	(۱/۰۳۷۷, ۰/۰۰۹۶)	(۱/۰۳۴۴, ۰/۰۱۴۰)	۲۱۴۸/۷۷۸۸	۲۱۵۴/۲۳۶۹
		کاکس	(۱/۰۳۹۸, ۰/۰۳۳۵)	(۱/۰۴۱۰, ۰/۰۶۲۲)	۲۲۶۹/۴۸۸۸	۲۲۷۵/۳۶۹۸
		AFT	(۱/۰۳۹۷, ۰/۰۳۳۰)	(۱/۰۴۰۲, ۰/۰۶۱۳)	۲۲۶۷/۸۴۱۱	۲۲۷۱/۳۰۲۰
		شکنندگی	(۱/۰۳۶۴, ۰/۰۳۲۳)	(۱/۰۳۹۴, ۰/۰۵۹۷)	۲۲۵۷/۰۶۹۸	۲۲۶۳/۴۱۰۵

لازم به ذکر مدل بقای فضایی به علت استفاده از اطلاعات مکانی مشاهدات در درصدهای بالای سانسور، عملکرد بهتری نسبت به سایر مدل‌ها دارد.

۵ مثال کاربردی

در این بخش به بررسی یک مثال کاربردی با استفاده از داده‌های بقای لوسمی حاد میلوئید[1] پرداخته می‌شود. لوسمی به معنی سرطان خون مربوط به بافت‌های خون‌ساز بدن است و سرطان خون حاد میلوئید یک نوع شایع و کشنده آن است.. داده‌های بقای لوسمی مورد استفاده در این مطالعه از داده‌های ثبت شده در پایگاه داده ثبت سرطان خون شمال غرب در بریتانیا بدست آمده است، که این پایگاه داده سوابق بروز و وضعیت بقای همه موارد لوسمی در شمال غربی انگلستان را در اختیار دارد. داده‌های بقای بیماران لوسمی براساس اطلاعات ثبت شده ۱۰۴۳ بیمار و سانسور

[1] Acute myeloid leukemia

جدول ۳. نتایج برازش مدل‌های بقا در حجم نمونه ۶۰۰ و سه تکرار با سانسور ۲۰ و ۸۰ درصد

سانسور	تکرار	مدل	$(\beta_1, Std.)$	$(\beta_2, Std.)$	DIC	BIC
	۵۰	بقای فضایی	(۱٫۰۲۱۳, ۰٫۰۰۲۶)	(۱٫۰۱۸۹, ۰٫۰۰۶۸)	۱۱۷۵٫۷۵۴۸	۱۱۸۱٫۳۶۹۴
		کاکس	(۱٫۰۱۴۴, ۰٫۰۰۸۴)	(۱٫۰۱۳۰, ۰٫۰۲۱۶)	۱۲۱۳٫۳۶۶۵	۱۲۱۸٫۳۰۲۹
		AFT	(۱٫۰۱۶۲, ۰٫۰۰۸۱)	(۱٫۰۱۶۲, ۰٫۰۲۰۹)	۱۲۱۲٫۴۲۴۵	۱۲۱۶٫۳۷۸۹
		شکنندگی	(۱٫۰۱۷۷, ۰٫۰۰۷۸)	(۱٫۰۱۷۷, ۰٫۰۱۹۲)	۱۲۱۱٫۶۲۷۴	۱۲۱۳٫۲۰۱۸
%۲۰	۱۰۰	بقای فضایی	(۱٫۰۳۷۲, ۰٫۰۰۲۰)	(۱٫۰۱۹۶, ۰٫۰۰۵۶)	۱۰۷۰٫۷۹۳۲	۱۰۷۶٫۵۲۶۸۹
		کاکس	(۱٫۰۳۹۹, ۰٫۰۰۷۴)	(۱٫۰۱۵۴, ۰٫۰۱۸۲)	۱۱۰۱٫۸۹۶۴	۱۱۰۹٫۷۴۵۱
		AFT	(۱٫۰۳۹۱, ۰٫۰۰۷۳)	(۱٫۰۱۷۱, ۰٫۰۱۷۵)	۱۱۰۰٫۴۳۶۶	۱۱۰۶٫۳۷۴۵
		شکنندگی	(۱٫۰۳۸۷, ۰٫۰۰۷۰)	(۱٫۰۱۸۲, ۰٫۰۱۶۹)	۱۰۹۷٫۳۵۲۷	۱۱۰۰٫۲۵۴۸
	۲۰۰	مدل بقای فضایی	(۱٫۰۳۱۰, ۰٫۰۰۱۶)	(۱٫۰۲۰۲, ۰٫۰۰۳۹)	۱۱۳۲٫۷۳۲۱	۱۱۳۹٫۶۵۲۱
		کاکس	(۱٫۰۲۶۵, ۰٫۰۰۷۰)	(۱٫۰۱۷۸, ۰٫۰۱۶۵)	۱۱۸۹٫۲۴۲۱	۱۱۹۵٫۵۲۱۴
		AFT	(۱٫۰۲۷۰, ۰٫۰۰۶۸)	(۱٫۰۱۸۴, ۰٫۰۱۵۹)	۱۱۸۸٫۸۹۶۲	۱۱۹۳٫۵۴۷۸
		شکنندگی	(۱٫۰۲۸۴, ۰٫۰۰۶۶)	(۱٫۰۱۹۴, ۰٫۰۱۵۰)	۱۱۸۷٫۹۰۸۴	۱۱۹۱٫۳۶۹۸
%۸۰	۵۰	بقای فضایی	(۱٫۰۲۰۱, ۰٫۰۰۹۲)	(۱٫۰۳۰۲, ۰٫۰۱۳۵)	۳۴۱۷٫۶۵۵۴	۳۴۲۱٫۶۵۴۷
		کاکس	(۱٫۰۱۸۶, ۰٫۰۳۱۲)	(۱٫۰۳۹۴, ۰٫۰۵۹۲)	۳۵۶۳٫۷۰۲۱	۳۵۷۱٫۴۱۷۸
		AFT	(۱٫۰۱۹۹, ۰٫۰۳۰۱)	(۱٫۰۳۸۲, ۰٫۰۵۸۵)	۳۵۵۵٫۴۱۷۷	۳۵۶۱٫۴۸۷۵
		شکنندگی	(۱٫۰۲۰۷, ۰٫۰۲۹۲)	(۱٫۰۳۶۴, ۰٫۰۵۷۷)	۳۵۴۴٫۱۰۵۴	۳۵۵۰٫۹۸۵۲
	۱۰۰	بقای فضایی	(۱٫۰۳۴۶, ۰٫۰۰۸۸)	(۱٫۰۲۷۶, ۰٫۰۱۲۸)	۳۲۸۸٫۱۳۶۴	۳۲۹۴٫۵۶۶۸
		کاکس	(۱٫۰۲۳۸, ۰٫۰۳۰۵)	(۱٫۰۳۷۰, ۰٫۰۵۷۱)	۳۴۷۵٫۸۹۱۲	۳۴۸۱٫۴۲۶۶
		AFT	(۱٫۰۲۵۰, ۰٫۰۲۹۳)	(۱٫۰۳۶۵, ۰٫۰۵۶۵)	۳۴۷۲٫۳۹۸۷	۳۴۷۹٫۶۵۴۴
		شکنندگی	(۱٫۰۲۵۹, ۰٫۰۲۸۵)	(۱٫۰۳۲۲, ۰٫۰۵۵۸)	۳۴۶۶٫۷۶۹۹	۳۴۷۱٫۶۵۷۷
	۲۰۰	مدل بقای فضایی	(۱٫۰۲۴۷, ۰٫۰۰۷۵)	(۱٫۰۲۴۲, ۰٫۰۱۲۰)	۳۱۹۳٫۸۹۵۱	۳۲۰۲٫۳۶۹۵
		کاکس	(۱٫۰۱۶۵, ۰٫۰۲۹۴)	(۱٫۰۳۳۸, ۰٫۰۵۵۳)	۳۲۸۹٫۵۱۲۳	۳۲۹۴٫۶۳۵۵
		AFT	(۱٫۰۱۹۷, ۰٫۰۲۸۳)	(۱٫۰۳۱۹, ۰٫۰۵۴۷)	۳۲۸۸٫۹۷۸۵	۳۲۹۱٫۵۴۷۷
		شکنندگی	(۱٫۰۱۹۲, ۰٫۰۲۷۶)	(۱٫۰۱۸۳, ۰٫۰۵۴۰)	۳۲۷۹٫۹۷۴۱	۳۲۸۳٫۵۴۷۱

۱۶ درصد طی سال‌های ۱۹۸۲ میلادی تا ۱۹۹۸ میلادی جمع‌آوری شده است. متغیرهای پیش‌بین در این داده‌های به صورت زیر است:

- سن: سن بیمار در زمان تشخیص بیماری سرطان خون که افرادی در رده سنی ۱۴ سال تا ۹۲ سال را در برمی‌گیرد.
- جنسیت: جنسیت بیماران مرد و زن را مشخص می‌کند که در آن عدد ۰ برای زنان و عدد ۱ برای مردان است..
- گلبول‌های سفید: تعداد و وضعیت گلبول‌های سفید در خون بیماران مبتلا به لوسمی حاد
- امتیاز تاونسند: به منظور تعیین سطح فقر خانواده و تحلیل عوامل مرتبط با فقر، از عواملی مانند درآمد، شرایط مسکن، آموزش، بهداشت و ارتباطات استفاده می‌کند، که مقادیر بالای آن نشان‌دهنده مناطق فقیرتری است این متغیر یک معیار کمی با دامنه‌ای از ۷− تا ۱۰ در داده‌های بقای لوسمی است..

در این داده‌ها موقعیت دقیق مکان زندگی بیماران و منطقه‌ای زندگی آنها نیز آورده شده است. در ادامه مدل‌های کاکس، شکنندگی و زمان شکست شتابیده نیز برای مقایسه به داده‌های بقای لوسمی برازانده شد، که نتایج برآورد این مدل‌ها در جدول ۴ آورده شده است. همانطور که ملاحظه می‌شود نتایج برازش مدل‌های بقا به داده‌های لوسمی حاد

جدول ۴. جدول نتایج داده‌های لوسمی حاد میلوئید در حجم ۱۰۴۳

مدل	$(\beta_1, Std.)$	$(\beta_2, Std.)$	$(\beta_3, Std.)$	$(\beta_4, Std.)$	DIC	BIC
بقای فضایی	(۱٫۰۳۰۷, ۰٫۰۰۰۴)	(۱٫۰۲۸۹, ۰٫۰۰۹۸)	(۱٫۰۰۹۲, ۰٫۰۰۰۱)	(۱٫۰۴۳۸, ۰٫۰۰۰۷)	۱۱۸۵۰	۱۱۸۵۴
کاکس	(۱٫۰۲۹۱, ۰٫۰۰۳۲)	(۱٫۰۳۶۰, ۰٫۱۱۲۷)	(۱٫۰۰۲۸, ۰٫۰۰۰۹)	(۱٫۰۲۷۲, ۰٫۰۲۶۳)	۱۱۹۵۵	۱۱۹۵۹
AFT	(۱٫۰۵۷۹, ۰٫۰۰۲۴)	(۱٫۰۸۸۰, ۰٫۰۶۰۷)	(۱٫۰۰۶۹, ۰٫۰۰۰۵)	(۱٫۰۷۴۶, ۰٫۰۱۶۴)	۱۱۹۰۲	۱۱۹۰۷
شکنندگی	(۱٫۰۳۱۹, ۰٫۰۰۲۱)	(۱٫۰۸۴۹, ۰٫۰۲۹۹)	(۱٫۰۰۳۱, ۰٫۰۰۰۴)	(۱٫۰۳۲۰, ۰٫۰۰۹۶)	۱۱۹۰۱	۱۱۹۰۵

میلوئید با حجم ۱۰۴۳ و سانسور ۱۶ درصد بدست آمده است، که مقادیر BIC برای مدل‌های بقای فضایی، کاکس، شکنندگی و زمان شکست شتابیده به ترتیب ۱۱۸۵۴, ۱۱۹۵۹, ۱۱۹۰۷, ۱۱۹۰۵ است.. طبق ملاک‌های DIC و BIC مدل بقای فضایی با توجه به مقدار کم DIC و BIC عملکرد بهتری نسبت به مدل‌های کاکس، شکنندگی و زمان شکست شتابیده را دارا است و مدل شکنندگی بعد از مدل بقای فضایی عملکرد بهتری نسبت به مدل‌های کاکس و زمان شکست شتابیده دارد. بر این اساس پارامتر رگرسیونی متغیر سن حاکی از آن است که با افزایش هر سال در متغیر سن خطر مرگ بیماران ۱٫۰۳۰۷ یعنی حدود ۰٫۰۳ درصد افزایش می‌یابد. پارامتر رگرسیونی متغیر جنسیت نشان می‌دهد که نرخ خطر مرگ در مردان ۱٫۰۲۸۹ یعنی حدود ۰٫۰۲ درصد بیشتر از زنان است. با افزایش یک واحد در تعداد گلبول‌های سفید خطر مرگ بیماران ۱٫۰۰۹۲ یعنی حدود ۰٫۰۰۹ درصد افزایش می‌یابد و با افزایش یک واحد در امتیاز تاونسند خطر مرگ در بیماران ۱٫۰۴۳۸ یعنی حدود ۰٫۰۴ افزایش می‌یابد.

بحث و نتیجه‌گیری

این مقاله به منظور بهبود کارایی مدل تحلیل بقا با در نظر گرفتن همبستگی فضایی بین مشاهدات به بررسی استفاده از یک مدل بقای فضایی در تحلیل داده‌های بقا پرداخته است. در حوزه تحلیل داده‌های بقا، مدل‌های کلاسیکی مانند مدل‌های کاکس، شکنندگی و زمان شکست شتابیده اغلب مورد توجه قرار می‌گیرند. اما وجود همبستگی فضایی بین مشاهدات ممکن است تأثیرات نادرست و گمراه کننده‌ای در نتایج این مدل‌ها ایجاد کند. به همین دلیل، استفاده از مدل بقای فضایی که در آن همبستگی فضایی بین مشاهدات نیز مدل‌سازی می‌شود، بهبود محسوسی در نتایج به همراه دارد. با این حال، در رویکرد فراوانی‌گرا، تحلیل مدل بقای فضایی به ویژه در حالت زمین‌آماری به دلیل پیچیدگی در توابع درستنمایی و وجود پارامترهای فضایی زمان‌بر و دشوار است. به همین دلیل، در این مقاله از رویکرد بیزی و استفاده از زنجیره مارکوف مونت کارلو در مدل بقای فضایی استفاده شده است. این روش، پیچیدگی را کاهش داده و نتایجی سریع‌تر و دقیق‌تر را فراهم می‌آورد. در نهایت مقایسه عملکرد مدل‌های کلاسیک و مدل بقای فضایی در تحلیل داده‌های بقا، نشان می‌دهد که استفاده از مدل بقای فضایی با رویکرد بیزی می‌تواند بهبود قابل توجهی در

تحلیل داده‌های بقا فراهم نماید.

تقدیر و تشکر

نویسندگان از سردبیر، داوران گرامی و ویراستار محترم مجله که نظرات و پیشنهادات ارزشمندشان همواره باعث اصلاح و بهبود مقالات ارسالی گردیده است تشکر و قدردانی می نمایند.

مراجع

محمدزاده، م. (۱۳۹۸)، آمار فضایی و کاربردهای آن، چاپ سوم، مرکز نشر آثار علمی دانشگاه تربیت مدرس، تهران.

مترجم، ک. (۱۴۰۰)، معرفی یک مدل بقای فضایی با اثرات تصادفی چوله گاوسی و کاربرد آن در تحلیل داده‌های بیماری کووید-۱۹، مجله علوم آماری، **۱۵**، ۵۶۷-۵۹۰

Banerjee, S., Wall, M. M. and Carlin, B. P. (2003), Frailty Modeling for Spatially Correlated Survival Data with Application to Infant Mortality in Minnesota, *Biostatistics*, **4**(1), 123-142.

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

Diggle, P. J., and J. A. Tawn, and R. A. Moyeed (1998), Model-based Geostatistics, *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, **47**, 299-350.

Gelfand, Alan E., and Sudipto Banerjee. (2017), Bayesian Modeling and Analysis of Geostatistical Data, *Annual Review of Statistics and Its Application*, **4**, 245-266.

Liu, Yulu. (2023), A Review of Survival Analysis Theory and Its Application. *Proceedings of the 2nd International Conference on Culture, Design and Social Development (CDSD 2022).*

Liopis-Cardona, Fran, Carmen Armero, and Gabriel Sanfélix-Gimeno. (2023), A Bayesian Multivariate Spatial Approach for Illness-death Survival Models.*Statistical Methods in Medical Research*, **32**, 1633-1648.

Miller, J. R. G. (2011), *Survival Analysis*, New York: Springer.

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

Rue, H. and Held, L. (2005), *Gaussian Markov Random Fields*, Chapman and Hall.

Tang, Y., Song, X., and Yi, G. Y. (2022), Bayesian Analysis under Accelerated Failure Time Models with Error-prone Time-to-event Outcomes. *Lifetime Data Analysis*, 1-30.

Thamrin, S. A., and Amran, J.aya, A. K., and Rahmi, S., and Ansariadi. (2017), Bayesian Inference for Spatial Parametric Proportional Hazards Model Using Spatsurv R, *In AIP Conference Proceeding*, **1827**(1), 200-215.

Thamrin, S. A., Jaya, A. K., and Mengersen, K. (2021), Bayesian Spatial Survival Modelling for Dengue Dever in Makassar, Indonesia. *Gaceta sanitaria*, **35**, 59-63.

Zhang, Z., Stringer, A., Brown, P. and Stafford, J. (2023), Bayesian Inference for Cox Proportional Hazard Models with Partial Likelihoods, Nonlinear Covariate Effects and Correlated Observations. *Statistical Methods in Medical Research*, **32**(1), 165-180.

Zhou, H. and Hanson, T. (2015), Bayesian Spatial Survival Models, *Nonparametric Bayesian Inference in Biostatistics*, 215-246.