

# عوامل مؤثر بر بقاء بیماران مبتلا به سرطان کولورکتال با استفاده از توابع مفصل

محمد رضا آخوند<sup>۱</sup>، نوشیروان کاظم نژاد<sup>۲</sup>، ابراهیم حاجی زاده<sup>۳</sup>، سید رضا فاطمی<sup>۴</sup>، علی قنبری مطلق<sup>۵</sup>

<sup>۱</sup> دانشجوی دوره دکتری دانشگاه تربیت مدرس، دانشکده پزشکی، گروه آمارزیستی، تهران، ایران

<sup>۲</sup> استاد دانشگاه تربیت مدرس، دانشکده پزشکی، گروه آمارزیستی، تهران، ایران

<sup>۳</sup> دانشیار دانشگاه تربیت مدرس، دانشکده پزشکی، گروه آمارزیستی، تهران، ایران

<sup>۴</sup> استادیار مرکز تحقیقات بیماری‌های گوارش و کبد، دانشگاه علوم پزشکی شهید بهشتی، تهران، ایران

<sup>۵</sup> استادیار مرکز تحقیقات سرطان، دانشگاه علوم پزشکی شهید بهشتی، تهران، ایران

نویسنده رابط: نوشیروان کاظم نژاد، نشانی: تهران، دانشگاه تربیت مدرس، دانشکده پزشکی، گروه آمارزیستی. تلفن: ۸۲۸۸۳۸۷۵، نمابر: ۸۲۸۸۴۵۵۵، پست الکترونیک:

kazem\_an@modares.ac.ir

تاریخ دریافت: ۱۳۸۸/۱۲/۱؛ پذیرش: ۱۳۸۹/۴/۱۲

**مقدمه و اهداف:** از جمله داده‌های بقاء چند متغیره، داده‌های ریسک‌های رقابتی هستند. یکی از روش‌های مدل‌سازی داده‌های ریسک‌های رقابتی استفاده از توابع مفصل است که هدف این مطالعه مدل‌سازی بیزی داده‌های ریسک‌های رقابتی به کمک توابع مفصل است.

**روش کار:** در این مطالعه از اطلاعات ثبت شده بیماران مبتلا به سرطان کولورکتال در مرکز تحقیقات گوارش و کبد دانشگاه علوم پزشکی شهید بهشتی استفاده شد. برای تحلیل داده‌ها، تابع درست‌نمایی به کمک تابع مفصل کلاسیک تشکیل گردید. در ادامه با انتخاب توزیع‌های پیشین مناسب برای پارامترها، توزیع پسین پارامترها با استفاده از الگوریتم متروپلیس-هستینگ و نمونه‌گیری برشی بدست آمد.

**نتایج:** در حالت تک متغیره بر زمان مرگ در اثر سرطان کولون و رکتوم عواملی همچون جنسیت، درجه تمایز یافتگی تومور، متاستاز به گره‌های لنفاوی، متاستاز به سایر ارگان‌ها و مرحله تومور دارای اثری معنی‌دار بودند. همچنین میزان نفوذ تومور به دیواره روده فقط بر بقاء بیماران مبتلا به سرطان کولون دارای اثری معنی‌دار بود. در حالت چندگانه نیز بر زمان مرگ در اثر سرطان کولون متغیرهای سن در زمان تشخیص، درجه تمایز یافتگی تومور و متاستاز به سایر ارگان‌ها دارای اثری معنی‌دار بودند، در حالیکه بر زمان مرگ در اثر سرطان رکتوم، متغیرهای درجه تمایز یافتگی و اندازه تومور اثر معنی‌داری داشتند.

**نتیجه‌گیری:** نتایج نشان داد که برخی از متغیرها ممکن است اثرات متفاوتی بر روی سرطان کولون و رکتوم داشته باشند که در نتیجه نیاز است تا تحقیقات بیشتری با متمایز در نظر گرفتن ریسک فاکتورهای مرتبط با سرطان کولون و رکتوم صورت گیرد.

**واژگان کلیدی:** تابع مفصل کلاسیک، ریسک‌های رقابتی، تحلیل بیزی، سرطان کولورکتال

## مقدمه

حادثه‌ای که مانع مشاهده حادثه مورد علاقه می‌شود و یا احتمال وقوع آن را تغییر می‌دهد، تعریف کرده‌اند.

مدل‌های مختلفی برای تحلیل داده‌های ریسک‌های رقابتی پیشنهاد شده است. از جمله مدل‌های مورد استفاده در تحلیل داده‌های ریسک‌های رقابتی، مدل مخاطره متناسب ویژه-علت کاکس<sup>۲</sup> است. در این روش برای تابع خطر ویژه-علت بعدی

در داده‌های بقاء پیشامد مورد نظر ممکن است به دلایل مختلفی روی دهد. برای مثال، اگر علاقه‌مند به تحلیل داده‌های زمان تا وقوع مرگ در اثر بیماری قلبی باشیم، برای برخی افراد تحت مطالعه ممکن است مرگ به دلایل دیگری غیر از بیماری قلبی، رخ دهد. این دلایل متفاوت شکست، به نام ریسک‌های رقابتی<sup>۱</sup> شناخته می‌شوند. گلی و همکاران (۱) ریسک رقابتی را به صورت

<sup>۲</sup> Cause Specific Cox Proportional Hazard Model

<sup>۱</sup> Competing Risks

لاولس (۱۴) نیز از مدل‌های تکه‌ای ثابت<sup>۴</sup> و توابع اسپلاین<sup>۵</sup> برای توابع خطر پایه در مدل‌های بقاء چند متغیره استفاده کردند. ژانگ و کلین (۱۵)، کریبر (۱۶) و کیاشف و همکاران (۱۷) از توابع مفصل برای تحلیل داده‌های ریسک‌های رقابتی استفاده کردند. هوانگ و ژانگ (۱۸) پس از مدل‌سازی حاشیه‌ها با استفاده از مدل‌های مخاطره متناسب از توابع مفصل برای مدل‌سازی توام داده‌ها استفاده کردند. چن (۱۹) نیز مدل‌های مخاطره متناسب و بخت‌های متناسب را برای حاشیه‌ها در تحلیل داده‌های ریسک‌های رقابتی وابسته به کار گرفت.

با وجود مزایای روش‌های بیزی نسبت به کلاسیک همچون نیرومندی در مقابل کم بودن اندازه نمونه و عدم نیاز به استفاده از توزیع‌های مجانبی برای برآوردها (۲۰) تاکنون روش‌های بیزی در تحلیل داده‌های مدل‌سازی شده با استفاده از توابع مفصل کمتر بکارگرفته شده‌اند و مقالات بسیار کمی در زمینه تحلیل بیزی داده‌های بقاء چند متغیره مدل‌سازی شده با استفاده از توابع مفصل می‌توان یافت. تنها مقالات منتشره در این زمینه مربوط می‌شود به گوستافون و همکاران (۲۱) که از نمونه‌گیری نقاط مهم<sup>۶</sup> برای برازش توابع مفصل به داده‌های بقاء دو متغیره استفاده کردند. آن‌ها ساختار وابستگی میان زمان‌های بقاء را از طریق تابع مفصل نرمال ایجاد کردند. رومئو و همکاران (۲۲) نیز روش‌های بیزی را در تحلیل داده‌های بقای دو متغیره مدل‌سازی شده با استفاده از توابع مفصل به کار گرفتند.

با توجه به مطالعات اندک در زمینه تحلیل بیزی داده‌های بقاء چند متغیره مدل‌سازی شده با استفاده از توابع مفصل و فقدان مطالعه‌ای در زمینه تحلیل بیزی داده‌های ریسک‌های رقابتی مدل‌سازی شده با استفاده از توابع مفصل، در این مطالعه قصد داریم پس از مدل‌سازی داده‌های ریسک‌های رقابتی به کمک تابع مفصل کلایتون، از روش‌های بیزی برای برآورد پارامترها استفاده کنیم. با توجه به اینکه محاسبه توزیع پسین پارامترها با استفاده از روش‌های تحلیلی امکان‌پذیر نیست، برای انجام تحلیل بیزی از روش شبیه‌سازی مؤنت کارلوی زنجیر مارکوفی<sup>۷</sup> استفاده می‌شود. این روش قادر است برای هر حجم نمونه دلخواه، استنباط نسبتاً دقیقی از پارامترها را ارائه دهد. در این مطالعه پس از معرفی توابع مفصل، چگونگی مدل‌سازی

نامتناهی در نظر گرفته می‌شود، در نتیجه اگر علاقه‌مند به برآورد احتمالات بقاء ویژه- علت<sup>۱</sup> باشیم، خطرات به دست آمده توابع پله‌ای غیر معقولی هستند که دارای نوار اطمینان بسیار عریضی- هستند (۲).

از دیگر روش‌های مورد استفاده در تحلیل داده‌های ریسک‌های رقابتی، مدل‌های کاملاً پارامتری است، که در این مدل‌ها در نظر گرفتن همبستگی میان زمان‌های بقاء باعث ایجاد مشکل شناسایی‌پذیری<sup>۲</sup> مدل می‌گردد. اما همانطوریکه می‌دانیم مدلی که دارای قابلیت شناسایی‌پذیری نباشد با داشتن تعداد متناهی مشاهده از آن نمی‌توان پارامترهای مدل را برآورد کرد و یا به عبارت دیگر مقادیر مختلف برای پارامترها می‌تواند منجر به توزیع‌های احتمالی یکسانی گردد. در نتیجه برآورد پارامترهای مدل امکان‌پذیر نیست. به دلیل وجود مشکل شناسایی‌پذیری در مدل‌های کاملاً پارامتری اغلب ساختاری مستقل برای ریسک‌های رقابتی وابسته در نظر گرفته می‌شود. به عبارت دیگر فرض می‌شود که زمان‌های شکست برای ریسک‌های مختلف دو به دو از هم مستقل هستند. بنابراین انتخاب مدل مستقل هم باعث سادگی و هم موجب رفع مشکل شناسایی‌پذیری مدل می‌گردد. اما در عمل، در بسیاری از موارد در نظر گرفتن چنین ساختاری می‌تواند منجر به نتایج گمراه کننده‌ای گردد (۳-۷).

کریبر (۸) و همچنین اسکلاو و کریبر (۲) مدلی را جهت ریسک‌های رقابتی وابسته بر مبنای توابع مفصل<sup>۳</sup> ارائه کردند که علاوه بر در نظر گرفتن همبستگی میان ریسک‌های مختلف، مشکل شناسایی‌پذیری مدل‌های کاملاً پارامتری را نداشت و در عین حال دامنه‌ای گسترده از همبستگی‌ها را از طریق پارامتر وابستگی میان زمان‌های بقاء فراهم می‌کرد. در این مدل توزیع‌های حاشیه‌ای به گونه‌ای کاملاً پارامتری انتخاب می‌شوند و می‌تواند تاثیر متغیرهای کمکی را بر روی ریسک‌های مختلف بدست آورد.

در سال‌های اخیر استفاده از تابع مفصل در زمینه‌های مختلف علمی گسترش فزاینده‌ای یافته است (۹-۱۱). آکز (۱۲) توابع مفصل را جهت مدل‌سازی همبستگی در مدل‌های شکنندگی دو متغیره به کار گرفت. شی و لوئیس (۱۳) روش‌های برآورد پارامتری و نیمه-پارامتری پارامتر همبستگی مدل‌های مفصل را در تحلیل داده‌های بقا دو متغیره مورد بررسی قرار دادند. هی و

<sup>۴</sup>Constant Model Picewise

<sup>۵</sup>Spline Function

<sup>۶</sup>Important Sampling

<sup>۷</sup>Markov Chain Monte Carlo

<sup>۱</sup>Cause Specific Survival Probabilities

<sup>۲</sup>Identifiability

<sup>۳</sup>Copula Functions

استفاده از رابطه زیر قابل محاسبه است :

(۳)

$$\tau_{\alpha}(T_1, T_2) = \frac{\alpha}{\alpha + 2}$$

شکل (۱) تابع چگالی توام وایبل به دست آمده بر مبنای مقادیر مختلف پارامتر همبستگی مفصل کلاپتون را نشان می‌دهد. شکل‌های زیر به ازاء مقادیر  $\alpha$  برابر صفر، ۲، ۴ و ۶ به دست آمده‌اند. همانطوریکه مشاهده می‌گردد با بزرگتر شدن مقدار پارامتر وابستگی  $\alpha$ ، همبستگی افزایش می‌یابد.

### تشکیل تابع درست‌نمایی داده‌های ریسک‌های رقابتی

داده‌های ریسک‌های رقابتی با  $m$  علت جهت شکست را در نظر بگیرید. همچنین فرض کنید  $T_{ij}$ ، زمان شکست به دلیل  $i$ -امین علت برای  $i$ -امین فرد در نمونه باشد که  $j = 1, 2, \dots, m$ ،  $i = 1, 2, \dots, n$  است. در این حالت یک مجموعه از داده‌ها به صورت  $\{z_{ij}, c_{ij}\}$  خواهیم داشت که در آن  $T_i = \min(T_{ij}, j = 1, 2, \dots, m)$  و  $z_i = \min(T_i, C_i)$  زمان‌های بقاء واقعی،  $1, 2, \dots, m$  زمان‌های سانسور شدن است. در اینجا ماتریس نشانگر وضعیت  $i$ -امین فرد برای  $i$ -امین علت به صورت  $c_{ij} = I(T_i = T_{ij})$  و بردار نشانگر وضعیت برای  $i$ -امین فرد به صورت  $c_i = \sum_{j=1}^m c_{ij}$  تعریف می‌گردد. همچنین فرض می‌شود که برای  $j \neq k$ ،  $T_{ij} \neq T_{ik}$ . در داده‌های ریسک‌های رقابتی با دو گروه از افراد مواجه هستیم، گروهی که حادثه برای آن‌ها در اثر یکی از ریسک‌ها بوقوع می‌پیوندد و گروه دیگر افرادی که تا پایان مطالعه حادثه‌ای برای آن‌ها رخ نمی‌دهد. برای فردی که حادثه برای وی براساس وقوع یکی از ریسک‌ها مثلا ریسک  $i$ -ام رخ می‌دهد، سهم وی در تابع درست‌نمایی به صورت زیر قابل محاسبه است:

(۴)

$$f^{(j)}(t) = \lim_{\Delta t \rightarrow 0} \frac{\Pr(T_1 > t, \dots, t \leq T_j \leq t + \Delta t, \dots, T_m > t | T \geq t)}{\Delta t} = -\frac{\partial}{\partial t_j} S(t_1, \dots, t_m) |_{t_k = t, \forall k}$$

و برای فردی که تا پایان مدت زمان مطالعه حادثه مورد نظر برای وی به وقوع نمی‌پیوندد، سهم وی در تابع درست‌نمایی برابر  $S_T(t) = \Pr(T_1 > t, \dots, T_m > t)$  می‌باشد. در نتیجه با فرض اینکه  $C_i$ ، متغیر سانسور شدن برای فرد  $i$ -ام، مستقل از زمان بقاء مربوط به  $i$ -امین ریسک  $T_{ij}$  باشد، می‌توان تابع درست‌نمایی را به صورت زیر نوشت (۲۸):

داده‌های بقاء ریسک‌های رقابتی با استفاده از تابع مفصل کلاپتون را ارائه می‌کنیم. سپس با انتخاب توزیع‌های پیشین مناسب برای پارامترها، از روش بیزی برای تحلیل داده‌های بقاء بیماران مبتلا به سرطان کولورکتال استفاده می‌کنیم.

## روش کار

### توابع مفصل

توابع مفصل به صورت توزیع‌های چندمتغیره‌ای که دارای توزیع‌های حاشیه‌ای یکنواخت می‌باشند تعریف می‌شوند. در حالت دومتغیره فرض کنید  $(T_1, T_2)$  متغیرهای تصادفی پیوسته با توابع بقاء حاشیه‌ای به ترتیب برابر با  $(S_1, S_2)$  باشند. در اینصورت تابع بقاء توام با استفاده از تابع مفصل  $C_{\alpha}$  به صورت زیر بدست می‌آید:

$$S(T_1, T_2) = C_{\alpha}(S_1(T_1), S_2(T_2))$$

خانواده‌های مختلفی از توابع مفصل وجود دارند که از جمله مهم‌ترین آن‌ها می‌توان به خانواده توابع مفصل ارشمیدسی اشاره کرد (۲۶-۲۳). در این خانواده از توابع مفصل تابع بقاء دو متغیره به صورت زیر تعریف می‌گردد:

(۱)

$S(t_1, t_2) = \varphi_{\alpha}(\varphi_{\alpha}^{-1}(S_1(t_1)) + \varphi_{\alpha}^{-1}(S_2(t_2)))$ ;  $\varphi_{\alpha}: [0, +\infty) \rightarrow [0, 1]$   
که  $\varphi_{\alpha}$  تابع مولد نامیده می‌شود،  $\varphi(0) = 0$ ،  $\varphi' < 0$ ،  $\varphi'' > 0$ ،  $1$  و  $\varphi_{\alpha}^{-1}$  معکوس تابع مولد است.  $\alpha$  هم نشان دهنده میزان همبستگی میان زمان‌های بقاء می‌باشد.

برای تحلیل داده‌ها نیاز به انتخاب یک تابع مفصل است، در این تحقیق با توجه به سادگی فرم محاسباتی، از تابع مفصل کلاپتون (۲۷) برای تشکیل تابع بقاء توام استفاده می‌شود. با استفاده از این تابع مفصل تابع بقاء توام به صورت زیر به دست می‌آید:

(۲)

$$S(t_1, t_2) = [S_1(t_1)^{-\alpha} + S_2(t_2)^{-\alpha} - 1]^{-\frac{1}{\alpha}} \quad \alpha \in \mathbb{R}^+$$

که در آن  $S_1(t_1)$ ،  $S_2(t_2)$  توابع بقاء حاشیه‌ای و  $\alpha$  پارامتر همبستگی است. هنگامیکه  $\alpha \rightarrow 0$  حاشیه‌ها مستقل خواهند بود یعنی  $S(t_1, t_2) = S(t_1) \times S(t_2)$ . با مشتق‌گیری از تابع بقاء توام بالا نسبت به  $t_1, t_2$  می‌توان تابع چگالی توام را به دست آورد. با استفاده از ضریب همبستگی تاوکندال، بهتر می‌توان میزان همبستگی میان متغیرها را درک کرد که در تابع مفصل کلاپتون با

(۵)

$$L = \prod_{i=1}^n \left( \prod_{j=1}^m [f^{(i)}(z_i)]^{c_{ij}} [S_T(z_i)]^{1-c_i} \right)$$

در حالتی که فقط دو دلیل علت شکست داشته باشیم (j = 2) با در نظر گرفتن توزیع وایبل برای حاشیه‌ها  $T_j \sim \text{weibull}(r_j, \lambda_j)$  با تابع بقاء به صورت زیر

(۶)

$S_j(t|r_j, \lambda_j) = e^{-\lambda_j t^{r_j}}$ ,  $r_j > 0, \lambda_j > 0, j = 1, 2$  و همچنین با وارد کردن متغیرهای کمکی از طریق  $\lambda_j = \exp(\beta_{0j} + \beta_{1j} X_{1j})$  تابع درستنمایی به صورت زیر به دست می‌آید:

(۷)

$$L(\alpha, r_1, r_2, \beta_1, \beta_2 | z, \Delta_1, \Delta_2, X) = \prod_{i=1}^n \left( r_1 z_i^{r_1-1} \right)^{\Delta_{i1}} \left( r_2 z_i^{r_2-1} \right)^{\Delta_{i2}} e^{\Delta_{i1}((\beta_{01} + \beta_{11} X_{11i}) + \alpha z_i^{r_1} \exp(\beta_{01} + \beta_{11} X_{11i}))} e^{\Delta_{i2}((\beta_{02} + \beta_{12} X_{12i}) + \alpha z_i^{r_2} \exp(\beta_{02} + \beta_{12} X_{12i}))} \left[ \left( e^{\alpha z_i^{r_1} \exp(\beta_{01} + \beta_{11} X_{11i})} \right) + \left( e^{\alpha z_i^{r_2} \exp(\beta_{02} + \beta_{12} X_{12i})} \right) - 1 \right]^{-\left( \frac{1}{\alpha} + \Delta_{i1} + \Delta_{i2} \right)}$$

که در آن  $\beta_j = (\beta_{0j}, \beta_{1j})^T$  است.

### تحلیل بیزی داده‌های ریسک‌های رقابتی

از مزیت‌های روش‌های بیزی نسبت به روش‌های کلاسیک این است که در صورتی که اطلاعات اضافی در مورد پارامترها در اختیار باشد می‌توان آن را از طریق توزیع پیشین وارد مدل نمود و در صورتی که چنین اطلاعاتی وجود نداشته باشد می‌توان از توزیع‌های پیشین ناآگاهی‌بخش که هیچگونه اطلاعات اضافی در مورد پارامترها را وارد مدل نمی‌کنند استفاده کرد.

در مطالعات بیزی دو نوع توزیع پیشین ناآگاهی‌بخش برای ضرائب رگرسیونی اغلب مورد استفاده قرار می‌گیرد: یکی توزیع پیشین ناسره<sup>۱</sup> به صورت  $\pi(\beta) \propto 1$  و دیگری توزیع پیشین سره<sup>۲</sup> ناآگاهی‌بخش نرمال که در آن با بزرگ در نظر گرفتن واریانس توزیع پیشین، توزیع نرمال تبدیل به توزیعی نسبتاً یکنواخت می‌گردد. با توجه به اینکه با انتخاب توزیع‌های پیشین ناسره ممکن است توزیع پسین ناسره گردد و استنباط درباره پارامترها را دچار مشکل کند اغلب توزیع پیشین ناآگاهی‌بخش نرمال برای

<sup>۱</sup>Improper

<sup>۲</sup>Proper

ضرائب رگرسیونی مورد استفاده قرار می‌گیرد. همچنین با توجه اینکه پارامترهای  $\alpha$  و  $r_j$  می‌توانند مقادیر غیرمنفی را اختیار کنند به ترتیب توزیع‌های پیشین ناآگاهی‌بخش گاما و نمایی برای آن‌ها انتخاب شد.

در نتیجه با در نظر گرفتن توزیع‌های پیشین‌های ناآگاهی‌بخش  $r_j \sim \text{exp}(r_{0j})$ ,  $\alpha \sim \text{gamma}(a, b)$  به صورت  $\beta_{0j} \sim \text{Normal}(c_{0j}, d_{0j}^2)$  و  $\beta_{1j} \sim \text{Normal}(c_{1j}, d_{1j}^2)$ ، توزیع پسین توام برای پارامترها  $(\alpha, r_1, r_2, \beta_1, \beta_2)$  به صورت زیر بدست می‌آید:

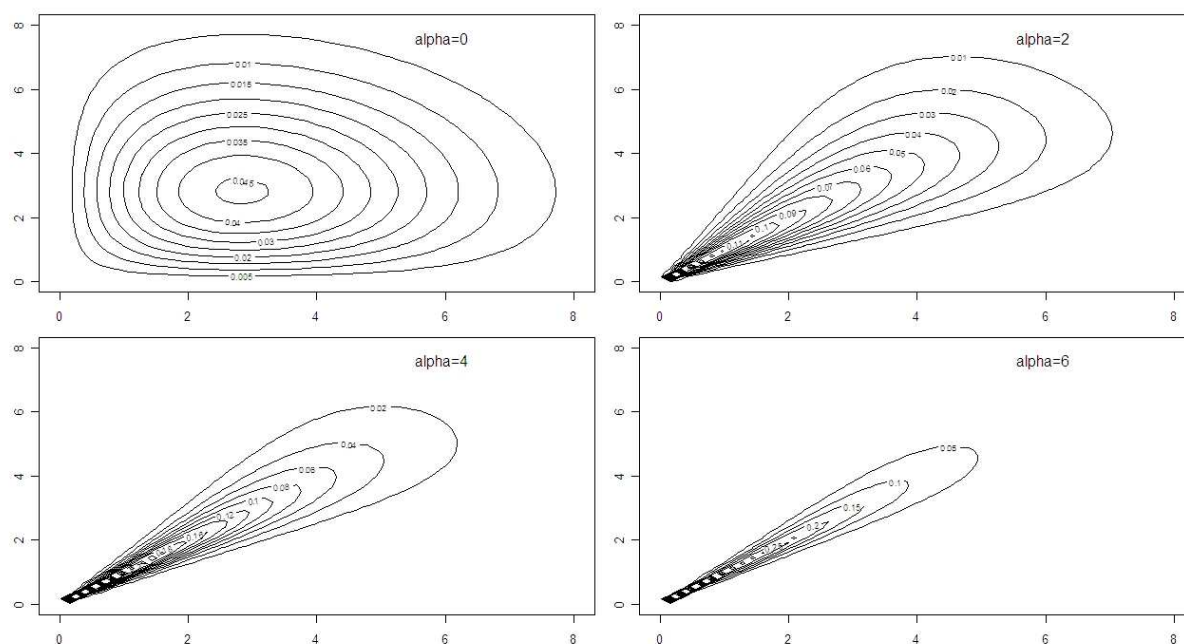
(۸)

$$\pi(\alpha, r_1, r_2, \beta_1, \beta_2 | z, \Delta_1, \Delta_2, X) \propto \prod_{i=1}^n \left[ \left( e^{\alpha z_i^{r_1} \exp(\beta_{01} + \beta_{11} X_{11i})} \right) + \left( e^{\alpha z_i^{r_2} \exp(\beta_{02} + \beta_{12} X_{12i})} \right) - 1 \right]^{-\left( \frac{1}{\alpha} + \Delta_{i1} + \Delta_{i2} \right)} e^{(\beta_{01} + \beta_{11} X_{11i}) \Delta_{i1} - (\beta_{01}^2 - 2\beta_{01} c_{01}) / 2d_{01}^2 - (\beta_{11}^2 - 2\beta_{11} c_{11}) / 2d_{11}^2} e^{(\beta_{02} + \beta_{12} X_{12i}) \Delta_{i2} - (\beta_{02}^2 - 2\beta_{02} c_{02}) / 2d_{02}^2 - (\beta_{12}^2 - 2\beta_{12} c_{12}) / 2d_{12}^2} \alpha^{a-1} e^{-b\alpha} e^{-r_{01} r_1 - r_{02} r_2} \left( r_1 z_i^{r_1-1} \right)^{\Delta_{i1}} \left( r_2 z_i^{r_2-1} \right)^{\Delta_{i2}} e^{\alpha(\Delta_{i1} z_i^{r_1} \exp(\beta_{01} + \beta_{11} X_{11i}) + \Delta_{i2} z_i^{r_2} \exp(\beta_{02} + \beta_{12} X_{12i}))}$$

به خاطر پیچیدگی و ابعاد زیاد توزیع توام پسین به دست آمده، امکان محاسبه توزیع پسین پارامترهای مدل به روش تحلیلی وجود ندارد. لذا برای تقریب توزیع پسین پارامترها از روش‌های مونت کارلوی زنجیره مارکوفی (MCMC) استفاده می‌گردد. در این روش نمونه‌های تصادفی از توزیع پسین تولید می‌شود و بر اساس این نمونه‌ها استنباط در مورد پارامترها صورت می‌گیرد. نمونه‌های ابتدایی که قبل از همگرایی زنجیره مارکوف تولید می‌شوند را دوره تطبیق می‌گویند که از نتایج آن در برآورد پارامترها استفاده نمی‌گردد (۳۰-۲۹).

### کاربرد: شناسایی عوامل مؤثر بر بقاء بیماران مبتلا به سرطان کولورکتال

سرطان کولورکتال چهارمین سرطان شایع در جهان است (۳۱). همچنین سومین دلیل مرگ بر اثر سرطان است (۳۲). سالیانه نزدیک به یک میلیون مورد جدید سرطان کولورکتال در سراسر جهان شناسایی می‌شود و نزدیک به نیم میلیون نفر به علت این بیماری می‌میرند (۳۳). بر اساس گزارش سالیانه اداره سرطان وابسته به مرکز مدیریت بیماری‌ها، سرطان کولورکتال به ترتیب



شکل ۱- توابع چگالی توام وایبل با پارامترهای شکل و مقیاس به ترتیب  $\lambda = 4, r = 2$ ، با استفاده از تابع مفصل کلایتون و مقادیر پارامتر همبستگی برابر  $\alpha = 0, 2, 4, 6$

برای این افراد مرگ می‌تواند بر اثر سرطان کولون و یا رکتوم بوقوع بپیوندد، می‌توان از مدل ریسک‌های رقابتی جهت تجزیه و تحلیل چنین داده‌ها استفاده نمود و اثر هر کدام از عوامل خطر را به صورت مجزا بر روی سرطان کولون و رکتوم مورد بررسی قرار داد. داده‌های استفاده شده در این تحقیق اطلاعات ۱۱۱۴ بیماران مبتلا به سرطان کولورکتال ثبت شده در مرکز ثبت تومور مرکز تحقیقات بیماری‌های گوارش و کبد وابسته به دانشگاه علوم پزشکی شهید بهشتی تهران در دوره زمانی ۱۳۸۰/۱۰/۱۱ تا ۱۳۸۵/۱۰/۱۱ می‌باشد. این افراد تا تاریخ ۱۳۸۵/۱۰/۱۱ پیگیری شده و وضعیت بقاء آن‌ها مشخص شده است. مرگ بیمار از طریق تماس با خانواده و اطرافیان بیمار مورد بررسی قرار گرفته است. علت فوت به صورت مرگ بر اثر سرطان کولون و مرگ بر اثر سرطان رکتوم در نظر گرفته شده و از مدل ریسک‌های رقابتی ارائه شده در بخش قبل جهت بررسی عوامل موثر بر بقاء بیماران استفاده گردید. عواملی که در این تحقیق اثرشان بر بقاء بیماران مبتلا به سرطان کولون و رکتوم مورد بررسی قرار گرفت عبارتند از: جنسیت، سن در زمان تشخیص، میزان نفوذ تومور به دیواره روده، متاستاز به گره‌های لنفاوی، متاستاز به سایر ارگان‌ها، درجه تمایز یافتگی تومور، نوع تومور، مرتبه تومور و اندازه تومور بود.

سومین و پنجمین سرطان شایع در میان زنان و مردان در ایران است (۳۴).

محققین معمولاً سرطان کولون و رکتوم را تحت یک نام و با عنوان سرطان کولورکتال گزارش می‌کنند، اما در سال‌های اخیر پژوهش‌های گوناگونی نشان داده‌اند که برخی متغیرها ممکن است اثرات متفاوتی بر روی سرطان کولون و رکتوم داشته باشند (۳۶-۳۵). در مطالعه‌ای که در سال ۲۰۰۸ توسط کمیته مشورتی راهنمای فعالیت فیزیکی صورت گرفت نتیجه‌گیری شد که فعالیت فیزیکی ارتباط معکوس معنی‌داری با سرطان کولون دارد اما ارتباط معنی‌داری میان مرگ در اثر سرطان رکتوم و فعالیت فیزیکی مشاهده نگردید (۳۷). همچنین در مطالعاتی نیز که توسط فلود و همکاران (۳۹-۳۸) و وی و همکاران (۴۰) صورت گرفته چنین نتیجه‌گیری شده که اثر متغیرهای تغذیه‌ای بر روی سرطان کولون و رکتوم متفاوت است. پدرس و همکاران (۴۱) نیز بیان کردند که میان مصرف الکل و سرطان رکتوم ارتباط معنی‌داری وجود دارد در حالیکه ارتباط معنی‌داری میان مصرف الکل و سرطان کولون مشاهده نگردید. فوچس و همکاران (۴۲) و وی و همکاران (۴۰) نیز نشان دادند که تأثیر سابقه خانودگی ابتلا به سرطان بر سرطان کولون و رکتوم متفاوت است. با توجه به اینکه

### روش اجرای شبیه‌سازی MCMC

### یافته‌ها

با به کارگیری توزیع پسین بدست آمده از عبارت (۸) در نرم افزار WinBugs (۴۳) و در نظر گرفتن توزیع پیشین ناآگاهی بخش برای پارامترها به صورت  $\alpha \sim \text{gamma}(1, 0.001)$  ،  $\beta_{kj} \sim \text{Normal}(0, 100^2)$  ،  $r_j \sim \text{exp}(0.001)$  ،  $k = 0, 1, j = 1, 2, \dots, K$  و استفاده از نمونه‌گیری برشی و الگوریتم متروپلیس-هستینگ می‌توان توزیع پسین برای پارامترها را به دست آورد. برای اطمینان از همگرایی زنجیره‌های مارکف از معیار همگرایی گوک (۴۴) استفاده شده است. پس از اجرای شبیه‌سازی در نرم‌افزار WinBugs نتایج حاصله برای بررسی همگرایی و انجام استنباط توسط بسته BOA (۴۵) در نرم افزار R (۴۶) مورد فراخوانی قرار گرفت. تمام شبیه‌سازی‌ها با استفاده از زنجیره‌هایی به طول ۴۰۰۰۰ و با در نظر گرفتن ۲۰۰۰۰ مشاهده اول به عنوان دوره تطبیق انجام گرفت.

تحلیل تک متغیره داده‌ها نشان داد که عواملی همچون جنسیت، درجه تمایز یافتگی تومور، میزان نفوذ تومور به دیواره روده، متاستاز به گره‌های لنفاوی، متاستاز به ارگان‌های دیگر و مرحله تومور، بر مرگ در اثر سرطان کولون دارای اثر معنی‌داری بودند. همچنین تحلیل تک متغیره برای سرطان رکتوم نشان داد که جز عامل میزان نفوذ تومور به دیواره روده بقیه عوامل با سرطان کولون مشترک است. بر اساس نتایج به دست آمده در جدول (۱) زنان نسبت به مردان دارای بقاء بیشتر و خطر کمتر مرگ در اثر سرطان کولون و رکتوم می‌باشند. همچنین درجه تمایز یافتگی تومور ارتباط معکوسی با خطر مرگ به دلیل سرطان کولون و رکتوم دارد. نسبت خطر بر اثر سرطان کولون در افرادی که نفوذ تومور به دیواره روده در آن‌ها مشاهده شده ۲/۳۹ برابر افرادی است که متاستاز به دیواره روده نداشته‌اند که این اثر از نظر آماری معنی‌دار بود. همچنین بیمارانی که متاستاز به گره‌های لنفاوی داشته‌اند نسبت به بیمارانی که متاستاز به گره‌های لنفاوی

جدول شماره ۱- نتایج حاصل از برازش مدل‌های ریسک‌های رقابتی به داده‌های بیماران مبتلا به سرطان کولورکتال در حالت تک متغیره

متغیر	سطوح	تعداد	درصد	سرطان کولون		سرطان رکتوم	
				H.R**	فاصله باورمند ۹۵٪	H.R	فاصله باورمند ۹۵٪
سن در زمان تشخیص	≤ ۵۰	۴۸۳	۴۳/۴	۱	-	۱	-
	> ۵۰	۶۳۱	۵۶/۶	۱/۰۸	(۰/۸۲ ، ۱/۳۸)	۱/۳۳	(۰/۹۵ ، ۱/۹۱)
جنسیت	مرد	۶۶۷	۵۹/۹	۱	-	۱	-
	زن	۴۴۷	۴۰/۱	۰/۷	(۰/۵۳ ، ۰/۹۲)*	۰/۵۸	(۰/۴ ، ۰/۸۳)*
درجه تمایز یافتگی تومور	خوب	۴۴۹	۴۰/۳	۱	-	۱	-
	متوسط	۲۹۲	۲۶/۲	۱/۱۹	(۰/۸۲ ، ۱/۶۵)	۱/۷	(۱/۱۱ ، ۲/۷)*
	ضعیف	۷۰	۶/۳	۲/۳۰	(۱/۴ ، ۳/۷۶)*	۱/۸۷	(۱/۱۲ ، ۳/۱۱)*
میزان نفوذ تومور به دیواره روده	عدم نفوذ	۱۱۳	۱۰/۱	۱	-	۱	-
	نفوذ کرده	۷۰۸	۶۳/۶	۲/۳۹	(۱/۳۷ ، ۴/۸)*	۱/۴۲	(۰/۷۷ ، ۲/۳۹)
متاستاز به گره‌های لنفاوی	مشاهده نشده	۴۱۹	۳۷/۶	۱	-	۱	-
	مشاهده شده	۳۶۱	۳۲/۴	۳/۳	(۲/۰۲ ، ۵/۴۴)*	۲/۲۴	(۱/۱۹ ، ۴/۰۶)*
متاستاز ارگان‌های دیگر	مشاهده نشده	۵۰۳	۴۵/۲	۱	-	۱	-
	مشاهده شده	۹۳	۸/۳	۴/۰۰	(۲/۷۵ ، ۵/۶۹)*	۲/۵۳	(۱/۴۳ ، ۴/۲۱)*
مرحله تومور	ابتدایی	۴۲۱	۳۷/۸	۱	-	۱	-
	پیشرفته	۴۱۱	۳۶/۹	۲/۴۷	(۱/۴۴ ، ۴/۰۳)*	۴/۴۷	(۱/۹۸ ، ۱۱/۹۳)*
اندازه تومور	≤ ۲۵	۱۱۳	۱۰/۱	۱	-	۱	-
	> ۲۵	۷۰۸	۶۳/۶	۱/۲۲	(۰/۷۹ ، ۱/۸۷)	۱/۴۶	(۰/۹۶ ، ۲/۲۱)
نوع مورفولوژی تومور	غیره	۲۱۱	۱۸/۹	۱	-	۱	-
	آدنوکارسینوما	۸۸۲	۷۹/۲	۰/۸۲	(۰/۶ ، ۱/۱۱)	۱/۱۸	(۰/۸۱ ، ۱/۸۶)

\* در سطح ۰/۰۵ معنی‌دار

\*\* نسبت خطر حاشیه‌ای (Marginal Hazard Ratio)

جدول شماره ۲- نتایج حاصل از برازش مدل‌های ریسک‌های رقابتی به داده‌های بیماران مبتلا به سرطان کولورکتال در حالت چندگانه

متغیر	ضریب (β)	انحراف معیار	فاصله باورمند <sup>†</sup> %۹۵	نسبت خطر (H.R)	فاصله باورمند %۹۵
مقدار ثابت	-۳/۸۳۳	۰/۶۵۳	(-۵/۱۶۹ ، -۲/۵۸)	-	---
سن در زمان تشخیص*	۰/۵۳۲	۰/۲۲۲	(۰/۰۹۹ ، ۰/۹۶۸)	۱/۷۰	(۱/۱۰ ، ۲/۶۳)
جنسیت	-۰/۳۱۴	۰/۲۲۶	(-۰/۷۶۳ ، ۰/۱۱۶)	۰/۷۳	(۰/۴۷ ، ۱/۱۲)
درجه تمایز متوسط	۰/۰۱	۰/۲۴۰	(-۰/۴۷۵ ، ۰/۴۷۳)	۱/۰۱	(۰/۶۲ ، ۱/۶۰)
یافتگی تومور* ضعیف*	۱/۱۰۹	۰/۳۵۲	(۰/۴۰۱ ، ۱/۷۹۳)	۳/۰۳	(۱/۴۹ ، ۶/۰۱)
میزان نفوذ تومور به دیواره روده	۰/۱۵۷	۰/۳۹۸	(-۰/۵۷۱ ، ۱/۰۰۷)	۱/۱۷	(۰/۵۶ ، ۲/۷۴)
متاستاز به گره‌های لنفاوی	۰/۷۴۹	۰/۴۸۸	(-۰/۲۰۲ ، ۱/۷۰۶)	۲/۱۱	(۰/۸۲ ، ۵/۵۱)
متاستاز به ارگان‌های دیگر*	۱/۴۴	۰/۳۱۶	(۰/۸۱۷ ، ۲/۰۵۸)	۴/۲۲	(۲/۲۶ ، ۷/۸۳)
مرحله تومور	-۰/۵۵۴	۰/۵۱۶	(-۱/۵۵۱ ، ۰/۴۶۳)	۰/۵۷	(۰/۲۱ ، ۱/۵۹)
اندازه تومور	۰/۶۳۹	۰/۴۴۸	(-۰/۱۸۴ ، ۱/۵۸۷)	۱/۸۹	(۰/۸۳ ، ۴/۸۹)
نوع مورفولوژی تومور	-۰/۴۱۳	۰/۳۰۴	(-۰/۹۹۶ ، ۰/۲۰۲)	۰/۶۶	(۰/۳۷ ، ۱/۲۲)
پارامتر شکل	۱/۱۳	۰/۰۹۶	(۰/۹۴۵ ، ۱/۳۲۳)	-	---
مقدار ثابت	-۷/۵۸۴	۱/۵۱۲	(-۱۱/۰۳ ، ۵/۰۱۷)	-	---
سن در زمان تشخیص	۰/۰۲	۰/۳۵۴	(-۰/۶۸۱ ، ۰/۷۱۱)	۱/۰۲	(۰/۵۱ ، ۲/۰۴)
جنسیت	۰/۰۹۶	۰/۳۵۲	(-۰/۶۰۳ ، ۰/۷۸۳)	۱/۱۰	(۰/۵۵ ، ۲/۱۹)
درجه تمایز متوسط*	۰/۷۵۸	۰/۳۵۴	(۰/۰۶۷ ، ۱/۴۵۳)	۲/۱۳	(۱/۰۷ ، ۴/۲۸)
یافتگی تومور* ضعیف*	۰/۸۵۹	۰/۳۸۵	(۰/۱۰۴ ، ۱/۶۳)	۲/۳۶	(۱/۱۱ ، ۵/۱)
میزان نفوذ تومور به دیواره روده	۰/۰۷۲	۰/۵۲۲	(-۰/۸۸۴ ، ۱/۱۷۵)	۱/۰۷	(۰/۴۱ ، ۳/۲۴)
متاستاز به گره‌های لنفاوی	۰/۸۳۳	۰/۹۳۶	(-۰/۸۹۱ ، ۲/۶۷۱)	۲/۳۰	(۰/۴۱ ، ۱۴/۴۵)
متاستاز به ارگان‌های دیگر	۰/۲۶۸	۰/۶۱۲	(-۱/۰۲۳ ، ۱/۳۸۱)	۱/۳۱	(۰/۳۶ ، ۳/۹۸)
مرحله تومور	۰/۱۶	۰/۹۶	(-۱/۷۴۷ ، ۱/۹۶)	۱/۱۷	(۰/۱۷ ، ۷/۱۰)
اندازه تومور*	۲/۲۹۳	۰/۶۱۱	(۱/۰۹۵ ، ۳/۴۹)	۹/۹	(۲/۹۹ ، ۳۲/۸)
نوع مورفولوژی تومور	۰/۲۷۸	۰/۵۹۱	(-۰/۷۸۲ ، ۱/۵۵۱)	۱/۳۲	(۰/۴۶ ، ۴/۷۲)
پارامتر شکل	۱/۳۹۲	۰/۱۷۵	(۱/۰۵۹ ، ۱/۷۴۸)	-	---
پارامتر همبستگی	۰/۶۸۴	۰/۵۵۵	(۰/۰۲۹ ، ۲/۰۷۱)	-	---

مرحله تمایز

مرحله تمایز

<sup>†</sup>Credible Interval

درجه تمایز یافتگی و اندازه تومور دارای اثری معنی‌دار بودند (جدول شماره ۲). خطر مرگ در اثر سرطان کولون در افراد با سن در زمان تشخیص بالای ۵۰ سال، ۱/۷ برابر خطر مرگ در این سرطان در افراد با سن در زمان تشخیص کمتر از ۵۰ سال است (( $HR = 1/7$  ;  $95\% CI = (1/1, 2/63)$ ). بیماران نیز که متاستاز به ارگان‌های دیگر در آن‌ها مشاهده بود دارای خطر بیشتر مرگ در اثر سرطان کولون بودند( $HR = 4/22$  ;  $95\% CI = (2/26, 7/83)$ ). اندازه تومور نیز به عنوان عاملی مؤثر بر مرگ در اثر سرطان رکتوم شناسایی شد به قسمی که خطر مرگ در اثر سرطان رکتوم در افراد با اندازه تومور بیشتر از ۲۵ میلی‌متر ۹/۹ برابر خطر مرگ در افراد با اندازه تومور کمتر از ۲۵ میلی‌متر بود

در آن‌ها مشاهده نشده نیز دارای خطر بیشتر مرگ بر اثر سرطان کولون و رکتوم می‌باشند. تحلیل مقدماتی همچنین نشان داد که افرادی که متاستاز به ارگان‌های دیگر داشته‌اند دارای بقای کمتری بوده‌اند. مرحله تومور نیز دارای ارتباط مستقیمی با خطر مرگ بر اثر هر دو سرطان مورد مطالعه داشت به نحوی که خطر مرگ در بیمار با مرحله پیشرفته تومور از خطر مرگ در بیمار با مرحله ابتدایی تومور بیشتر بود.

در تحلیل چندگانه داده‌ها با استفاده از تابع مفصل کلایتون، عوامل سن در زمان تشخیص، درجه تمایز یافتگی تومور، متاستاز به ارگان‌های دیگر دارای اثر معنی‌داری بر مرگ در اثر سرطان کولون بودند، در حالیکه بر مرگ در اثر سرطان رکتوم، متغیرهای

میان درجه تمایز یافتگی و سرطان رکتوم مشاهده شده است، اما ارتباط درجه تمایز یافتگی تومور با سرطان کولون معنی‌دار نبود. میزان نفوذ تومور به دیواره روده فقط در سرطان کولون به عنوان عاملی معنی‌دار بود که این مطلب در مطالعه لیانگ و همکاران (۴۹) نیز مشاهده می‌گردد. متاستاز به گره‌های لنفاوی نیز به عنوان عاملی مؤثر بر بقاء بیماران هم در سرطان کولون و هم در سرطان رکتوم مشاهده گردید. در تحقیقی نیز که توسط سو و همکاران (۵۰) صورت گرفته است نتیجه مشابهی به دست آمده است. پارک و همکاران (۴۸) نیز نشان دادند که متاستاز به گره‌های لنفاوی عاملی مؤثر در مرگ بر اثر سرطان کولون است. متاستاز به سایر ارگان‌ها نیز به عنوان عاملی مؤثر بر خطر مرگ بر اثر سرطان کولون و رکتوم بود که در تحقیق انجام شده توسط پارک و همکاران (۴۸) نیز متاستاز به سایر ارگان‌ها دارای اثر معنی‌دار بر بقاء بیماران مبتلا به سرطان کولون شناخته شد. در مطالعه‌ای نیز که توسط سو و همکاران (۵۰) انجام گرفته متاستاز به سایر ارگان‌ها هم بر مرگ بر اثر سرطان کولون و هم بر مرگ بر اثر سرطان رکتوم به عنوان عاملی معنی‌دار گزارش شده است. میان مرحله تومور نیز با سرطان کولون و رکتوم ارتباط معنی‌داری مشاهده گردید که در مطالعه لیانگ و همکاران (۴۹) و هریسون و همکاران (۵۱) نیز مرحله تومور به عنوان عاملی مؤثر بر بقاء بیماران مبتلا به سرطان کولون تشخیص داده شده است. سو و همکاران (۵۰) نیز مرحله بیماری را عاملی مؤثر بر بقاء بیماران مبتلا به سرطان کولون و رکتوم تشخیص دادند.

در تحلیل چند متغیره نیز سن در زمان تشخیص بیماری به عنوان عاملی مؤثر بر مرگ در اثر سرطان کولون شناسایی شد. در مطالعه‌ای که توسط پارک و همکاران (۴۸) صورت گرفته سن در زمان تشخیص به عنوان عاملی مؤثر در سرطان کولون و رکتوم بیان شده، در حالیکه مطالعه‌ای که توسط لیانگ و همکاران (۴۹) و چین و همکاران (۵۲) انجام شده است، سن در زمان تشخیص به عنوان عاملی مهم در مرگ بر اثر سرطان کولون گزارش شده است. متاستاز به سایر ارگان‌ها نیز دارای اثری معنی‌دار بر مرگ در اثر سرطان کولون بود که در تحقیق انجام شده توسط پارک و همکاران (۴۸) نیز متاستاز به سایر ارگانها دارای اثر معنی‌داری بر بقاء بیماران مبتلا به سرطان کولون شناخته شده است. در تحلیلی نیز که توسط سو و همکاران (۵۰) انجام گرفته متاستاز به سایر ارگان‌ها هم بر مرگ بر اثر سرطان کولون و هم بر مرگ بر اثر سرطان رکتوم به عنوان عاملی مؤثر بیان شده است. بر مرگ در اثر سرطان رکتوم، اندازه تومور دارای اثری معنی‌دار بود که در مطالعه

(HR = ۹/۹ ; ۰/۹۵ CI = (۲/۹۹ , ۳۲/۸)) درجه تمایز یافتگی تومور نیز ارتباط معکوس معنی‌داری با خطر مرگ در اثر سرطان‌های کولون و رکتوم داشت. در بیماران مبتلا به سرطان کولون خطر مرگ در افراد با درجه تمایز یافتگی ضعیف ۳/۰۳ برابر افراد با درجه تمایز یافتگی خوب بود (HR = ۱/۶,۴۹/۰۱) ; ۰/۹۵ CI = (۳/۰۳ ; HR = ۳/۰۳). همچنین در بیماران مبتلا به سرطان رکتوم خطر مرگ در افراد با درجه تمایز یافتگی تومور متوسط ۲/۱۳ برابر و در افراد با درجه تمایز یافتگی تومور ضعیف ۲/۳۶ برابر افراد با درجه تمایز یافتگی خوب بدست آمد (HR = ۲/۱۳ ; ۰/۹۵ CI = (۱/۶,۰۷/۲۸)) ; (HR = ۲/۳۶ ; ۰/۹۵ CI = (۱/۵,۱۱/۱)).

بر اساس برآورد مقدار پسین پارامتر همبستگی در جدول (۲) و با استفاده از رابطه (۳) مقدار ضریب همبستگی تاوکندال برابر ۰/۲۵۵ برآورد می‌گردد و یک فاصله اطمینان ۰/۹۵٪ برای آن به برابر (۰/۱۴، ۰/۵۰۹) بدست می‌آید که نشان از وجود همبستگی میان زمان مرگ در اثر سرطان کولون و رکتوم دارد. یکی از فرض‌های در نظر گرفته شده در تشکیل تابع درستنمایی، در نظر گرفتن توزیع وایبل برای حاشیه‌ها بود. که جهت بررسی این فرض از آزمون کلموگروف-اسمیرنوف در نرم افزار R و بسته MASS استفاده گردید (۴۷). مقدار -احتمال به دست آمده برای داده‌های سرطان کولون برابر ۰/۶۷۷ و برای داده‌های بقاء حاصل از سرطان رکتوم ۰/۳۲۹ بود که نشان از برازش مناسب توزیع وایبل به داده‌ها دارد.

## بحث

در تجزیه و تحلیل تک متغیره اطلاعات، جنسیت دارای رابطه‌ای معنی‌دار با مرگ بر اثر سرطان کولون و رکتوم بود و مردان نسبت به زنان دارای خطر بیشتر مرگ در اثر سرطان کولون و رکتوم بودند. اما در حالت چند متغیره ارتباط معنی‌داری میان جنسیت و بقاء بیماران در هر دو نوع سرطان مشاهده نگردید. مطالعه انجام شده توسط وی و همکاران (۴۰) نیز این نتیجه را تأیید می‌کند. پارک و همکاران (۴۸) ارتباط معنی‌داری را فقط میان جنسیت و مرگ بر اثر سرطان رکتوم گزارش کردند اما این ارتباط در مورد سرطان کولون معنی‌دار نبود. درجه تمایز یافتگی تومور نیز در روش تک متغیره و چندمتغیره اثری معنی‌دار بر روی سرطان کولون و رکتوم داشت. در تحقیقی نیز که توسط لیانگ و همکاران (۴۹) انجام شده است ارتباط معنی‌داری میان بقاء بیماران مبتلا به سرطان کولون و درجه تمایز یافتگی مشاهده است. در مطالعه صورت گرفته توسط پارک و همکاران (۴۸) نیز ارتباط معنی‌داری

### نتیجه گیری

در این تحقیق ما ابتدا چگونگی مدل سازی داده های ریسک های رقابتی را با استفاده از تابع مفصل کلایتون ارائه گردید. سپس روش های بیزی را تحلیل داده های ریسک های رقابتی مدل سازی شده با استفاده از تابع مفصل کلایتون به کار گرفته شد. این کار را می توان اولین گام در تحلیل بیزی داده های ریسک های رقابتی مدل سازی شده با استفاده از توابع مفصل در نظر گرفت، چرا که انعطاف پذیری توابع مفصل این امکان را فراهم می کند که از سایر توزیع های قابل استفاده در تحلیل داده های بقاء همچون گاما، لگ-نرمال، لگ-لجستیک و ... برای توزیع های حاشیه ای استفاده نمود و علاوه بر آن می توان توزیع های حاشیه ای را به صورتی غیرهمسان در نظر گرفت. همچنین برای انعطاف پذیری بیشتر مدل می توان حاشیه ها را به صورت نیمه پارامتری (۱۴) و یا کاملاً ناپارامتری (۵۳-۵۵) انتخاب کرد. با وجود خانواده های مختلفی از توابع مفصل برای انواع گوناگونی از همبستگی ها، می توان از سایر خانواده های توابع مفصل نیز جهت مدل سازی داده های ریسک های رقابتی استفاده کرد.

انجام شده توسط پارک و همکاران (۴۸) نیز ارتباط معنی داری میان خطر مرگ در اثر سرطان رکتوم و اندازه تومور مشاهده گردید اما این ارتباط در مورد سرطان کولون معنی دار نبود که منطبق با نتیجه بدست آمده در این مطالعه است.

همانطوریکه در نتایج ارائه شده در بالا ملاحظه می گردد برخی متغیرها دارای اثرات متفاوتی بر روی بقاء بیماران مبتلا به سرطان کولون و رکتوم هستند که در برخی از مطالعاتی که به آن ها اشاره شد بر این نکته تاکید شده است. لی و لای (۳۶) ضمن بررسی تفاوت های میان نواحی مختلف روده بزرگ پیشنهاد کرده اند که روده بزرگ به سه ناحیه تقسیم شده و ریسک فاکتورهای مربوط به هر بخش به صورت جداگانه مورد بررسی قرار گیرد.

در کلیه مطالعات صورت گرفته برای تجزیه و تحلیل داده ها به جای استفاده از روش ریسک های رقابتی از رگرسیون کاکس و با در نظر گرفتن مرگ در اثر علت دیگر به عنوان مشاهده ای سانسور شده استفاده شده است که همانطوریکه هوانگ و ژانگ (۱۸) و همچنین چن (۱۹) اشاره کرده اند در نظر گرفتن استقلال میان زمان های سانسور می تواند منجر به نتایج آریبی گردد.

### منابع

- Gooley TA, Crowley J, Storer BE, Leisenring W, et al. *Stat in Med.* 1999; 18: 695.
- Escarela G, Carriere JF. Fitting competing risks with an assumed copula. *Statistical Methods in Medical Research.* 2003; 12: 333-49.
- Bunea C, Bedford T. The effect of model uncertainty on maintenance optimization. *IEEE Transactions on Reliability.* 2002; 51: 486-93.
- Tsiatis A. A Nonidentifiability Aspect of the Problem of Competing Risks. *Proceedings of the National Academy of Sciences.* 1975; 72: 20-2.
- Cox DR. *Renewal Theory.* London 1962.
- Haile SR. Inference on competing risks in breast cancer data: University of Pittsburgh; 2008.
- Zheng M, Klein JP. Estimates of marginal survival for dependent competing risks based on assumed copula. *Biometrics.* 1995; 82: 127-38.
- Carriere JF. A select and ultimate parametric model. *Transaction of Society of Actuaries.* 1994b; 46: 75-97.
- Nelsen R. *An Introduction to Copulas.* Second Edition ed. New York: Springer; 2006.
- Joe H. *Multivariate Models and Dependence Concepts.* London: Chapman & Hall; 1997.
- Cherubini U, Luciano E, Vecchiato W. *Copula Methods in Finance.* John Wiley & Sons; 2004.
- Oakes D. Bivariate survival models induced by frailties. *Journal of the American Statistical Association.* 1989; 84: 487-93.
- Shih JH, Louis TA. Inference on the association parameter in copula methods for bivariate survival data. *Biometrics.* 1995; 51: 1384-9.
- He W, Lawless J. Flexible Maximum Likelihood Methods for Bivariate Proportional Hazards Models. *Biometrics.* 2003; 59: 837-48.
- Zheng M, Klein JP. Estimates of marginal survival for dependent competing risks based on assumed copula. *Biometrics.* 1995; 82: 127-38.
- Carriere JF. The effect of removing cancer as a cause of death when it is correlated with other causes. *Actuarial Research Clearing House.* 1994; 1: 153-63.
- Kaishev V, Dimitrova D, Haberman S. Modeling the joint distribution of competing risks survival times using copula functions. *Insurance: Mathematics and Econometrics.* 2007; 41: 339-61.
- Huang X, Zhang N. Regression Survival Analysis with an Assumed Copula for Dependent Censoring: A Sensitivity Analysis Approach. *Biometrics.* 2008; 64: 1090-9.
- Chen YH. Semiparametric marginal regression analysis for dependent competing risks under an assumed copula. *Journal of the Royal Statistical Society Series B.* 2010; 72: 235-51.
- Ibrahim JG, Chen MH, Sinha D. *Bayesian Survival Analysis.* New York: Springer 2001.
- Gustafson P, Aeschliman D, Levy A. A simple approach to fitting Bayesian survival models. *Lifetime Data Analysis.* 2003; 9: 5-19.
- Romeo JS, Tanaka NI, Pedrosa-Lima AC. Bivariate survival modeling: a Bayesian approach based on Copulas. *Life Time Data Analysis.* 2006; 12: 205-22.
- Genest C, Mackay J. The joy of copulas: Bivariate distributions with uniform marginals. *The American Statistician.* 1986; 40: 280-3.
- Hougaard P. Survival models for heterogeneous populations derived from stable distributions. *Biometrika.* 1986; 73: 387-96.
- Hougaard P. A class of multivariate failure time distributions. *Biometrika.* 1986; 73: 671-8.
- Frank M. on the simulation associativity of  $F(x,y)$  and  $x+y-F(x,y)$ . *Aequationes Mathematicae.* 1979; 19: 194-226.

- 27- Clayton DG. A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika*. 1978; 65: 141-51.
- 28- Eldant-Johnson R, Johnson N. *Survival models and data analysis*. New York: Wiley; 1980.
- 29- Gilks W, Richardson S, Spiegelhalter D. *Markov Chain Monte Carlo in practice* Chapman & Hall (1996). Chapman & Hall; 1996.
- 30- Neal R. Slice Sampling. *Annals of Statistics*. 2003;31:705-67.
- 31- Boyle P, Langman JS. ABC of colorectal cancer: Epidemiology. *BMJ*. 2000; 321: 805-8.
- 32- Parkin D. Global cancer statistics in the year 2000. *Lancet Oncology*. 2001; 2: 533-43.
- 33- Stone WL, Krishnan K, Campbell S, et al. Tocopherols and the treatment of colon cancer. *Annals of the New York Academy of Sciences*. 2004; 1031: 223-33.
- 34- Islamic Republic of Iran, Ministry of Health and Medical Education, Office of Deputy Minister for Health Center for disease control, cancer office. *Iranian Annual National Cancer Registration Report*. 2005-2006
- 35- Liang S, Carlin BP, Gelfand AE. Analysis of MINNESOTA colon and rectum cancer point patterns with spatial and nonspatial covariate information. *The Annals of Applied Statistics*. 2009; 3: 943-62.
- 36- Li FY, Lai MD. Colorectal cancer, one entity or three. *Journal of Zhejiang University SCIENCE B*. 2009; 10: 219-29.
- 37- Physical Activity Guidelines Advisory Committee Report. US Dept Health and Human Services, Washington, DC. 2008.
- 38- Flood A, Rastogi T, Wirf'alt E, Mitrou PN, Reedy J, Subar AF, et al. Dietary patterns as identified by factor analysis and colorectal cancer among middle-aged Americans. *American Journal of Clinical Nutrition*. 2008; 88: 176-84.
- 39- Flood A, Peters U, Chatterjee N, Lacey JVJ, Schairer C, et al. Calcium from diet and supplements is associated with reduced risk of colorectal cancer in a prospective cohort of women. *Cancer Epidemiology Biomarkers & Prevention*. 2005; 14: 126-32.
- 40- Wei EK, Giovannucci E, Wu K, Rosner B, Fuchs CS, Willett WC, et al. Comparison of risk factors for colon and rectal cancer. *International Journal of Cancer*. 2003; 108: 433-42.
- 41- Pedersen A, Johansen C, Gronbaek M. Relations between amount and type of alcohol and colon and rectal cancer in a Danish population based cohort study. *Gut*. 2003; 52: 861-7.
- 42- Fuchs C, Giovannucci E, Colditz G, Hunter D, Speizer F, Willett W. A prospective study of family history and the risk of colorectal cancer. *New England Journal of Medicine*. 1994; 331:1669-74.
- 43- Spiegelhalter D, Thomas A, Best N. *WinBugs 1.4*. Computer program. Imperial College & MRC Biostatistics unit, IPH Cambridge UK. 2003.
- 44- Geweke J, In: Bernarod J, Berger, JO, Dawid, AP, Smith, AFM. *Bayesian Statistics*. oxford: oxford University Press; 1992.
- 45- Smith B. <http://www.public-health.uiowa.edu/boa>. 2005.
- 46- Venables WN, Smith DM. <http://www.r-project.org/>. 2009.
- 47- Ricci V. Fitting distributions with R. <http://www.wfsforg/licenses/licenseshtml#FDL> 2005.
- 48- Park YJ, Park KJ, Park JG, Lee KU, Choe KJ, Kim JP. Prognostic Factors in 2230 Korean Colorectal Cancer Patients: Analysis of Consecutively Operated Cases. *World Journal of Surgery*. 1999; 23: 721-6.
- 49- Liang H, Wang XN, Wang BG, Pan Y, Liu N, Wang DC, et al. Prognostic factors of young patients with colon cancer after surgery. *World Journal of Gastroenterology* 7; 12:1458-62. 2006 12: 1458-62.
- 50- Xu F, Di M, Dong J, Wang F, Jin Y, Zhu Y, et al. Influence of clinical and pathomorphological parameters on prognosis in colon carcinoma and rectal carcinoma. *Zhejiang Da Xue Xue Bao Yi Xue Ban*. 2006; 35: 303-10.
- 51- Harrison LE, Guillem JG, Paty P, Cohen AM. Preoperative carcinoembryonic antigen predicts outcomes in node-negative colon cancer patients: a multivariate analysis of 572 patients. 1997; 185: 55-9.
- 52- Chin CC, Wang JY, Yeh CY, Kuo YH, Wen-Shih Huang WS, Yeh CH. Metastatic lymph node ratio is a more precise predictor of prognosis than number of lymph node metastases in stage III colon cancer. *International Journal of Colorectal Diseases*. 2009; 24: 1297-302.
- 53- Saliens-Torres VH, Pereira C, Tiwari R. Bayesian non-parametric estimation in a competing risks models or a series systems. *Journal of Nonparametrics Statistics*. 2002; 14: 449-58.
- 54- Hjort N. Nonparametrics Bayes estimators based on Beta processes in models for life history of data. *Annals of Statistics*. 1990; 18: 1259-94.
- 55- Susarla V, Van Ryzin J. Non-parametric Bayesian estimation of survival curves from incomplete observations. *Journal of the American Statistical Association*. 1976; 71: 897-902.

Iranian Journal of Epidemiology 2011; 6(4): 40-49.

**Original Article**

# Investigation of Influential Factors Affecting Survival Rate of Patients with Colorectal Cancer using Copula Function

Akhoond MA<sup>1</sup>, Kazemnejad A<sup>1</sup>, Hajizadeh E<sup>1</sup>, Fatemi SR<sup>2</sup>, Motlagh A<sup>3</sup>

1- Department of Biostatistics, Medical Science Faculty, Tarbiat Modares University, Tehran, Iran

2- Research Center for Gastroenterology and Liver Disease, Shahid Beheshti University of Medical Sciences, Tehran, Iran.

3- Cancer Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran

**Corresponding author:** Kazemnejad A, kazem\_an@modares.ac.ir

**Background & objectives:** Competing risk data is one of the multivariate survival data. Competing risk data can be modelled using copula function. In this study we propose a bayesian modelling approach of competing risk data using the copula function.

**Methods:** We used the data from colorectal cancer registry in Tehran. After constructing likelihood function using Clayton copula by choosing appropriate prior distribution for parameters, we obtained the posterior distribution of parameters using the Metropolis-Hastings algorithms and Slice sampling.

**Results:** The results of univariate analysis showed that sex, histology of tumor, extent of wall penetration, lymph node metastasis, distant metastasis and pathological stage of tumor were significantly associated with colon cancer and sex, histology of tumor, lymph node metastasis, distant metastasis and pathological stage of tumor were significantly related to rectal cancer. In the multivariate analysis, age at diagnosis, tumor grade and distant metastasis were significant prognostic factors for colon cancer and tumor grade and size of the tumor were significant prognostic factors of rectal cancer

**Conclusions:** As we showed some variables may have different impacts on colon and rectum cancers, consequently, further studies are needed to be conducted considering risk factors of these cancers separately.

**Keywords:** Clayton Copula, Competing risk data, Bayesian approach, Colorectal cancer