مجله دانشگاه علوم پزشکی خراسان شمالی
۱۳۹۰؛ ۳ (ویژه نامه آمار زیستی و اپیدمیولوژی):۶۴ - ۵۳

Journal of North Khorasan University of Medical Sciences
2011;3(Biostatistics and Epidemiology Supplement):53-64

تحلیل داده‌های بقای بیماران سرطان پستان، با مدل خطرهای متناسب کاکس برای سانسور وابسته بر اساس تابع مفصل

محمدتقی شاکری[۱]، وحید قوامی قنبرآبادی[۲]، حبیب اله اسماعیلی[۱]، هادی جباری نوقابی[۳]، فاطمه همایی شاندیز[۴]، حسین باغیشنی[۵]*

[۱] دانشیار آمار زیستی، گروه آمارزیستی و اپیدمیولوژی، دانشکده بهداشت، دانشگاه علوم پزشکی مشهد
[۲] کارشناس ارشد آمار زیستی، مدیریت آمار و فن آوری اطلاعات، دانشگاه علوم پزشکی مشهد
[۳] استادیار گروه آمار، دانشکده علوم ریاضی، دانشگاه فردوسی مشهد
[۴] دانشیار گروه انکولوژی و رادیوتراپی، بیمارستان امید، دانشگاه علوم پزشکی مشهد
[۵] استادیار گروه آمار، دنشکده علوم ریاضی، دانشگاه صنعتی شاهرود
***نویسنده مسئول:** دانشکده بهداشت، دانشگاه علوم پزشکی مشهد
پست الکترونیک: EsmailyH@mums.ac.ir

چکیده

زمینه و هدف: *در مدل رایج خطرهای متناسب کاکس یکی از فرضیات اساسی برقراری فرض استقلال میان زمان سانسور و زمان پیشامد می‌باشد. در مطالعات بالینی زمانی که سانسور به دلایلی همچون خروج فرد از مطالعه یا خطر رقابتی رخ می دهد همواره نگرانی در خصوص اعتبار نتایج مبتنی بر فرض استقلال سانسورها وجود دارد. ارائه راهکاری جهت بررسی فرض استقلال و نیز تعمیم مدل خطرهای متناسب کاکس در حالت عدم برقراری فرض استقلال مفید می‌باشد.*

مواد و روش کار: *برای رسیدن به این هدف به معرفی تعمیمی از مدل خطرهای متناسب کاکس با استفاده از تابع مفصل می‌پردازیم. این مدل امکان انجام آنالیز حساسیت، جهت بررسی فرض استقلال را فراهم می‌نماید. تعمیم تابع درستنمایی و برآورد پارامترها با یک روش عددی با استفاده از نرم‌افزار R انجام شد. در ادامه با استفاده از این مدل، برقراری فرض استقلال زمان سانسور و پیشامد در مجموعه داده-های مربوط به بیماران دارای سرطان پستان که طی سال‌های ۱۳۸۰ تا ۱۳۸۸جهت درمان به بخش آنکولوژی بیمارستان‌های قائم (عج) و امید مراجعه نموده بودند مورد بررسی قرار گرفت.*

یافته‌ها: *با استفاده از شبیه‌سازی ابتدا مناسب بودن مدل ارائه شده اثبات گردید سپس تاثیرات عدم برقراری فرض استقلال بین زمان سانسور و پیشامد بر روی برآورد پارامترها بررسی گردید. در مجموعه داده‌های واقعی برقراری فرض استقلال مورد تایید قرار گرفت.*

نتیجه گیری:*عدم برقراری فرض استقلال می‌تواند باعث نتایج اریب در مدل خطرهای متناسب کاکس گردد که میزان اریبی با میزان وابستگی میان سانسور و پیشامد ارتباط مستقیم دارد. در مجموعه داده‌های مربوط به بیماران دارای سرطان پستان مشخص گردید که تشخیص بیماری در مراحل اولیه نقش مهمی در موفقیت درمان و عدم عود بیماری‌دارد.*

واژه های کلیدی: *مدل خطرهای متناسب کاکس، آنالیز حساسیت، تابع مفصل، خطرات رقابتی، سرطان*

[Downloaded from journal.nkums.ac.ir on 2026-11-04]

مقدمه

درتحقیقات مختلف به‌خصوص مطالعات پزشکی گاهی هدف محقق، عوامل مؤثر بر زمان رخ دادن یک پیشامد[1] خاص می‌باشد؛ این پیشامد خاص می‌تواند مرگ بر اثر یک بیماری مشخص، عود یک بیماری، پاسخ به یک درمان ویژه، رد کردن عضو پیوندی و ... باشد. داده‌های مربوط به چنین مطالعاتی را داده‌های بقا (زمان تا پیشامد)[2] گویند. درحقیقت متغیر اصلی مورد توجه، زمان تا رخ دادن یک پدیده خاص می‌باشد که به آن زمان بقا[3] نیز گفته می‌شود. یکی از مشخصات داده‌های بقا وجود داده‌های سانسور شده[4] می‌باشد. به عبارتی زمان بقای واقعی ممکن است بیشتر از زمان مطالعه باشد و بنابراین نتوان آن را اندازه‌گیری کرد؛ به عنوان مثال در یک مطالعه ۱۰ ساله رتینوپاتی دیابتی، ممکن است برخی از بیماران در طی ۱۰ سال دچار رتینوپاتی نشوند و لذا زمان دقیق این پیشامد برای ایشان نامشخص است [۱]. در مطالعات مربوط به تحلیل بقا، یکی از روش‌های متداول برای پیش‌بینی احتمال بقا، استفاده از مدل خطرهای متناسب کاکس[5] می‌باشد؛ در این روش یکی از فرض‌های اساسی، استقلال میان زمان سانسور و زمان پیشامد مورد نظراست[۲]. در مطالعات بالینی زمانی که یک سانسور به دلایلی همچون کناره‌گیری بیمار[6] از مطالعه یا یک مخاطره رقابتی[7] رخ می دهد، همواره نگرانی در خصوص اعتبار نتایج مبتنی بر فرض استقلال سانسورها وجود دارد [۳]. به عنوان مثال ممکن است صعب العلاج بودن بیماری، ناامیدی از درمان، هزینه های درمان، شدت بیماری و غیره در تصمیم بیمار به ادامه درمان تأثیر گذاشته و از مطالعه کناره‌گیری نماید و از طرفی محقق با توجه به سوابق بیمار احتمال بیشتری را برای مشاهده پیشامد مورد نظر توسط وی قائل باشد. این نگرانی به خصوص در زمانی که خروج فرد از مطالعه و رخ ندادن رویداد مورد نظر تا پایان مطالعه برای فرد به عنوان سانسورهای مستقل[8] در نظر گرفته می‌شوند بروز می‌کند. یکی از راهکارهای بررسی وجود فرض استقلال انجام آنالیز حساسیت[9] است [۴]. آنالیز حساسیت، یک مطالعه در جهت بررسی نتایج یک روش آماری، در صورت تغییر مفروضات اساسی آن روش آماری است. به عبارت دیگر تغییرات در خروجی یک مدل ریاضی به چه اندازه می تواند به منابع ورودی آن مدل تخصیص یابد. حال اگر بتوان مدل خطرهای متناسب کاکس را به نحوی تعمیم داد که میزان وابستگی میان سانسور و پیشامد در مدل لحاظ گردد، آن گاه می‌توان با فرض کردن درجات خاصی از وابستگی نتایج حاصل بر روی برآورد پارامترها و معنی‌دار بودن اثر آن ها را بررسی نمود و به عبارتی یک آنالیز حساسیت انجام داد. یکی از روش‌های انجام آنالیز حساسیت استفاده از تابع مفصل[10] می‌باشد [۳]. زینگ و کلین[11] (۱۹۹۵) یک تابع مفصل را برای توزیع توأم مخاطرات رقابتی فرض نموده و توزیع‌های حاشیه‌ای آن را به روش ناپارامتری و تحت تابع مفصل مذکور برآورد نمودند. روش آن ها به دلیل امکان تغییر وابستگی میان توابع حاشیه‌ای در تابع مفصل و نوع تابع مفصل به کار رفته یک وسیله خوب آنالیز حساسیت برای برآوردگر کاپلان می‌باشد. آن ها نشان دادند زمانی که تابع مفصل مشخص باشد داده‌های مربوط به مخاطرات رقابتی جهت تعریف توابع حاشیه‌ای و ایجاد یک برآوردگر مناسب و سازگار کفایت می‌کند. زمانی که پیشامد مورد نظر و سانسور مستقل باشند این برآوردگر پیشنهادی به برآوردگر کاپلان مایر تبدیل می‌شود [۵].

زینگ و کلین (۱۹۹۴) در مطالعه‌ای دیگر بر اساس نتایج مطالعه قبل خود که در بالا به آن اشاره شد روش تابع مفصل را جهت ساختن یک برآوردگر برای توابع حاشیه‌ای بقا در داده‌های دارای مخاطرات رقابتی وابسته به‌کار بردند [۶]. قابل ذکر است که این مقاله با این که دیرتر از مقاله

[1] -event
[2]- time to event data
[3]- survival Time
[4] -cencored data
[5]- cox proportional hazards model
[6] -patient withdrawal
[7] -competing risk

[8]- independent censors
[9] -sensitivity analysis
[10] -copula
[11] -Zheng and Klein

[Downloaded from journal.nkums.ac.ir on 2026-11-04]

قبل که در بالا اشاره شد نوشته شده است اما زودتر به چاپ رسیده است. هانگ و چنگ[1] (۲۰۰۸) تابع مفصل ارایه شده توسط زینگ و کلین (۱۹۹۵) را جهت انجام یک آنالیز حساسیت برای مدل خطرهای متناسب کاکس تعمیم دادند. اهمیت کار ایشان از لحاظ تعمیم به مدل مشهور کاکس قابل توجه است [۳]. در این تحقیق ما با استفاده از نتایج مطالعات فوق و با استفاده از مدل تعمیم یافته ارائه شده توسط هانگ و چنگ (۲۰۰۸)، با استفاده از توابع مفصل به انجام آنالیز حساسیت می پردازیم. جهت پیاده‌سازی این روش و تعمیم تابع درستنمایی و برآورد پارامترها با یک روش عددی، از نرم‌افزار R استفاده نمودیم. در ادامه با استفاده از شبیه‌سازی دقت مدل ارائه شده را بررسی نموده همچنین تأثیر عدم برقراری فرض استقلال را در برآورد پارامترهای مدل نشان می‌دهیم. سپس مجموعه داده‌های مربوط به بیماران دارای سرطان پستان که طی سال‌های ۱۳۸۰ تا ۱۳۸۸ به بخش آنکولوژی بیمارستان‌های قائم (عج) و امید مشهد مراجعه نموده‌اند با استفاده از روش ارائه شده تحلیل می‌کنیم.

روش کار

با توجه به استفاده از تابع مفصل در ابتدا به معرفی تابع مفصل می‌پردازیم.

تابع مفصل: تابع مفصل، تابعی است که تابع توزیع چندمتغیره چند بعدی را به توابع حاشیه‌ای یک بعدی آن مرتبط می‌سازد. به طور کلی توابع مفصل، یک راه راحت جهت مدل‌بندی دو یا چند متغیر وابسته بدون نیاز به اطلاع از توزیع توأم یا حتی توزیع‌های حاشیه‌ای آن ها فراهم می‌نمایند. برای هر متغیر پیوسته، تابع توزیع از یک توزیع یکنواخت تبعیت می‌کند لذا توابع مفصل بر اساس متغیرهای دارای توزیع یکنواخت ایجاد می‌شوند [۷]. ما در این تحقیق از توابع مفصل دو بعدی استفاده می‌نماییم؛ فرض کنید که u و v در بازه $[0,1]$ باشند و α پارامتر تابع مفصل باشد که درجه وابستگی را نشان داده و در هر تابع مفصل ارتباطی بین آن و تاو کندال برقرار است. در این تحقیق با توجه به استفاده از تابع مفصل فرانک[2] به معرفی آن در ذیل می‌پردازیم:

تابع مفصل فرانک: تابع مفصل فرانک به صورت زیر تعریف می‌شود:

(۱)

$$C(u,v;\alpha)=\log_\alpha\{1+\frac{(\alpha^u-1)(\alpha^v-1)}{\alpha-1}\},\alpha>0,\alpha\neq 1.$$

تبدیلات میان τ کندال و پارامتر وابستگی α در تابع مفصل قابل محاسبه می‌باشد. ارتباط میان τ و α در تابع مفصل فرانک به صورت زیر است [۷]:

(۲)

$$\tau=1+4(-\log\alpha)^{-1}\{D_1(-\log\alpha)-1\},\alpha>0,\alpha\neq 1.$$

که در آن $D_k(X)$ (تابع دیبای[3]) به صورت زیر تعریف می‌شود که k یک عدد صحیح مثبت است:

(۳)

$$D_k(x)=\frac{k}{x^k}\int_0^x\frac{t^k}{e^t-1}dt.$$

مدل‌های با زمان پیشامد و سانسور آگاهی بخش: فرض می‌کنیم T زمان پیشامد، C زمان سانسور آگاهی بخش و S زمان سانسور بدون اطلاع باشند. زمان مشاهده شده یا به عبارتی زمان بقا را به صورت $X=\min(T,C,S)$ نشان می‌دهیم. همچنین دو تابع نشانگر $\delta_1=I(X=T)$ و $\delta_2=I(X=C)$ را تعریف می‌نماییم. دو متغیر مستقل Z و W با ابعاد به ترتیب $p\times n$ و $q\times n$ را که با زمان‌های سانسور آگاهی بخش و زمان پیشامد مرتبط هستند، در نظر می‌گیریم. توجه داریم که مشاهده T (یا C) سبب می‌شود تا C (یا T) در یک بازه کوتاه مشاهده نگردد، لذا می‌توان T و C را به عنوان مخاطرات رقابتی در نظر گرفت.

حال توابع خطر و توابع توزیع تجمعی حاشیه‌ای را به صورت زیر تعریف می‌کنیم:

- توابع خطر برای زمان پیشامد T_i و زمان سانسور وابسته C_i به ترتیب به صورت زیر تعریف می‌گردند:

$$\lambda_i(t|Z_i,W_i)=\lambda_0(t)\exp(Z_i'\beta) \quad \text{و} \quad \psi_i(t|Z_i,W_i)=\psi_0(t)\exp(W_i'\beta_c)$$

که در آن‌ها، $\lambda_o(t)$ و $\psi_o(t)$ توابع خطر پایه نامشخص، β و β_c پارامترهای مجهول به ترتیب با ابعاد $p\times 1$ و $q\times 1$ می‌باشند. این متغیرها می‌توانند

[1]- Huang and Zhang,
[2]- Frank copula
[3]- Debye function

[Downloaded from journal.nkums.ac.ir on 2026-11-04]

کاملا یکسان، در برخی یکسان ودر برخی مجزا و یا کاملا مجزا از یکدیگر باشند. توجه داریم که با این که تنها Z_i یـا W_i درسمت راست تساوی‌ها قرار دارند ولی هر دو Z_i و W_i در برآوردهای پارامترها نقش دارند، لذا دقیق‌تر این است که توابع خطر به همراه هر دو متغیر نوشته شود.

- توابع توزیع تجمعی حاشیه‌ای برای T_i و C_i به صورت زیر تعریف می‌گردند:

$$F_i(t\,|\,Z_i,W_i)=1-S_i(t\,|\,Z_i,W_i)=1-\exp\{-\Lambda_o(t)\exp(Z_i'\beta)\}$$

$$G_i(t\,|\,Z_i,W_i)=1-R_i(t\,|\,Z_i,W_i)=1-\exp\{-\Psi_o(t)\exp(W_i'\beta_c)\}$$

کـه $S_i(t)$ و $R_i(t)$ توابع بقا به ترتیب برای T و C و $\Lambda_o(t)$ و $\Psi_o(t)$ توابع خطرتجمعی پایه به ترتیب برایT و C می‌باشند. حال برای یک تابع مفصل مفروض مانند C با پارامتر α، تابع توزیع توأم تجمعی T_i و C_i به صورت زیر قابل مدل‌بندی است:

$$H_i(t,c)=\Pr(T_i\le t,C_i\le c)=C\{F_i(t),G_i(c);\alpha\}$$

به طورکلی تا زمانی‌که فرد پیشامد مورد نظر را تجربه نکرده باشد دارای جرم[1] واحد می‌باشد، اما به محض مشاهده پیشامد، تمام جرم خود را از دست می‌دهد. بدیهی است درحالت برقراری فرض استقلال میان سانسور و پیشامد، درصورتی که فرد سانسور گردد جرمش به صورت یکنواخت بر روی تمام نقاط پیشامد بعد از خود توزیع می‌گردد و این در حالی است که در حالت وجود سانسور وابسته این توزیع جرم به صورت یکنواخت انجام نمی‌شود. سؤالی که مطرح می‌باشد این است که در حالت سانسور وابسته جرم این افراد به چه صورت بر روی نقاط پیشامد بعد از خود توزیع می‌گردد؟

توابع احتمال شرطی بقا: فرض کنید x_i ، که $i=1,\ldots,n$ نقاط زمانی مرتب شده به صورت صعودی و بدون وجود گره[2] باشند. اگر فرد i ام در زمان x_i $i=1,\ldots,n$ سانسور شده باشد (سانسور آگاهی بخش) سپس برای زمان پیشامد $x_j > x_i$ می‌توانیم احتمال این که فرد i ام در زمان x_j ازدست رفته باشد را محاسبه نماییم. زینگ وکلین (۱۹۹۴) با انجام محاسباتی تحت فرض تابع توزیع توأم مشخص شده توسط تابع مفصل C برای $t > x_i$ نشان دادند که:

$$\Pr(T_i\ge t\,|\,T_i\ge x_i,C_i=x_i)=\frac{\Pr(T_i\ge t,C_i=x_i)}{\Pr(T_i\ge x_i,C_i=x_i)}=\frac{1-C_v\{F_i(t),G_i(x_i);\alpha\}}{1-C_v\{F_i(x_i),G_i(x_i);\alpha\}}$$

که درآن

$$C_v(a,b;\alpha)=\frac{\partial C(u,v;\alpha)}{\partial v}\Big|_{(u,v)=(a,b)}$$

حال تابع احتمال بقای شرطی[3] فوق را با $P_i(t)$ نشان می‌دهیم [۶].

جرمی که شخص سانسورشده iام در زمان x_j از دست می‌دهد را با $D_i(x_j)$ نمایش می‌دهیم که x_j زمان پیشامدفرد j ام می‌باشد و $x_j > x_i$. حال داریم:

$$D_i(x_j)=P_i(x_{j-1})-P_i(x_j) \qquad (۴)$$

برای سانسور آگاهی‌بخش نیز چون β_c در $P_i(t)$ دخیل است، لذا بایستی این احتمال مجدداً برآورد گردد. جهت درک بهتر، پیشامدها و سانسورها را در چارچوب مخاطرات رقابتی تعریف می‌نماییم؛ لذا می‌توانیم سانسور وابسته را به عنوان پیشامد درنظرگرفته و توابع مشابه را برآورد کنیم. به طور خاص برای فرد i ام که در زمان x_i پیشامد را مشاهده نموده است، می‌توانیم احتمال این که فرد در زمان $c > x_i$ به صورت وابسته سانسور گردد را محاسبه نماییم. برای $c > x_i$ داریم:

$$Q_i(c)=\Pr(C_i\ge c\,|\,C_i\ge x_i,T_i=x_i)=\frac{\Pr(C_i\ge c,T_i=x_i)}{\Pr(C_i\ge x_i,T_i=x_i)}=\frac{1-C_u\{F_i(x_i),G_i(c);\alpha\}}{1-C_u\{F_i(x_i),G_i(x_i);\alpha\}}$$

که درآن

$$C_u(a,b;\alpha)=\frac{\partial C(u,v;\alpha)}{\partial u}\Big|_{(u,v)=(a,b)}.$$

فرض کنیم که $E_i(x_j)$ بیان‌کننده جرمی است که شخص i ام که در زمان x_i پیشامد را مشاهده نموده است در زمان سانسور وابسته x_j از دست بدهد. داریم:

$$E_i(x_j)=Q_i(x_{j-1})-Q_i(x_j)$$

[1]- mass
[2] -tie
[3] -conditional survival probability function

[Downloaded from journal.nkums.ac.ir on 2026-11-04]

حال به معرفی توابع درستنمایی جزئی که تعمیمی از تابع درستنمایی جزئی مدل خطرهای متناسب کاکس است می‌پردازیم.

توابع درستنمایی جزئی: جهت درنظرگرفتن هم‌زمان پیشامد و سانسور آگاهی‌بخش، توابع نشان‌گر $D_i(x_j)$ و $E_i(x_j)$ را که بر خلاف توابع نشانگر معمول، هر مقدار بین صفر و یک را شامل می‌شود و بر اساس تابع مفصل تعریف شده‌اند، درنظر می‌گیریم. برای هر فرد با سانسور آگاهی‌بخش، فرض می‌کنیم که سهم آن در تابع درستنمایی به صورت تدریجی در هر نقطه‌ای که پیشامد مشاهده می‌شود، کاهش یافته و به وسیله $P_i(x_j)$ و $D_i(x_j)$ نشان داده می‌شود. به صورت شهودی این فرد نسبت به حالت معمول که فوراً پیشامد مورد نظر را مشاهده می‌نماید، به صورت تدریجی پیشامد را مشاهده می‌نماید.

حال تابع درستنمایی جزئی تعمیم‌یافته کاکس را به برای پیشامد به صورت زیر تعریف می‌کنیم:

$$L_j^{(T)}(\beta)=\prod_{i=1}^{j}\{\frac{P_i(x_j)\exp(Z_i^{'}\beta)}{\sum_{k=1}^{n}P_k(x_j)\exp(Z_k^{'}\beta)}\}^{D_i(x_j)}$$

$$L^{(T)}(\beta)=\prod_{j=1}^{n}L_j^{(T)}(\beta)=\prod_{j=1}^{n}\prod_{i=1}^{j}\{\frac{P_i(x_j)\exp(Z_i^{'}\beta)}{\sum_{k=1}^{n}P_k(x_j)\exp(Z_k^{'}\beta)}\}^{D_i(x_j)}$$

که در آن $L_j^{(T)}(\beta)$ تابع درستنمایی برای زمان x_j می‌باشد. حال جهت خوش‌تعریف شدن تساوی فوق بایستی برخی تعدیلات را اعمال نماییم. اولاً برای $j\le k$ مقدار $P_k(x_j)$ را برابر یک درنظر می‌گیریم؛ زیرا طبق تعریف $P_k(x_j)$ بیانگر احتمال زنده ماندن فرد k ام در زمان x_j به ازاء $j>k$ با فرض زنده ماندن تا زمان k است و این بدین معنا است که برای $j\le k$ فرد k ام زنده می‌ماند به عبارتی $P_k(x_j)=1$. ثانیاً توجه داریم که چون مشاهدات به صورت صعودی مرتب شده‌اند، لذا برای هر i که $i=1,...,n$ فرد i ام در زمان x_i پیشامد را مشاهده کرده است؛ بنابراین برای فرد i ام که پیشامد را مشاهده کرده است برای $j>i$ داریم $P_i(x_j)=0$. ثالثا با توجه به تعریف $D_i(x_j)$ برای فرد i ام که پیشامد مورد نظر را تجربه کرده است به جای استفاده از تساوی [۴] برای $j>i$ فرض می‌کنیم: $D_i(x_i)=1$ و $D_i(x_j)=0$ و این بدین معنی است که فرد i ام هنگامی که پیشامد را در زمان x_i تجربه کرده است تمام جرم[1] خود را از دست داده است.

در ادامه یک تابع درستنمایی جزئی تعمیم‌یافته کاکس را برای زمانی که سانسور وابسته را به عنوان پیشامد مورد نظر در نظر می‌گیریم، به صورت زیر تعریف می‌کنیم:

$$L_j^{(C)}(\beta)=\prod_{i=1}^{j}\{\frac{Q_i(x_j)\exp(W_i^{'}\beta_c)}{\sum_{k=1}^{n}Q_k(x_j)\exp(W_k^{'}\beta_c)}\}^{E_i(x_j)}$$

$$L^{(C)}(\beta_c)=\prod_{j=1}^{n}L_j^{(C)}(\beta_c)=\prod_{j=1}^{n}\prod_{i=1}^{j}\{\frac{Q_i(x_j)\exp(W_i^{'}\beta)}{\sum_{k=1}^{n}Q_k(x_j)\exp(W_k^{'}\beta)}\}^{E_i(x_j)}$$

که در آن $L_j^{(C)}(\beta_c)$ تابع درستنمایی برای زمان سانسور وابسته x_j می‌باشد.جهت خوش‌تعریف شدن تساوی فوق به صورت مشابه بایستی برخی تعدیلات را اعمال نماییم. اولاً برای $j\le k$ مقدار $Q_k(x_j)$ را برابر یک درنظر می‌گیریم؛ ثانیاً این که برای یک فردi ام که به صورت وابسته سانسور شده است برای $j>i$ ما $Q_i(x_j)=0$ فرض می‌نماییم. ثالثاً با توجه به تعریف $E_i(x_j)$ برای فرد i ام که به صورت وابسته سانسور شده است، فرض می‌کنیم: $E_i(x_i)=1$ و $E_i(x_j)=0$.

همچنین جهت سانسورهای مستقل نیز نیاز به تنظیمات دیگری نیز وجود دارد. برای فرد i ام که به صورت مستقل سانسور شده است مشابه دلایل فوق الذکر برای $j\le i$ داریم $P_i(x_j)=Q_i(x_j)=1$ و برای $j>i$ نیز داریم $P_i(x_j)=Q_i(x_j)=0$. همچنین در سانسورهای مستقل برای تمام jها فرض می‌کنیم $E_i(x_j)=D_i(x_j)=0$ و این بدین معنا است که فرد i ام هیچ مقدار از جرم خود را از دست نمی‌دهد. به صورت خلاصه در این روش با پیشامد مورد نظر و سانسورهای مستقل همانند مدل معمول کاکس برخورد می‌شود و تنها با سانسورهای آگاهی‌بخش به صورت متفاوتی برخورد شده است.

[1]- mass

[Downloaded from journal.nkums.ac.ir on 2026-11-04]

حال بایستی β و β_c از طریق ماکسیمم ساختن تابع درستنمایی جزئی توأم تعمیم یافته زیر برآورد گردند:

$$L(\beta,\beta_c)=L^{(T)}(\beta)L^{(C)}(\beta_c)$$

از آنجایی‌که توابع درستنمایی جزئی توأم تعمیم‌یافته‌ای که در قسمت قبل معرفی شدند، شامل مقادیر مجهولی همچون $P_i(x_j),Q_i(x_j),D_i(x_j),E_i(x_j)$ هستند؛ لذا جهت برآورد پارامترها از روش تکرارهای عددی استفاده می‌شود. روش تکرار عددی به شرح ذیل می باشد:

- مرحله اول: با فرض استقلال سانسورها و در نظر گرفتن دو مدل کاکس برای پیشامد مورد نظر و نیز سانسور وابسته مقادیر اولیه $\widehat{\beta}^{(0)},\widehat{\beta}_c^{(0)}$ را به دست آورده و سپس با استفاده از این مقادیر اولیه، با روش برسلو توابع خطر تجمعی پایه ($\Lambda_0^{(m)}$ و $\Psi_0^{(m)}$) را برآورد می‌نماییم. همچنین m را برابر صفر درنظر می‌گیریم.
- مرحله دوم:

$$\widehat{\beta}^{(0)},\widehat{\beta}_c^{(0)}\Rightarrow\widehat{S}_0^{(0)}(t),\widehat{R}_0^{(0)}(t)\Rightarrow\widehat{F}_i^{(0)}(.),\widehat{G}_i^{(0)}(.)\Rightarrow$$
$$\Rightarrow\widehat{P}_i^{(0)}(.),\widehat{Q}_i^{(0)}(.)\Rightarrow\widehat{D}_i^{(0)}(.),\widehat{E}_i^{(0)}(.).$$

سپس $L(\beta,\beta_c)$ را نسبت به پارامترهای β,β_c ماکسیم کرده و مقادیر آن را با $\widehat{\beta}^{(1)},\widehat{\beta}_c^{(1)}$ نشان داده و m را برابر یک در نظر می‌گیریم.

- مرحله سوم: با استفاده از برآوردهای جدید پارامترها خواهیم داشت:

$$\widehat{\beta}^{(m)},\widehat{\beta}_c^{(m)}\Rightarrow\widehat{S}_0^{(m)}(t),\widehat{R}_0^{(m)}(t)\Rightarrow\widehat{F}_i^{(m)}(.),\widehat{G}_i^{(m)}(.)\Rightarrow$$
$$\Rightarrow\widehat{P}_i^{(m)}(.),\widehat{Q}_i^{(m)}(.)\Rightarrow\widehat{D}_i^{(m)}(.),\widehat{E}_i^{(m)}(.).$$

- مرحله چهارم: دوباره $L(\beta,\beta_c)$ را نسبت به پارامترهای β,β_c ماکسیم کرده و مقادیر آن را با $\widehat{\beta}^{(m+1)},\widehat{\beta}_c^{(m+1)}$ نشان می‌دهیم.
- مرحله پنجم: m=m+1 و به مرحله سوم برگشته، تا زمانی که همگرایی به وجود آید، این عمل را ادامه می‌دهیم.

در پایان می‌توانیم توابع خطر تجمعی $\widehat{\Psi}(.),\widehat{\Lambda}(.)$ و توابع بقای $\widehat{R}(.),\widehat{S}(.)$ را برآورد نماییم.

قابل ذکر است که جهت پیاده‌سازی الگوریتم فوق و نیز شبیه‌سازی که در ادامه خواهد آمد از نرم افزار R.2.12.1 استفاده شده است.

یافته‌ها

در این بخش ابتدا با استفاده از شبیه‌سازی به بررسی دقت مدل پرداخته و در ادامه مدل تعمیم یافته ارائه شده در این مقاله را جهت بررسی فرض استقلال زمان سانسور و زمان پیشامد بر روی یک مجموعه داده واقعی مربوط به بیماران دارای سرطان پستان که جهت درمان به بخش آنکولوژی بیمارستان‌های قائم(عج) و امید مشهد مراجعه نموده‌اند استفاده می‌کنیم.

شبیه‌سازی: در این بخش با استفاده از شبیه‌سازی[1] به دنبال بررسی دقت مدل تعمیم‌یافته در برآورد پارامترها و نیز بررسی تأثیر استفاده از وابستگی‌های گوناگون در برآورد پارامترها می‌باشیم. در ابتدا تعداد ۳۰۰ مجموعه داده بقا را که شرط استقلال زمان پیشامد و زمان سانسور در آن ها برقرار نبوده، به حجم ۲۰۰ بر اساس معادلات مشخص بقا با پارامترهای معلوم تولید می‌نماییم. جهت این امر از تابع مفصل فرانک با مقدار پارامتر معلوم ۰.۰۰۰۰۰۰۰۱۲۵۸ که معادل τ کندال ۰.۸ می‌باشد، استفاده شده است. این مجموعه داده‌ها را با استفاده از روش‌های ذیل مدل‌بندی می‌نماییم:

- ✓ مدل خطرهای متناسب کاکس
- ✓ مدل معرفی شده در این تحقیق با تابع مفصل فرانک و میزان وابستگی معادل τ ی کندال ۰/۰۸، (α=۰/۰۰۰۰۰۰۰۱۲۵۸)
- ✓ مدل معرفی شده در این تحقیق با تابع مفصل فرانک و میزان وابستگی معادل τ ی کندال ۰/۰۵، (α=۰/۰۰۳۲۱۵)
- ✓ مدل معرفی شده در این تحقیق با تابع مفصل فرانک و میزان وابستگی معادل τ ی کندال ۰/۰۲، (α=۰/۱۵۵۵۴)

قابل ذکر است که ارتباط میان τ ی کندال و پارامتر تابع مفصل فرانک با استفاده از نرم‌افزار Maple و براساس روابط ۱ و ۲ محاسبه گردید.

[1]-simulation

[Downloaded from journal.nkums.ac.ir on 2026-11-04]

نحوه تولید متغیرهای وابسته زمان پیشامد و سانسور آگاهی بخش: جهت تولید داده‌ها دو متغیر مستقل یکی از توزیع برنولی با احتمال ۰.۵ و دیگری از توزیع یکنواخت در بازه (۵،۵-) ایجاد می‌نماییم. سپس با استفاده از توزیع وایبل، توزیع‌های حاشیه‌ای مربوط به زمان‌های پیشامد و سانسور آگاهی بخش را به صورت زیر به دست می‌آوریم؛ برای زمانtداریم:

$$S(t)=\exp\{-0.5t^{2}\exp(\beta_1 B+\beta_2 A)\} \qquad R(t)=\exp\{-0.5t\exp(\beta_{1c} B+\beta_{2c} A)\}$$

که درآن $\beta_1=0.5,\beta_2=0.1,\beta_{1c}=0.6,\beta_{2c}=0.3$

.

تابع توزیع توأم T وC را با استفاده از تابع مفصل فرانک به صورت زیر مشخص کرده‌ایم:

$$J(t,c;\alpha)=\log_\alpha\{1+\frac{(\alpha^{F(t)}-1)(\alpha^{G(c)}-1)}{\alpha-1}\}$$

جهت تولید زمان های پیشامد و سانسور آگاهی بخش که از یکدیگر مستقل نمی‌باشند، با استفاده از تابع مفصل فرانک به صورت زیر عمل می‌کنیم [۸]:

- مرحله اول: دو متغیر تصادفی مستقل u و v' از توزیع یکنواخت (۱و۰) تولید می‌نماییم.
- مرحله دوم: قرار می دهیم:

$$v=\log_\alpha\{1+\frac{v'(\alpha-1)}{v'+(1-v')\alpha^{u}}\}$$

سپس توابع توزیع تجمعی u و v' را که در رابطه‌ی مربوط به تابع مفصل فرانک صدق می‌کند و متعاقب آن $S(t)$و $R(t)$را به دست آورده و با قرار دادن در روابط (۲۹.۳) و با فرض معلوم بودن α ، t_T و t_C به صورت زیر حاصل می‌شوند:

$$t_C=\sqrt{\frac{-2\log(S(t))}{\exp(Z_i'\beta)}} \qquad t_T=\frac{-2\log(R(t))}{\exp(W_i'\beta_c)}$$

علاوه بر این متغیر تصادفی S_i از توزیع یکنواخت (۰، ۵) را جهت زمان سانسور مستقل تولید می‌نماییم. بنابراین با توجه به تعریف سانسور راست مشاهدات به صورت $X_i=\min(T_i,C_i,S_i)$ تعریف می‌گردند. پس از قرار دادن تمامی متغیرها در یک مجموعه داده‌ها و مرتب کردن آن ها بر اساس متغیر زمان X_i به صورت صعودی، روش ارایه شده در فصل قبل را به همراه مدل متداول خطرهای متناسب کاکس بر روی این مجموعه داده پیاده‌سازی نموده تا برآوردهای مربوط به پارامترهای β و β_c را به دست آوریم. قابل ذکر است که جهت تولید این مجموعه داده به حجم ۲۰۰ از مقدار۰.۰۰۰۰۰۰۰۱۲۵۸=α که معادل τ کندال ۰.۸ می‌باشد استفاده نموده و به همین ترتیب۳۰۰مجموعه داده تولید می‌نماییم. در ابتدا روش متداول مدل خطرهای متناسب کاکس را جهت برآورد ضرائب پارامترها در ۳۰۰ مجموعه داده‌ی شبیه‌سازی شده استفاده می‌نماییم. سپس میانگین برآوردهای حاصل به همراه میانگین مربعات خطا (MSE) را محاسبه می‌کنیم. هدف از این قسمت تأثیر عدم برقراری فرض استقلال در برآوردپارامترها می‌باشد. همان‌طور که ملاحظه می‌شود عدم برقراری فرض استقلال در مدل خطرهای متناسب کاکس سبب اریبی شدید نتایج از مقادیر حقیقی می‌شود. در جدول (۱) نتایج حاصل از این مرحله آورده شده است.

در ادامه با توجه به این که داده‌ها از تابع مفصل فرانک با $\tau=0.8$ تولید شده‌اند، با انتخاب صحیح میزان وابستگی، به دنبال دقت روش پیشنهادی می‌باشیم. همانطور که در جدول(۱) ملاحظه می‌گردد، پارامترهای برآورد شده به مقدار واقعی آن ها بسیار نزدیک است. ضمناً میزان مقادیر MSE نسبت به استفاده از مدل خطرهای متناسب کاکس کوچکتر می‌باشد. لذا این نتایج نشان‌دهنده دقت روش پیشنهادی در برآورد پارامترها است. حال با استفاده از تابع مفصل فرانک و میزان وابستگی معادل ۰/۰۲ =τ و ۰/۰۵ =τ، تأثیر میزان وابستگی بر میزان اریبی نتایج مربوط به برآورد پارامترها را مورد بررسی قرار می‌دهیم. همانطور که ملاحظه می‌گردد در صورتی که درجه وابستگی (مثلاً τ کندال) به اشتباه اختصاص یافته و مورد استفاده قرار گیرد شاهد اریبی قابل ملاحظه‌ای در نتایج مربوط به برآورد پارامترها خواهیم بود به طوری‌که هرچه میزان انحراف از وابستگی حقیقی بیشتر باشد، میزان اریبی از مقادیر حقیقی پارامترها نیز بیشتر خواهد بود.

تحلیل داده‌های بیماران سرطان پستان: پرونده مربوط به ۹۹۶ بیمار مبتلا به سرطان پستان که طی سال‌های ۱۳۸۰ لغایت ۱۳۸۸ جهت درمان به بخش آنکولوژی بیمارستان‌های قائم (عج) و امید مشهد مراجعه نموده‌اند

[Downloaded from journal.nkums.ac.ir on 2026-11-04]

مورد بررسی قرار گرفت که ۵۵۹ بیمار شرایط ورود به مطالعه را پیدا کردند. قابل ذکر است که شرایط خروج از مطالعه عبارت بودند از عدم جراحی بیمار، عدم اتمام دوره درمان به طور کامل و قرار داشتن در مرحله IV بیماری. شروع مطالعه برای افراد زمان جراحی ایشان است. میانگین مدت زمان پیگیری افراد ۳۲.۹ ماه و انحراف معیار آن ۲۱.۹ ماه می‌باشد. هدف اصلی این مطالعه بررسی تأثیر BMI[1] بر بقای بدون بیماری بیماران دارای سرطان پستان پس از درمان ایشان است. از آنجایی که برخی عوامل همچون مرحله بیماری[2] و سن بیمار از عوامل مؤثر شناخته شده بر روی بقای بیماران می‌باشد لذا این متغیرها نیز به مدل اضافه گردیده است. میانگین سنی افراد ۴۷.۷۷ سال و انحراف معیار آن ۱۱.۳۵ سال می‌باشد. میانه سن افراد نیز ۴۷ سال، جوان‌ترین بیمار ۲۵ سال و مسن‌ترین بیمار نیز ۸۰ سال سن داشته‌اند. میانگین BMI افراد ۲۷.۳۷ و انحراف معیار آن نیز ۵.۲۹ می‌باشد. میانه BMI افراد نیز ۲۷.۲۴، کمترین BMI برابر ۱۴.۰۳ وبیشترین آن نیز ۵۴.۸۲ می‌باشد. ۶۱درصد از بیماران در مرحله اولیه بیماری (مراحل I و IIA و IIB) و ۳۹ درصد نیز در مرحله پیشرفته بیماری (مراحل IIIA ، IIIB و IIIC) قرار دارند. از تعداد ۵۵۹ بیمار وارد شده در مطالعه تعداد ۱۵۱ نفر عود[3] داشته‌اند که به عنوان پیشامد مورد نظر در نظر گرفته می‌شود. از تعداد ۴۰۸ نفر باقیمانده تعداد ۵۵ نفر که آخرین وضعیت بالینی ایشان از نظر پزشک مساعد نبوده است از ادامه همکاری با پزشک خودداری نموده‌اند. بدیهی است با توجه به نظر پزشک این افراد را به عنوان سانسور آگاهی بخش درنظر می‌گیریم. تعداد ۳۵۳ نفر باقیمانده نیز در آخرین وضعیت مشاهده شده دارای شرایط مساعدی بوده‌اند که مشاهدات مربوط به این افراد به عنوان سانسور بدون اطلاع یا سانسور مستقل در نظر گرفته می‌شود. بدیهی است استفاده از مدل خطرهای متناسب کاکس در شرایط برقراری فرض استقلال در مورد تمامی مشاهدات سانسور شده اعم از سانسور آگاهی‌بخش یا سانسور بدون اطلاع صحیح می‌باشد؛ لذا انجام آنالیز حساسیت جهت بررسی برقراری یا عدم برقراری فرض استقلال میان زمان‌های پیشامد و زمان‌های مربوط به سانسور آگاهی بخش مفید می‌باشد. یکی از متغیرهای مورد بررسی در این تحقیق گروه سنی بیماران (age groups) می‌باشد که بیماران همانند مطالعه مورد اشاره در منبع [۹] به دو گروه زیر ۳۵ سال، و بزرگتر یا مساوی ۳۵ سال تقسیم شدند. متغیر دیگر، شاخص توده بدنی (BMI) می‌باشد که بیماران بر اساس معیار معرفی شده توسط سازمان بهداشت جهانی (WHO)[4] که افراد با BMI بالاتر از ۳۰ را به عنوان افراد چاق[5] معرفی می‌کند به دو گروه با BMI کمتر از ۳۰ و بزرگتر یا مساوی ۳۰ تقسیم شدند [۱۰]. متغیر دیگر نیز مرحله بیماری (stage) است که بیماران در دو گروه مرحله اولیه[6] و مرحله پیشرفته[7] تقسیم‌بندی شدند. قابل ذکر است با توجه به نظر پزشک متخصص، بیماران دارای مراحل I و IIA و IIB جزء مرحله اولیه و بیماران دارای مراحل IIIA ، IIIB و IIIC جزء مرحله پیشرفته بیماری قرار گرفتند. متغیر وابسته نیز مدت زمان بقای بدون بیماری[8] می‌باشد که به ماه گزارش شده است. قابل ذکر است که در مورد بیماران سانسور شده مقدار این متغیر برابر مدت زمانی است که فرد در مطالعه قرار داشته است.

جهت بررسی فرض استقلال میان زمان‌های پیشامد و زمان‌های مربوط به سانسور آگاهی‌بخش با استفاده از مدل تعمیم‌یافته به انجام آنالیز حساسیت پرداختیم. برای این منظور ازتابع مفصل فرانک و با τ کندال با مقادیر صفر و ۰.۲ و ۰.۵و ۰.۸ استفاده شد. نتایج در جداول (۲) تا (۴) قابل مشاهده است. برای برآورد انحراف معیار پارامترها از خودگردان سازی با ۱۰۰ بار تکرار استفاده شده است. P-مقدارها نیز با توجه به این مطلب که $(\frac{\hat{\beta}}{SE(\hat{\beta})})^2$ دارای توزیع کی دو با یک درجه آزادی است، محاسبه گردیده‌اند. ضمناً در تمام موارد سطح اول متغیر، به عنوان

[1]- body mass index
[2] -stage
[3] -relapse
[4] -World Health Organization
[5] -obese
[6] -early stage
[7]- advance stage
[8] -disease- free survival

[Downloaded from journal.nkums.ac.ir on 2026-11-04]

مبنا در نظر گرفته شده است. همانطور که ملاحظه می‌شود با توجه به نتایج به دست آمده تغییر میزان وابستگی از استقلال تا τ کندال برابر ۰.۸ تأثیری بر معنی‌داری اثر متغیرهای مستقل نگذاشته و تغییرات میزان برآورد نیز ناچیز است. با توجه به نتایج به دست آمده می‌توان برآوردهای حاصل از مدل خطرهای متناسب کاکس را بر مبنای فرض استقلال معتبر دانست. در ادامه با توجه به نتایج به‌دست آمده می‌توان از مدل خطرهای متناسب کاکس جهت تحلیل بیشتر داده‌ها استفاده نمود. ابتدا آزمون متناسب بودن خطرها[1] را بر اساس باقیمانده‌های وزنی که توسط گرامسک[2] و تورنیا[3] در سال ۱۹۹۴ ارایه

جدول۱: مقایسه نتایج حاصل از استفاده از مدل های مختلف

روش تولید داده‌ها : تابع مفصل فرانک با تاو کندال ۰.۸ (n=200,T=34.5% ,C=51.5%,S=14%)

مدل استفاده شده	مدل خطرهای متناسب کاکس		مدل تعمیم یافته با تابع مفصل فرانک با ۰/۰۸ $\tau=$		مدل تعمیم یافته با تابع مفصل فرانک با ۰/۰۵ $\tau=$		مدل تعمیم یافته با تابع مفصل فرانک با $\tau=$ ۰/۰۲
پارامتر	مقدار حقیقی	مقدار برآورد	$\sqrt{MSE}$	مقدار برآورد	$\sqrt{MSE}$	مقدار برآورد	$\sqrt{MSE}$	مقدار برآورد
β_1	۰/۰۵	۰/۲۲۶	۰/۲۸	۰/۴۷	۰/۰۳۱	۰/۳۹	۰/۱۵۲	۰/۳۳	۰/۱۵۵
β_2	۰/۱	۰/۰۵۴	۰/۰۵۶	۰/۱۰	۰/۰۰۹	۰/۰۷	۰/۰۲۷	۰/۰۶	۰/۰۴۲
β_{c1}	۰/۶	۰/۸۴۱	۰/۲۵	۰/۶۱	۰/۰۱۵	۰/۶۷	۰/۰۸۷	۰/۷۷	۰/۱۸۴
β_{c2}	۰/۳	۰/۳۴۱	۰/۰۵	۰/۲۹	۰/۰۰۹	۰/۳۱	۰/۰۱۸	۰/۳۲	۰/۰۲۵

جدول ۲: نتایج مربوط به برآورد پارامتر متغیر گروه سنی

	$\tau=$۰	$\tau=$۰/۲	$\tau=$۰/۵	$\tau=$۰/۸
برآورد پارامتر	-۰/۳۷۴	-۰/۳۵۲	-۰/۳۶۱	-۰/۳۳۸
انحراف معیار پارامتر	۰/۲۴۹	۰/۲۰۹	۰/۱۹۲	۰/۱۸۹
کران پائین بازه اطمینان	-۰/۸۶۲	-۰/۷۶۲	-۰/۷۳۷	-۰/۷۰۸
کران بالای بازه اطمینان	۰/۱۱۴	۰/۰۵۸	۰/۰۱۵	۰/۰۳۲
مقدارP	۰/۱۳۳	۰/۰۹۲۱	۰/۰۶۰	۰/۰۷۳

[1]- proportıonal hazards assumptıon
[2] -Grambsch
[3] -Therneau

[Downloaded from journal.nkums.ac.ir on 2026-11-04]

جدول۳ : نتایج مربوط به برارود پارامتر متغیر سطوح BMI

	τ=۰	τ=۰/۰۲	τ=۰/۵	۰ τ=/۰۸
برآورد پارامتر	۰/۲۸۵	۰/۲۹۲	۰/۲۷۹	۰/۲۵۷
انحراف معیار پارامتر	۰/۱۸۱	۰/۱۷۲	۰/۱۶۱	۰/۱۶۸
کران پائین بازه اطمینان	-۰/۰۶۹	-۰/۰۴۵	-۰/۰۳۶	-۰/۰۷۲
کران بالای بازه اطمینان	۰/۶۴۰	۰/۶۲۹	۰/۵۹۴	۰/۷۸۰
مقدارP	۰/۱۱۵	۰/۰۸۹	۰/۰۸۳۱	۰/۱۲۶

جدول۴: نتایج مربوط به برآورد پارامتر متغیر مرحله بیماری

	τ=۰	τ=۰.۲	τ=۰.۵	τ=۰.۸
برآورد پارامتر	۰.۵۵۶	۰.۵۲۷	۰.۴۹۲	۰.۴۶۹
انحراف معیار پارامتر	۰.۱۶۴	۰.۱۴۹	۰.۱۳۶	۰.۱۲۷
کران پائین بازه اطمینان	۰.۲۳۴	۰.۲۳۵	۰.۲۲۵	۰.۲۲۰
کران بالای بازه اطمینان	۰.۷۲	۰.۸۱۹	۰.۷۵۸	۰.۷۱۸
مقدارP	۰.۰۰۱	۰.۰۰۰۴	۰.۰۰۰۲	۰.۰۰۰۲

جدول۵: نتایج آزمون متناسب بودن خطرها در مورد هر یک از متغیرهای توضیحی

متغیر	آماره آزمون (کی دو)	درجه آزادی	P-Value
مرحله بیماری	۱.۸۳۰۸	۱	۰.۱۷۶
BMI سطح	۰.۰۵۱۱	۱	۰.۸۲۱۲
گروه سنی	۱.۳۴۱۵	۱	۰.۲۴۶۸

جدول۶: متغیرهایی که بر اساس روش گام به گام پیشرونده، در مدل باقی می‌مانند

متغیر	برآورد(B)	خطای استاندارد	آماره والد	درجه آزادی	P-Value	Exp(B)
مرحله بیماری	۰/۵۳۹	۰/۱۶۳	۱۰/۸۸۱	۱	۰/۰۰۱	۱/۷۱۴

[Downloaded from journal.nkums.ac.ir on 2026-11-04]

جدول(۷): متغیرهایی که بر اساس روش گام به گام پیشرونده، در مدل باقی نمی‌مانند

	Score	df	P Value
BMI سطح	۱/۹۳۰	۱	۰/۱۶۵
گروه سنی	۱/۶۹۰	۱	۰/۱۹۴
*گروه سنی BMI سطح	۲/۰۷۰	۱	۰/۱۵۰
گروه سنی*مرحله بیماری	۰/۰۳۸	۱	۰/۸۴۵
*مرحله بیماریBMI سطح	۱/۶۷۳	۱	۰/۱۹۶
*مرحله بیماری*گروه سنیBMI سطح	۲/۳۴۸	۱	۰/۱۲۵

شده است را انجام می‌دهیم [۱۱]. جهت انجام این آزمون از دستور cox.zph در نرم‌افزار R استفاده نمودیم .همان-طوری که در جدول (۵) مشاهده می‌شود فرض متناسب بودن خطرها در مورد هیچکدام از متغیرها در سطح ۰.۰۵ رد نمی‌شود.در ادامه با استفاده از روش گام به گام پیشرونده[1] مبتنی بر آزمون والد[2] به انتخاب متغیرهای مستقل می‌پردازیم. در ابتدا متغیرهای گروه سنی (agegroup)، سطح شاخص توده بدن (BMI)، مرحله بیماری (stage) مدل را به همراه اثرات متقابل آن ها به مدل وارد می‌نماییم. بر اساس نتایج به دست آمده، تنها متغیر مرحله بیماری در مدل باقی می‌ماند.

با توجه به نتایج فوق بر اساس مدل خطرهای متناسب کاکس تنها اثر مرحله بیماری در بقای بیماران معنی دار است به طوری‌که خطر عود بیماری در مرحله پیشرفته ۱.۷ برابر خطر عود در بیمارانی است که در مرحله اولیه بیماری قرار داشته ومورد درمان قرار گرفته‌اند.

بحث

یکی از رایج‌ترین روش‌های تحلیل داده‌های بقا استفاده از مدل خطرهای متناسب کاکس می‌باشد. دلایل رایج بودن این روش مواردی همچون استفاده از زمان بقا، در نظر گرفتن اطلاعات داده‌های سانسور شده و عدم وابستگی به شکل توزیع داده‌ها می‌باشد [۱۲]. علیرغم این ویژگی‌ها درصورتیکه فرض استقلال میان مشاهدات سانسور و پیشامد، که یکی از فرض های اساسی مدل خطرهای متناسب کاکس می‌باشد، برقرار نباشد ممکن است شاهد اریبی در برآورد پارامترها باشیم. این میزان از اریبی به میزان و جهت وابستگی میان زمان سانسور و زمان پیشامد بستگی دارد لذا انجام آنالیز حساسیت با هدف بررسی نتایج حاصل از تغییرات میزان وابستگی بر خروجی مدل مفید می‌باشد. به عبارتی ما پایداری نتایج مدل خطرهای متناسب کاکس را با تغییر درجه ارتباط میان زمان‌های سانسور و زمان پیشامد بررسی می‌نماییم. ما با استفاده از توابع مفصل، مدل خطرهای متناسب کاکس را به نحوی تعمیم دادیم که امکان لحاظ نمودن ارتباط میان سانسور و پیشامد در مدل وجود داشته باشد. یکی از مزایای استفاده از توابع مفصل این است که بدون اطلاع از توزیع-های حاشیه‌ای می‌توانیم توزیع توأم مشاهدات را به دست آوریم. بدیهی است چنین خاصیتی جهت به‌دست آوردن توزیع توأم سانسور و پیشامد مناسب می‌باشد. از آنجاییکه انتخاب نادرست تابع مفصل سبب اریبی ناچیز در برآورد پارامترها می‌گردد لذا ما در این تحقیق از تابع مفصل فرانک به دلیل خواص آن از جمله امکان تغییر تاو کندال در بین مقادیر ۱- تا ۱+ استفاده نمودیم [۳]. در مجموعه داده‌های مربوط به بیماران دارای سرطان پستان که طی سال‌های ۱۳۸۰ لغایت ۱۳۸۸ به بخش آنکولوژی

[1] -stepwise forward
[2] -Wald test

[Downloaded from journal.nkums.ac.ir on 2026-11-04]

بیمارستان‌های قائم (عج) و امید مراجعه نموده بودند، ۹.۸ درصد از بیماران پس از درمان از ادامه همکاری با پزشک خودداری نموده بودند که با توجه به بررسی‌های به‌عمل آمده توسط پزشک متخصص وضعیت احتمال عود بیماری در ایشان بالا بود لذا برقراری فرض استقلال میان زمان سانسور و زمان پیشامد در مورد این افراد مورد تردید واقع شد. بنابراین با روش ارایه شده اقدام به برآورد پارامترهای مدل در خصوص متغیرهای گروه سنی، سطح BMI و مرحله بیماری با استفاده از روش ارایه شده و با فرض ضرائب وابستگی تاو کندال برابر صفر، ۰.۲ ، ۰.۵ و ۰.۸ نمودیم. نتایج نشان‌دهنده عدم نوسانات زیاد در برآورد پارامترها و نیز عدم تغییر در معنی‌داری اثر آن ها بود. با توجه به این موارد به این نتیجه رسیدیم که نتایج حاصل از برازش مدل خطرهای متناسب کاکس تحت فرض استقلال سانسور و پیشامد در مورد این افراد می‌تواند برقرار باشد. لذا با توجه به برقراری فرض استقلال با استفاده از مدل خطرهای متناسب کاکس به تحلیل مجدد اطلاعات پرداختیم.

نتیجه گیری

نتایج حاصل از استفاده از مدل خطرهای متناسب کاکس نشان‌دهنده تأثیر مرحله بیماری در بقای بیماران بود به نحوی که خطر عود بیماری در افرادی که در زمان تشخیص بیماری در مرحله پیشرفته بیماری قرار داشته‌اند ۱.۷ برابر بیمارانی است که در زمان تشخیص اولیه بیماری در مرحله اولیه بیماری قرار داشته‌اند. بنابراین تشخیص بیماری در مراحل اولیه نقش مهمی در موفقیت درمان و عدم عود بیماری‌دارد.

تشکر و قدردانی

این مقاله حاصل پایان نامه تحقیقاتی کارشناسی ارشد آمار زیستی می باشد . نویسندگان از معاونت محترم پژوهشی دانشگاه علوم پزشکی مشهد که هزینه های اجرای این تحقیق را تقبل نموده اند تقدیر و تشکر می نمایند .

References

1. Kleinbaum DG, Klein M, Survival Analysi A Self-Learning Text, Second edition, Springer, New York 2005; 1-42.
2. Cox DR, Regression analysis and life table (with discussion)', Journal of the Royal Statistical Society, Series B 1972; 34: 187-222.
3.Huang X, Zhang N, Regression Survival Analysis with an Assumed Copula for Dependent Censoring: A Sensitivity Analysis Approach, Biometrics 2008; 64: 1090–1099.
4. Kleinbaum DG, Klein M, Survival Analysi A Self-Learning Text, Second edition, Springer, New York 2005; 403-407.
5. Zheng M , Klein J. P, Estimates of marginal survival for dependent competing risks based on an assumed copula, Biometrika 1995; 82: 127-138.
6.Zheng M, Klein JP, A self-consistent estimator of marginal survival functions based on dependent competing risk data and an assumed copula, Communications in Statistics, Part A - Theory and Methods 1994; 23 :2299-2311.
7. Nelsen RB, An Introduction to Copulas, Springer, New York,2006.
8.Nelsen RB, Properties of a one-parameter family of bivariate distributions with specified marginals, Communications in Statistics, Part A: Theory and Methods1986;15:3277-3285.
9. Fazl alizadeh A, Hatami A, The effect of reducing the course of radiotherapy after mastectomy on rate of local recurrence in breast cancer patients, Pejouhandeh,2004; 9:209-214[Persian]
10. Available from: URL: http://apps.who.int/bmi
11. Grambsch P, Therneau T, Proportional hazards tests and diagnostics based on weighted residuals, *Biometrika* 1994; 81: 515-26.
12. Kleinbaum DG, Klein M, Survival Analysi A Self-Learning Text, Second edition, Springer, New York, 2005 ;83-127.

[Downloaded from journal.nkums.ac.ir on 2026-11-04]

Analysis of survival data of patients with breast cancer with Cox proportional hazards model for dependent censoring based on copula function

Shakeri MT[1], Ghavami Ghanbarabadi V[2], Esmaeili H[1], Gabbari Noghabi H[3], Homaei Shandiz F[4], Baghishani H[5]*

[1]Associated professor, Department Of Environmental Statistic and Epidemiology, School Of Health, Mashhad University Of Medical Sciences
[2]M.Sc of Environmental Statistic, Statistic Management and Information Technology, Mashhad University of Medical Sciences
[3]Associated Professor, Department Of Statistic, School of Mathematical Sciences, Mashhad University of Medical Sciences
[4]Associated Professor, Department of Oncology and Radiotherapy, Omid Hospital, Mashhad University of Medical Sciences
[5]Assistant Professor, Department Of Statistic, School of Mathematical Sciences, Shahrood University

***Corresponding Author:** Shahid Fakori Square, Town Of Danesh and Salamat, Management Of Statistic and Information Sciences, Mashhad , Iran*
Email:ghavamighv1@mums.ac.ir

Submitted: 2011 Dec 14 **Accepted:** 26 Jan 201

Abstract

Background & objectives: *In the common Cox proportional hazards model one of the basic assumptions is independence between censoring time and event time. In clinical studies, when censoring is caused by competing risks or patient withdrawal, there is always a concern about the validity of treatment effect estimates that are obtained under the assumption of independent censoring. Introduction a solution for checking the assumption of independent and extending the Cox model for dependent censoring is useful.*

Material &Methods: *To achieve this goal, we use copula function and extend the Cox model. This model can also perform a sensitivity analysis for checking the assumption of independent. We generalize the likelihood function in R software and estimate the parameters by using a iteration algorithm. We apply the proposed method to the data of breast cancer patients at Ghaem and Omid hospitals and check the independence assumption.*

Results: *We show by Simulation that this algorithm works well and show the effect of dependent censoring on estimated parameters. We apply the proposed method to the data of breast cancer patients at Ghaem and Omid hospitals and show that the independence assumption is true.*

Conclusion: *Not assumption of independence causes bias in estimation in Cox model and the bias is dependent on the degree of association between event and censoring times. In dataset of breast cancer patients the primary diagnosis is an important role in Success in the treatment.*

Key words: *Cox proportional hazards model, sensitivity analysis, copula function, competing risks, breast cancer.*

[Downloaded from journal.nkums.ac.ir on 2026-11-04]

[Downloaded from journal.nkums.ac.ir on 2026-11-04]