

مدل‌بندی عوامل موثر بر طول مدت بیکاری با استفاده از تجزیه و تحلیل بقا

فاطمه زنجیران[†] و کیومرث مترجم^{‡*}

[†] مؤسسه مهرآفرینان راهیان سلامت
[‡] دانشگاه تربیت مدرس

چکیده. در بسیاری از پژوهش‌های آماری با متغیرهایی تحت عنوان مدت زمان بیکاری، مدت زمان ابتلا به بیماری تا مرگ، طول عمر یک قطعه، مدت زمان تصویب لایحه‌ای در مجلس و موارد مشابه دیگری رو به رو می‌شویم. در این‌گونه مطالعات هدف یافتن و مدل‌بندی منابع تأثیرگذار بر مدت زمان مورد بررسی است تا با شناخت و تغییر منابع خطر، بتوان مدت زمان مذکور را کنترل نمود. در این مقاله برای مدل‌بندی عوامل موثر بر طول مدت زمان بیکاری از مفاهیم تجزیه و تحلیل بقا استفاده شده است. برای بررسی تأثیرات متغیرهای تبیینی بر طول مدت زمان بیکاری از مدل مخاطرات متناسب کاکس و شکنندگی استفاده شده است. داده‌های این مطالعه برای بررسی تأثیر متغیرهای جنسیت، سن، وضعیت تأهل و سطح تحصیلات به‌عنوان متغیرهای پیشگوی موثر بر مدت زمان بیکاری در کلانشهر تهران گردآوری شده است. نمونه شامل ۵۴۶ نفر است که اطلاعات مربوط به تاریخ شروع و تاریخ پایان دوره‌ی بیکاری برای هر فرد از ابتدای فروردین سال ۱۳۹۷ تا ابتدای فروردین سال ۱۳۹۸ ثبت شده است. نتایج به‌دست آمده نشان می‌دهند که طول دوره بیکاری افراد متأثر از متغیرهای وضعیت تأهل، سطح تحصیلات و جنسیت است. واژگان کلیدی: مدل شکنندگی، مدل خطرهای متناسب کاکس، داده بقا، سانسور، مدت زمان بیکاری.

* نویسنده‌ی عهده‌دار مکاتبات

دریافت: ۱۳۹۹/۱۰/۳۰، پذیرش: ۱۳۹۹/۱۲/۲۰.

۱- مقدمه

آنالیز بقاء^۱، تحلیل ماندگاری، یا تحلیل بقاء یکی از مباحث علم آمار است که در رشته‌های مختلفی از جمله علوم کامپیوتر، اپیدمیولوژی، اقتصاد و علوم اجتماعی کاربرد فراوان دارد. تحلیل بقاء به مجموعه‌ای از روش‌های آماری گفته می‌شود که در آن‌ها متغیر مطلوب، زمان رخداد یک پدیده است. این موضوع در علوم مهندسی، نظریه قابلیت اطمینان نامیده می‌شود. ویژگی خاص تحلیل بقاء این است که با داده‌های سانسور شده وفق دارد و از این رو از اطلاعات مربوط به داده‌هایی که در زمان ارزیابی، هنوز حادثه مورد نظر را تجربه نکرده‌اند نیز استفاده می‌نماید. تحلیل بقاء به‌عنوان یک روش آماری که اساساً برای تحلیل داده‌های طول عمر متولد شده است می‌تواند در بسیاری دیگر از زمینه‌ها از جمله مسائل اقتصادی و اجتماعی مورد استفاده قرار گیرد. این روش آماری اطلاعات حاصل از داده‌های سانسور شده و مشاهده شده را با یکدیگر ترکیب نموده و بدین ترتیب تحلیل آماری داده‌های سانسور شده را امکان‌پذیر می‌کند. علاوه بر این، خصوصیت غیر خطی داده‌های زمان بقا را نیز مورد توجه قرار می‌دهد.

۲- ادبیات تحقیق

کاکس در سال ۱۹۷۲ مدلی را برای تحلیل داده‌های بقا معرفی نمود که تا کنون به‌صورت گسترده مورد استفاده قرار گرفته است. در شرایطی که هدف تحقیق، بررسی همزمان اثرات چند متغیر بر زمان بقا باشد، مدل کاکس کاربرد فراوانی دارد. مدل معرفی شده توسط کاکس، مدل مخاطرات متناسب^۲ نیز نام گرفته است. باید توجه داشت که استفاده از این مدل مستلزم برقراری فرضیاتی است که یکی از مهم‌ترین آنها استقلال داده‌های بقا است که در صورت تشخیص نادرست یا بی‌توجهی به آن، می‌تواند موجب استخراج نتایج نادرست شود. از دیگر ویژگی‌های مدل کاکس این است که بدون در نظر گرفتن هیچ‌گونه فرض توزیعی درباره تابع خطر پایه، می‌توان آن را به داده‌های بقا برازش داد، اگرچه در مدل پارامتری کاکس، می‌توان برای تابع خطر پایه، فرض توزیعی نیز در نظر گرفت. بر این اساس، در مدل کاکس می‌توان مدل‌های پارامتری و نیمه پارامتری را به داده‌های بقا برازش داد. سادگی و انعطاف‌پذیری این

مدل در محاسبات سبب شده است مورد توجه طیف گسترده‌ای از کاربران قرار گیرد، اما باید توجه داشت که این مدل به علت در نظر نگرفتن عوامل خطر ناشناخته لزوماً بهترین مدل را به داده‌ها برازش نمی‌دهد. یکی از نکات مورد توجه محققان در تحلیل داده‌های بقا وجود عوامل خطر ناشناخته است. گاهی به دلیل شرایط خاص مطالعه یا وجود برخی محدودیت‌های اقتصادی امکان اندازه‌گیری برخی از عوامل خطر در قالب متغیرهای تبیینی وجود ندارد. در نظر نگرفتن برخی عوامل تأثیرگذار بر زمان بقا یا غیر قابل مشاهده بودن این عوامل موجب پیدایش عوامل خطر ناشناخته در مدل‌بندی داده‌های بقا می‌شود و سبب می‌شود بخش عمده‌ای از تغییرات تابع خطر که می‌توانست با وجود عوامل ناشناخته در مدل تبیین شود، با جمله خطای مدل جمع شده، باعث افزایش تغییرات تابع خطر، تشخیص نادرست تابع بقا و برآورد نامناسب ضرایب رگرسیونی مدل شود. با وجود این که مدل مخاطرات متناسب کاکس یکی از پرکاربردترین مدل‌ها در برازش به داده‌های بقا است، اما منظور کردن عوامل خطر ناشناخته مؤثر بر زمان بقا در این مدل میسر نیست؛ یک راه حل برای وارد کردن اثر عوامل خطر ناشناخته در مدل بقا استفاده از مدل‌های شکنندگی^۳ است [۴]. واپل و همکاران برای اولین بار در سال ۱۹۷۹ عبارت شکنندگی را برای مدل‌های بقای تک متغیره مورد استفاده قرار دادند [۱۲]. سپس کلایتون و کازیک در سال ۱۹۸۵ مدل‌های شکنندگی را به کار بردند [۳]. در این مدل‌ها یک متغیر تصادفی با تکیه‌گاه مثبت و میانگین یک در مدل کاکس ضرب می‌شود و به همین دلیل مدل‌های شکنندگی را کاکس اصلاح شده نیز می‌گویند [۱].

یکی از کاربردهای مهم تحلیل بقا علاوه بر استفاده آن در تحلیل داده‌های طول عمر بکارگیری آن در تحلیل داده‌های اقتصادی و موضوعاتی مانند بیکاری است. بیکاری و پیامدهای آن در جوامع امروزی به صورت یک معضل درآمده و دغدغه‌ی ایجاد اشتغال و کاهش بیکاری، همه‌ی سطوح مدیریتی کشورها را به چالش کشانده است. در این میان یکی از شاخص‌های مهمی که می‌تواند به شناخت بیشتر معضل بیکاری کمک کند، طول مدت بیکاری افراد است.

منظور از مدت زمان بیکاری، مدت زمانی است که یک فرد بیکار و جویای کار است. مدت زمان بیکاری، متغیر مهمی است که تغییرات بازار کار را هنگام تجزیه و تحلیل

رابطه بین جریان اشتغال، بیکاری و بازارکار توضیح می‌دهد و به طور کلی، در مدل‌سازی ایجاد مشاغل، مورد استفاده قرار می‌گیرد [۲].

از این رو مطالعه شاخص‌های بازارکار با تجزیه و تحلیل متغیرهای ثابت، مانند نرخ اشتغال و تعداد بیکاران، به منظور اتخاذ سیاست‌های مناسب در بازارکار، مطالعه ورود و خروج افراد به بازارکار و توانایی افراد در یافتن سریع مشاغل جایگزین و همچنین تأثیر مدت زمان بیکاری بر گروه‌های مختلف نیروی کار بسیار ضروری به نظر می‌رسد [۷].

طول مدت بیکاری از اهمیت اقتصادی و اجتماعی فراوانی برخوردار است؛ بازارکاری که در آن تعداد افراد کمی بیکار می‌شوند اما برای مدت طولانی بیکار می‌مانند ناامن‌تر از بازارکاری است که در آن تعداد افراد بیشتری بیکار می‌شوند اما فقط برای مدت کوتاهی در آن موقعیت باقی می‌مانند. بدیهی است که در دوران رکود اقتصادی، نرخ بیکاری و طول مدت بیکاری افزایش می‌یابد که این امر پیامدهای مهمی در هزینه‌های خانوار و توانایی پرداخت مالی آنان دارد بنا بر این مدیریت این موضوع به سیاست‌های خاص بازارکار نیاز دارد. در حال حاضر این پیامدها با کاهش نرخ خروج بیکاری (به دلیل اشتغال یا پایان دوره قانونی بیکاری) و بیکاری بلند مدت اهمیت بیشتری پیدا می‌کنند. در سطح فردی نیز بیکاری طولانی مدت نه تنها افراد را از معیشت و فعالیت معنادار در زندگی محروم می‌کند بلکه موجب نارضایتی عمومی و افزایش پدیده‌های منفی اجتماعی نیز می‌شود [۹].

برای شناسایی عواملی که در طول مدت بیکاری و نرخ خروج تأثیر می‌گذارند، از روش‌های تجزیه و تحلیل بقا در تحقیقات بازارکار استفاده می‌شود. در سال‌های گذشته، مدل‌های مدت در مطالعات اقتصادی و علوم اجتماعی برای تجزیه و تحلیل طول مدت بیکاری و طول مدت اعتصابات کارگری استفاده شده است [۷].

محققین بسیاری از رویکرد تحلیل بقا برای تجزیه و تحلیل داده‌های بیکاری استفاده نموده‌اند به عنوان یکی از محققین پیشرو در این زمینه، لنکستر در سال ۱۹۷۹ مدت زمان بیکاری را با استفاده از اثر تصادفی مورد بررسی قرار داد و استفاده از مدل خطر را در مطالعات اقتصادی متداول نمود [۸]. دارابای و همکاران در سال ۲۰۱۷ روش‌های تجزیه و تحلیل بقا را برای تحلیل مدت زمان بیکاری در کشور رومانی مورد استفاده قرار دادند آنها برای برآورد مدت زمان بیکاری از روش کاپلان مهیر کمک

گرفتند [۷]. نونیا و انجیوهو در سال ۲۰۱۸ مدت زمان بیکاری را در کشور آفریقای جنوبی با استفاده از تابع بقا مدل‌سازی کردند [۱۱].

با توجه به اهمیت مسئله بیکاری به‌عنوان یک معضل اقتصادی-اجتماعی در کشورمان، در این مقاله به تحلیل داده‌های مربوط به مدت زمان بیکاری پرداخته می‌شود. در ادامه و در بخش (۳) معرفی مدل‌های بقا در تحلیل داده‌های بیکاری مورد توجه قرار گرفته است. در بخش (۴) به تشریح داده‌های مربوط به مدت زمان بیکاری پرداخته شده است سپس اطلاعات با استفاده از روش‌های تحلیل بقا مورد تجزیه و تحلیل قرار خواهد گرفت در انتها نیز به بحث در خصوص نتایج می‌پردازیم و پیشنهاداتی جهت مطالعات آتی مطرح خواهد شد.

۳- مدل‌بندی زمان‌های بقا

در بخش قبل به ویژگی‌های داده‌های بقا و امکان ورود سانسور در تحلیل بقا اشاره شد. در این بخش نحوه مدل‌بندی داده‌های بقا بر اساس متغیرهای تبیینی مورد بررسی قرار می‌گیرد. علاوه بر این، مدل‌بندی داده‌های بقا بر اساس تابع خطر با مدل‌های مخاطرات متناسب، شکنندگی و شکنندگی مشترک نیز معرفی می‌شوند.

۳-۱- مدل خطرهای متناسب

یکی از مدل‌های متداول برای مدل‌بندی داده‌های بقا، مدل خطرهای متناسب معروف به مدل کاکس است [۴]. در این مدل تابع خطر به صورت زیر تعریف می‌شود:

$$h(t|X) = h_0(t)g(X)$$

که در آن $h_0(\cdot)$ تابع خطر پایه، $X' = (X_1, \dots, X_k)$ متغیرهای تبیینی و $g(\cdot)$ تابعی مثبت است. در این مدل معمولاً $g(X) = e^{\beta'X}$ در نظر گرفته می‌شود، که در آن بردار $\beta' = (\beta_1, \dots, \beta_k)$ پارامترهای اثرات ثابت متغیرهای تبیینی یا پارامترهای رگرسیونی مدل هستند. بنا بر این داریم:

$$(1) \quad h(t|X) = h_0(t)e^{\beta'X}$$

وقتی تابع خطر پایه یک مدل پارامتری در نظر گرفته شود مدل (۱) کاملاً پارامتری خواهد بود و از طریق تابع درستنمایی می‌توان به برآوردهای ماکسیمم درستنمایی پارامترها دست یافت. به‌طور مثال تابع خطر (۱) بر اساس خطر پایه وایبول به‌صورت:

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

به‌دست می‌آید.

اما اگر تابع خطر پایه نامعلوم باشد، مدل (۱) نیمه‌پارامتری خواهد بود. در این حالت نمی‌توان تابع درستنمایی مدل را به روش فوق به دست آورد. کاکس یک تابع درستنمایی جزئی برای برآورد پارامترهای این مدل نیمه‌پارامتری معرفی کرد [۴] و [۵]. تابع درستنمایی زمان‌های بقا بر اساس رابطه (۲) به‌صورت

$$(2) \quad \prod_{i=1}^n (h_0(t_i) e^{\beta'X_i})^{\delta_i} \exp(-H_0(t_i) e^{\beta'X_i})$$

به‌دست می‌آید. برای محاسبه تابع درستنمایی نیمه‌پارامتری مدل، $h_0(\cdot)$ یک تابع گسسته در نظر گرفته می‌شود که به غیر از زمان‌های t_i ای که اتفاق مورد نظر رخ داده باشد صفر است. بنا بر این تابع توزیع خطر پایه به‌صورت

$$H_0(t) = \sum_{t_j \leq t} h_0(t_j)$$

به‌دست می‌آید. با قرار دادن این تابع خطر پایه در (۲) داریم

$$(3) \quad \prod_{i=1}^n (h_0(t_i) e^{\beta'X_i} \exp(-h_0(t_i) \sum_{j \in R(t_i)} e^{\beta'X_j}))^{\delta_i}$$

که در آن $R(t)$ مجموعه تمام افراد یا واحدهای آماری در معرض خطر است یعنی تمام افرادی که رخداد مورد نظر تا زمان t برای آن‌ها رخ نداده است. با گرفتن لگاریتم و مشتق جزئی تابع (۳) نسبت به $h_0(t_i)$ و برابر صفر قرار دادن آن داریم

$$(4) \quad \hat{h}_0(t_i) = \frac{\delta_i}{\sum_{j \in R(t_i)} e^{\beta'X_j}}$$

بنا بر این تابع درستنمایی جزئی به‌صورت

$$L(\beta) = \prod_{i=1}^n \left(\frac{e^{\beta'X}}{\sum_{j \in R(t_i)} e^{\beta'X}} \right)^{\delta_i}$$

حاصل می‌شود، که مستقل از تابع خطر پایه است و مشتق دوم آن همواره نامثبت است. برآورد تابع تجمعی خطر پایه بر اساس (۴) به صورت

$$\hat{H}(t) = \sum_{i=1}^n \frac{I(t_i \leq t) \delta_i}{\sum_{j \in R(t_i)} e^{\beta'X}}$$

به دست می‌آید.

۳-۲- مدل شکنندگی

در استفاده از مدل (۱) فرض می‌شود که تغییرپذیری فردی در تابع خطر از طریق متغیرهای تبیینی به‌طور کامل قابل تبیین هستند. اما اگر منابع خطر دیگری وجود داشته باشند که بر زمان بقا تأثیرگذار بوده و در عین حال غیر قابل مشاهده یا غیر قابل اندازه‌گیری باشند، برای در نظر گرفتن این عوامل خطر از مدل‌های شکنندگی استفاده می‌شود.

وایل و همکاران در سال ۱۹۷۹ برای اولین بار عبارت شکنندگی را برای مدل‌های بقای تک متغیره مورد استفاده قرار دادند [۱۲]. سپس کلایتون و کازیک مدل‌های شکنندگی را به‌کار بردند [۳]. بعد از آن لین و لی عملکرد مدل کاکس را در هنگام وجود عوامل ناشناخته مورد بررسی قرار دادند [۱۰]. شکنندگی یک مؤلفه تصادفی است که برای لحاظ کردن اثر عوامل مشاهده نشده یا ناشناخته وارد مدل می‌شود که توسط متغیرهای تبیینی، امکان ورود آنها در مدل میسر نیست.

در واقع مدل‌های شکنندگی نوع خاصی از مدل‌های اثرات تصادفی برای مدل‌بندی داده‌های بقا هستند. مدل‌های شکنندگی با وارد کردن یک اثر تصادفی ضربی به مدل (۱) به صورت

$$(5) \quad h(t | \mathbf{X}, \mathbf{W}) = Wh.(t) e^{\beta'X}$$

تعریف می‌شود که در آن بردار متغیرهای تبیینی، $X' = (X_1, \dots, X_k)$ بردار متغیرهای تبیینی، $\beta' = (\beta_1, \dots, \beta_k)$ بردار اثرات ثابت و $\mathbf{W}' = (W_1, \dots, W_n)$ بردار متغیر تصادفی نامنفی شکنندگی است. در این مدل \mathbf{W} دارای توزیع استاندارد شده‌ای است به‌طوری که امید ریاضی آن برابر بردار یک است.

در مدل شکنندگی (۵) به طور کلی اثرات تصادفی W_1, \dots, W_n غیر قابل مشاهده‌اند و برای استنباط در مورد این اثرات ابتدا فرض می‌شود که این اثرات از یک توزیع آماری با تکیه‌گاه مثبت پیروی می‌کنند. بنا بر این به شرط مشاهده‌ی W_1, \dots, W_n ، تابع درست‌نمایی به صورت

$$L(\beta|W) = \prod_{i=1}^n (W_i h_{\circ}(t_i) e^{\beta' X_i})^{\delta_i} \exp(-W_i H_{\circ}(t_i) e^{\beta' X_i})$$

و تابع بقا به صورت

$$S(t|W) = e^{\int_0^t h(s|W) ds} = e^{-W \int_0^t h_{\circ}(s) ds} = e^{-WH_{\circ}(t)}$$

به دست خواهد آمد و داریم

$$S(t) = E(S(t|W)) = E\left(e^{-WH_{\circ}(t)}\right) = LA\left(H_{\circ}(t)\right)$$

که در آن $LA(\cdot)$ تبدیل لاپلاس است.

تبدیل لاپلاس متغیر تصادفی مثبت T^* به صورت

$$LA(u) = E(e^{-uT^*}) = \int_0^{\infty} e^{-uT^*} f(t) dt$$

تعریف می‌شود. بنا بر این داریم

$$f(t) = -h_{\circ}(t) LA'\left(H_{\circ}(t)\right)$$

$$h(t) = h_{\circ}(t) \frac{LA'\left(H_{\circ}(t)\right)}{LA\left(H_{\circ}(t)\right)}$$

$$E(W) = -LA'\left(\circ\right)$$

$$Var(W) = LA''\left(\circ\right) - \left(LA'\left(\circ\right)\right)^2$$

اگر در مدل شکنندگی معرفی شده مؤلفه شکنندگی نشان‌دهنده‌ی همبستگی درون‌خوشه‌ای باشد، این مدل را مدل شکنندگی مشترک نامیده می‌شود [۱۳]. در این مدل بر خلاف مدل شکنندگی ارائه شده در (۵) به جای در نظر گرفتن مؤلفه شکنندگی

برای هر فرد، برای هر خوشه یک مؤلفه شکنندگی در نظر گرفته می‌شود این نوع از مدهای شکنندگی در مطالعات طول عمر دوقلوها رایج است. در مواقعی که عوامل ناشناخته یا غیر قابل محاسبه ناشی از خوشه‌ای بودن داده‌ها باشد از مدل شکنندگی مشترک استفاده می‌شود.

۴- معرفی داده‌های مطالعه

داده‌های این مطالعه برای بررسی تأثیر متغیرهای جنسیت، سن، میزان تحصیلات و وضعیت تأهل به‌عنوان متغیرهای تبیینی مؤثر بر مدت زمان بیکاری گردآوری و با استفاده از روش‌های بقا مورد تجزیه و تحلیل قرار گرفته است. نمونه شامل ۵۴۶ نفر در کلانشهر تهران است که است که اطلاعات مربوط به تاریخ شروع و تاریخ پایان دوره بیکاری برای هر شخص از ابتدای فروردین سال ۱۳۹۷ تا ابتدای فروردین سال ۱۳۹۸ ثبت شده است. داده‌ها نشان می‌دهد که ۳۵/۳۵ درصد بیکاران نمونه را زنان و ۶۴/۶۵ درصد را مردان تشکیل می‌دهند. در خلال مطالعه یک‌ساله ۳۲۹ نفر موفق به یافتن شغل و ۲۱۷ نفر در پایان مطالعه کماکان بیکار بوده‌اند و این حیث سانسور راست ۳۹/۷ درصدی در داده‌ها وجود دارد. لازم به ذکر است که حداقل مدت زمان دوره بیکاری برای افراد این مطالعه ۳۶ روز و حداکثر ۳۴۰ روز است. میانگین مدت زمان بیکاری نیز ۲۶۵ روز بوده است. در ادامه ابتدا با استفاده از روش کاپلان مهیر به توصیف داده‌ها پرداخته می‌شود سپس با استفاده از مدل‌های کاکس و شکنندگی نتایج تحلیلی بیشتری از داده‌ها را ارائه خواهد شد.

۵- نتایج و یافته‌ها

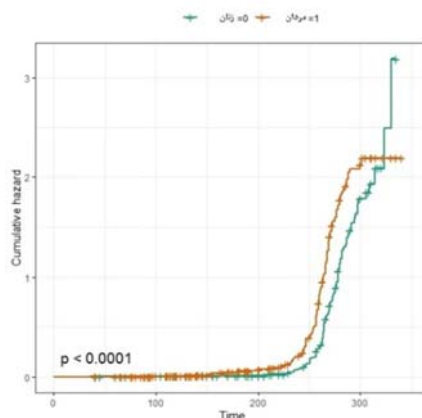
۵-۱- نتایج تجزیه و تحلیل کاپلان مهیر

اطلاعات توصیفی مربوط به نمونه در جدول ۱ نشان داده شده است. اعداد داخل پرانتز بیانگر انحراف استاندارد است. با توجه به جدول ۱ میانگین مدت بیکاری برای زنان ۲۷۵ روز و در مردان ۲۲۲ روز است.

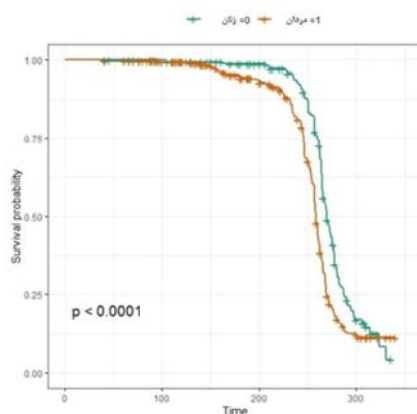
جدول ۱- اطلاعات توصیفی مربوط به داده‌های طول دوره‌ی بیکاری

متغیر	طبقه	تعداد درصد سانسور متوسط سن	متوسط زمان بیکاری
کل نمونه	-	۳۹/۷ ۵۴۶ (۰/۳۲۳)	۲۶۴/۵۲۸ (۱/۹۶۵)
جنسیت	مردان	۳۸/۵ ۳۵۳ (۰/۴۲۲)	۲۲۱/۲۵۹ (۲/۵۷۴)
	زنان	۴۲/۰ ۱۹۳ (۰/۴۷۲)	۲۷۴/۴۲۳ (۲/۶۸۴)
سطح تحصیلات	ابتدایی و سیکل	۴۳/۶ ۳۹ (۱/۲۰۸)	۲۸۲/۰۹۰ (۵/۴۶۱)
	دوره متوسطه و دیپلم	۴۰/۱ ۱۳۷ (۰/۵۰۴)	۲۷۴/۱۱۵ (۲/۹۷۸)
	کاردانی و کارشناسی	۳۸/۶ ۲۷۲ (۰/۴۲۲)	۲۶۴/۰۰۰ (۲/۵۴۹)
	کارشناسی ارشد و بالاتر	۴۰/۸ ۹۸ (۰/۹۶۷)	۲۴۳/۴۵۷ (۵/۸۵۷)
گروه سنی	۱۹-۲۵	۳۳/۳ ۱۲ (۰/۵۳۷)	۱۷۳/۷۵۵ (۲۳/۶۲۴)
	۲۶-۳۲	۳۹/۱ ۶۹ (۰/۲۳۶)	۲۵۱/۴۶۹ (۶/۴۶۶)
	۳۳-۳۹	۴۰/۲ ۲۴۹ (۰/۱۲۹)	۲۶۳/۸۴۲ (۲/۵۷۶)
	۴۰-۴۶	۳۹/۴ ۱۸۰ (۰/۱۶۸)	۲۷۲/۲۰۳ (۲/۶۷۶)
	۴۷-۵۳	۴۱/۷ ۳۶ (۰/۴۱۵)	۲۷۶/۲۸۹ (۷/۳۲۹)
	وضعیت تأهل	متأهل	۳۸/۱ ۳۸۳ (۰/۳۴۸)
مجرد	۴۳/۶ ۱۶۳ (۰/۶۶۱)	۲۴۹/۶۰۵ (۴/۵۵۱)	

شکل ۱، نشان می‌دهد احتمال بیکار ماندن زنان به‌طور نسبی بیشتر از مردان است همچنین در شکل ۲، تابع خطر نشان می‌دهد مردان نسبت به زنان در معرض خطر بیشتری (خطر در اینجا یافتن شغل است) و از این رو می‌توان گفت در مدت زمان کمتری به شغل جدید دست پیدا خواهند کرد. یکی دیگر از عوامل مؤثر در مدت زمان بیکاری افزایش سطوح تحصیلی است.

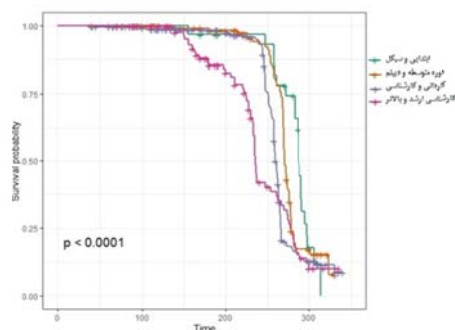


شکل ۱- نمودار تابع بقا برای عامل جنسیت

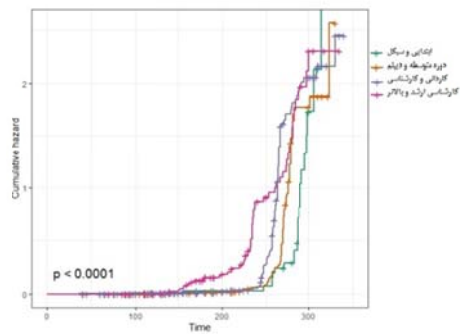


شکل ۲- نمودار تابع خطر برای عامل جنسیت

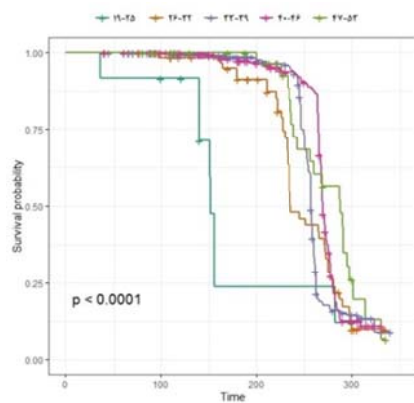
از شکل ۳ و ۴ می‌توان نتیجه گرفت، احتمال بیکار ماندن برای افراد کم‌سواد بیشتر است و کمترین احتمال بیکاری برای افراد دارای تحصیلات دانشگاهی است. به نظر می‌رسد میانگین مدت بیکاری با افزایش سن افزایش یافته است. شکل ۵ و ۶ به‌طور تقریبی این نتیجه را به ذهن متبادر می‌کند.



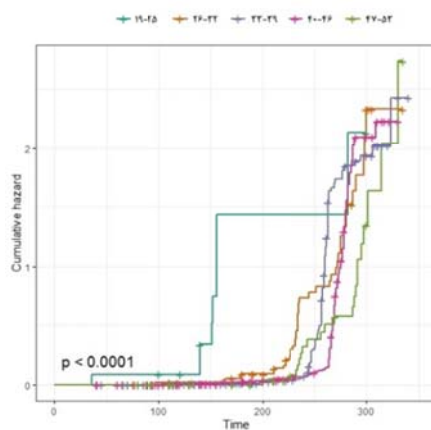
شکل ۳- نمودار تابع بقا برای عامل تحصیلات



شکل ۴- نمودار تابع خطر برای عامل تحصیلات

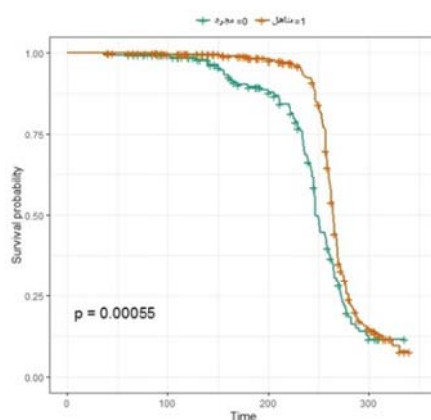


شکل ۵- نمودار تابع بقا برای عامل سن

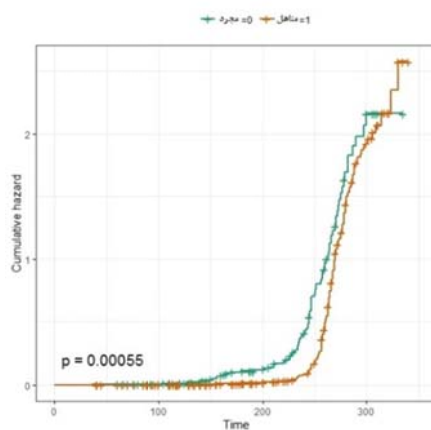


شکل ۶- نمودار تابع خطر برای عامل سن

همچنین در شکل ۷ و ۸ مشاهده می‌شود که افراد مجرد نسبت به افراد متأهل مدت زمان کمتری بیکار می‌مانند و از این رو خطر نسبی برای افراد مجرد بیشتر از افراد متأهل است.



شکل ۷- نمودار تابع بقا برای عامل وضعیت تأهل



شکل ۸- نمودار تابع خطر برای عامل وضعیت تأهل

لازم به ذکر است که روش کاپلان مهیر اساساً یک روش توصیفی است و نتایجی که برای اثر متغیرهای تبیینی در بالا ذکر شد صرفاً بر اساس شکل تابع بقا و تابع خطر تجربی است. بدیهی است که برای بررسی معنی‌دار همزمان متغیرهای تبیینی بر مدت

زمان بیکاری تنها استفاده از این روش کافی نبوده و لازم است تا با به‌کارگیری مدل کاکس به بررسی معنی داری عوامل مؤثر بر مدت زمان بیکاری پرداخته شود.

۵-۲- نتایج تجزیه و تحلیل کاکس

همان‌گونه که در بخش قبلی گفته شد روش کاپلان مه‌یر برای تحلیل توصیفی توابع بقا و خطر مورد استفاده قرار می‌گیرد و امکان سنجش تأثیر همزمان متغیرهای تبیینی بر زمان بقا را ندارد. بدین سبب از مدل مخاطرات متناسب کاکس برای تحلیل دقیق‌تر داده‌ها استفاده شده است با توجه به نتایج به‌دست آمده در جدول ۲، فرض $H_0: \beta = 0$ در سطح معنی‌داری ۹۵٪ رد می‌شود که این امر نشان از تأثیر معنی‌دار متغیرهای تبیینی بر طول مدت زمان بیکاری دارد. آماره‌های والد^۴، رتبه لگاریتمی^۵ و نسبت درست‌نمایی^۶ نیز به اتفاق معنی‌داری مدل و رد فرض صفر را تأیید می‌کنند.

جدول ۲- جدول ضرایب مدل کاکس

فاصله اطمینان ۹۵٪ برای Exp(β)	Exp(β)	sig	Z	s.e	β	
(۱/۱۳۷ ۱/۸۱۶)	۱/۴۳۷	۰/۰۰۲	۳/۰۳۳	۰/۱۱۲	۰/۳۶۳	جنسیت
(۰/۹۶۴ ۱/۰۳۳)	۰/۹۸۳	۰/۱۰۱	-۱/۶۴۰	۰/۰۱۰	-۰/۰۱۷	سن
(۰/۸۰۱ ۲/۰۶۹)	۱/۲۸۸	۰/۲۹۷	۱/۰۴۳	۰/۲۴۲	۰/۲۵۲	تحصیلات گروه ۲
(۱/۱۸۷ ۲/۹۳۱)	۱/۸۶۵	۰/۰۰۷	۲/۷۰۱	۰/۲۳۱	۰/۶۲۳	تحصیلات گروه ۳
(۱/۳۹۹ ۳/۷۷۸)	۲/۲۹۹	۰/۰۰۱	۳/۲۸۳	۰/۲۵۳	۰/۸۳۲	تحصیلات گروه ۴
(۰/۵۷۹ ۰/۹۷۱)	۰/۷۵۰	۰/۰۲۹	-۲/۱۸۷	۰/۱۳۲	-۰/۲۸۹	وضعیت تأهل

در این جدول β برآورد بردار ضرایب رگرسیون است، مثبت بودن این ضریب برای عامل جنسیت نشان می‌دهد که خطر (در اینجا یافتن شغل) برای مردان بیشتر است، $\exp(\beta)$ تغییر پیش‌بینی شده در میزان خطر برای هر واحد افزایش در متغیر تبیینی است. خطر پایان دوره بیکاری برای مردان بیکار ۴۳/۷ درصد نسبت به زنان بیکار بیشتر است. منفی بودن مقدار β برای عامل وضعیت تأهل نشان می‌دهد که خطر برای افراد متأهل کمتر از افراد مجرد است، بنا بر این با ثابت نگهداشتن سایر متغیرها، مجرد بودن خطر را ۲۵٪ افزایش می‌دهد. نرخ خطر سطوح تحصیلات گروه

۳) (کاردانی و کارشناسی) نسبت به گروه ۴ (کارشناسی ارشد و بالاتر) بیشتر است. بنا بر این نرخ خطر با سطح تحصیلات بالاتر افزایش می‌یابد. نتیجه آزمون χ^2 نیبوس که در جدول ۳ آمده است و برای ارزیابی کلی مدل کاکس به کار گرفته می‌شود، نتایج توصیفی به‌دست آمده از روش کاپلان مایر را تأیید می‌کند و نشان از معنی‌داری مدل کاکس در برازش به داده‌ها دارد.

جدول ۳- آزمون χ^2 نیبوس برای ارزیابی کلی مدل کاکس

نمره کلی		Block		Step		نمره کلی	
مجذور کای	df	مجذور کای	sig	مجذور کای	df	مجذور کای	df
۵۱/۹۴۹	۵	۵۲/۳۵۴	۰/۰۰۰	۵۲/۳۵۴	۴	۰/۰۰۰	۴

اما مدل کاکس به علت در نظر نگرفتن عوامل خطر ناشناخته لزوماً بهترین مدل را به داده‌ها برازش نمی‌دهد. بنا بر این برای کنکاش بیشتر و استخراج نتایج دقیق‌تر در ادامه مدل‌های شکنندگی به داده‌های زمان بیکاری برازش داده شده است.

۵-۳- نتایج تجزیه و تحلیل مدل شکنندگی

به‌منظور تعیین عوامل ناشناخته بر مدت زمان بیکاری، از مدل‌های شکنندگی گاوسی، گاما و t استفاده شد. نتایج حاصل از برازندگی مدل‌های شکنندگی تا حدودی مشابه به نتایج حاصل از مدل کاکس است. باید توجه داشت که معنی‌داری مؤلفه شکنندگی در این مدل‌ها نشان از تأیید مدل‌های شکنندگی و وجود عوامل خطر ناشناخته بر مدت زمان بیکاری است که توسط متغیرهای تبیینی اندازه‌گیری شده در داده‌ها تبیین نشده است. نتایج حاصل از برازش این مدل‌ها در جدول‌های ۴، ۵ و ۶ ارائه شده است.

برای مقایسه‌ی مدل‌های فوق و همچنین مقایسه آنها با نتایج حاصل از برازش مدل کاکس از ملاک آکائیک استفاده شد که نتایج آن در جدول ۷ آمده است. مقادیر ملاک آکائیک نشان می‌دهد که کمترین مقدار مربوط به مدل شکنندگی t است، از این رو مدل شکنندگی t به‌عنوان کاراترین مدل در این بررسی شناخته شد. طبق نتایج مدل شکنندگی t خطر پایان دوره بیکاری برای مردان بیکار ۳۳/۳ درصد نسبت به زنان بیکار بیشتر است. برای عامل وضعیت تأهل نرخ خطر برای افراد متأهل کمتر از

افراد مجرد است و مجرد بودن خطر (یافتن شغل جدید) را $26/73$ درصد افزایش می‌دهد. نرخ خطر سطوح تحصیلات گروه ۳ (کارדانی و کارشناسی) نسبت به گروه ۴ (کارشناسی ارشد و بالاتر) بیشتر است. بنا بر این نرخ خطر با سطح تحصیلات بالاتر افزایش می‌یابد. همچنین اثر متغیر سن بر مدت زمان بیکاری در هیچکدام از مدل‌ها تأیید نشده است.

جدول ۴- جدول ضرایب مدل شکنندگی گاوسی

فاصله اطمینان ۹۵% برای Exp(β)	Exp(β)	sig	کای دو	s.e	β	
($1/063$ $1/873$)	$1/357$	$0/014$	$6/02$	$0/124$	$0/304$	جنسیت
($0/968$ $1/001$)	$0/989$	$0/290$	$1/13$	$0/019$	$-0/012$	سن
($1/763$ $2/044$)	$1/250$	$0/380$	$0/78$	$0/251$	$0/223$	تحصیلات گروه ۲
($1/076$ $2/723$)	$1/713$	$0/023$	$5/14$	$0/238$	$0/539$	تحصیلات گروه ۳
($1/315$ $3/721$)	$2/212$	$0/003$	$8/96$	$0/266$	$0/794$	تحصیلات گروه ۴
($0/564$ $0/961$)	$0/736$	$0/024$	$5/09$	$0/136$	$-0/306$	وضعیت تأهل
		$0/026$	$27/06$			مؤلفه شکنندگی

جدول ۵- جدول ضرایب مدل شکنندگی گاما

فاصله اطمینان ۹۵% برای Exp(β)	Exp(β)	sig	کای دو	s.e	β	
($1/060$ $1/740$)	$1/364$	$0/012$	$6/28$	$0/124$	$0/311$	جنسیت
($0/967$ $1/009$)	$0/988$	$0/250$	$1/35$	$0/011$	$-0/013$	سن
($1/762$ $2/040$)	$1/248$	$0/380$	$0/77$	$0/251$	$0/221$	تحصیلات گروه ۲
($1/082$ $2/742$)	$1/723$	$0/022$	$5/26$	$0/238$	$0/544$	تحصیلات گروه ۳
($1/324$ $3/736$)	$2/228$	$0/003$	$9/12$	$0/265$	$0/799$	تحصیلات گروه ۴
($0/567$ $0/962$)	$0/739$	$0/025$	$5/04$	$0/135$	$-0/304$	وضعیت تأهل
		$0/039$	$23/69$			مؤلفه شکنندگی

جدول ۶- جدول ضرایب مدل شکنندگی t

فاصله اطمینان ۹۵٪ برای Exp(β)	Exp(β)	sig	کای دو	s.e	β	
(۱/۰۴۰ ۱/۷۰۲)	۱/۳۳۳	۰/۰۲۳	۵/۱۵	۰/۱۲۶	۰/۲۸۵	جنسیت
(۰/۹۶۷ ۱/۰۱۳)	۰/۹۹۱	۰/۴۱۰۰	۰/۶۹	۰/۰۱۱	-۰/۰۱۰	سن
(۱/۰۷۶ ۲/۰۶۰)	۱/۲۵۱	۰/۳۸۰۰	۰/۷۷	۰/۲۵۴	۰/۲۲۴	تحصیلات گروه ۲
(۱/۰۴۷ ۲/۶۷۸)	۱/۶۷۴	۰/۰۳۱۰	۴/۶۳	۰/۲۴۰	۰/۵۱۵	تحصیلات گروه ۳
(۱/۲۹۱ ۳/۷۱۹)	۲/۱۹۱	۰/۰۰۳۶	۸/۴۵	۰/۲۷۰	۰/۷۸۵	تحصیلات گروه ۴
(۰/۵۶۰ ۰/۹۵۹)	۰/۷۳۲	۰/۰۲۳۰	۵/۱۵	۰/۱۳۷	-۰/۳۱۲	وضعیت تأهل
		۰/۰۴۶۰	۳۱/۶۶			مؤلفه شکنندگی

جدول ۷- مقایسه‌ی مدل‌های شکنندگی مشترک بقا

ت	گاما	گاوسی	کاکس	معیارها
۳۳۸۸/۴۹۹	۳۳۹۰/۸۸۴	۳۳۸۹/۴۰۱	۳۴۰۰/۸۶۳	آکائیک
۱۰۲/۷	۸۷/۳۲	۹۱/۶۷	۵۰/۹۱	نسبت درست‌نمایی

۶- بحث و نتیجه‌گیری

در این مقاله با استفاده از روش‌های تحلیل بقا نظیر جداول کاپلان مه‌یر، مدل کاکس و مدل‌های شکنندگی، تأثیر متغیرهای تبیینی بر مدت زمان بیکاری بررسی شد. مدل‌های مخاطرات متناسب کاکس و مدل‌های شکنندگی، به نتایج تقریباً مشابهی دست یافتند. با استفاده از معیار آکائیک مدلی که برازش بهتری به داده‌ها داشت مدل شکنندگی t تعیین شد و نتایج حاصل از این مدل نشان داد با توجه به متغیر جنسیت، مدت بیکاری برای زنان نسبت به مردان بیشتر است. همچنین شانس اشتغال مجدد برای مردان در دوره تحلیل شده بیشتر از زنان است. افراد دارای تحصیلات کارشناسی ارشد و بالاتر شانس بیشتری در اشتغال مجدد در مقایسه با افراد با تحصیلات پایین‌تر دارند. همچنین افراد مجرد شانس اشتغال مجدد بیشتری نسبت به افراد متأهل دارند. نکته مهم اینکه معنی‌دار شدن مؤلفه‌های شکنندگی نشان از وجود عوامل ناشناخته‌ای دارد که در قالب متغیرهای تبیینی قابل اندازه‌گیری نبوده است.

۷- پیشنهادها

با توجه به آنچه در بحث و نتیجه‌گیری بیان شد، پیشنهاد می‌شود در مطالعات آتی متغیرهای تبیینی بیشتری در گردآوری اطلاعات مورد توجه قرار گیرد. با توجه به نرخ‌های متفاوت بیکاری در استان‌های مختلف کشور و همچنین ارتباط احتمالی فضایی این نرخ در استان‌های مجاور می‌توان با وارد کردن این اثر در مدل بقا کمک شایانی به بهبود برآوردهای مدل بقا نمود. بنا بر این پیشنهاد می‌شود در مطالعات آتی مدل بقای فضایی برای تحلیل و پیشگویی مدت زمان بیکاری مورد استفاده محققین این حوزه قرار گیرد.

تقدیر و تشکر

نویسندگان از داوران محترم مجله که نظرات ارزنده آن‌ها موجب بهبود مقاله خواهد شد، تشکر و قدردانی می‌نمایند.

توضیحات

1. Survival Analysis
2. Proportional Hazards Model
3. Frality Models
4. Wald test
5. Score (logrank) test
6. Likelihood ratio test

مرجع‌ها

- [۱] مترجم، ک.؛ محمدزاده، م. و آبیاری، آ. (۱۳۹۴). مدل‌بندی فضایی داده‌های بقای سانسور شده، مجله پژوهش‌های ریاضی، جلد ۱، شماره ۲، صص ۶۱-۷۰.
- [2] Ciucă, V. and Matei, M. (2010). Survival Analysis for the Unemployment Duration, Proceedings of the 5th WSEAS

- International Conference on Economy and Management Transformation (EMT'10). 1, 354-359.
- [3] Clayton, D.G. and Cuzik, J. (1985). Multivariate Generalization of the Proportional Hazards Model, *Journal of the Royal Statistical Society, Series A*, **148**, 82-117.
- [4] Cox, D.R. (1972). Regression Models and Life-Tables, *Journal of the Royal Statistical Society, Series B*, **34**, 187-220.
- [5] Cox, D.R. (1975). Partial Likelihood, *Biometrika*, **62**, 269-276.
- [6] Danacia, D.E. and Babucea, A.G. (2010). Using Survival Analysis in Economics, *Scientific Annals of the Alexandru Ioan, Cuza University of Iasi*. 439-450.
- [7] Darabă, G., Sandu, C.B. and Jaba, E. (2017). Estimation of Unemployment Duration in Botoșani County Using Survival Analysis, *Ovidius University Annals, Economic Sciences Series*, **17**, 155-161.
- [8] Lancaster, T. (1979). Econometric Methods for the Duration of Unemployment, *Econometrics*, **47**, 939-956.
- [9] Lewandowska-Gwarda, K. (2018). Geographically Weighted Regression in the Analysis of Unemployment in Poland, *International Journal of Geo-Information*, **17**.
- [10] Lin, D.Y. and Wei, L.J. (1989). The Robust Inference for the Cox Proportional Hazards Models, *Journal of the American Statistical Association*, **84**, 1074-1078.
- [11] Nonyana, J.Z. and Njuho, P.M. (2018). Modelling the Length of Time Spent in an Unemployment State in South Africa, *South African Journal of Science*, **114**, 1-7.
- [12] Vaupel, J.W., Manton, K.G. and Stallard, E. (1979). The Impact of Heterogeneity in Individual Frailty on the Dynamics of Mortality, *Demography*, **16**, 439-454.

- [13] Wienke, A. (2011). *Frality Models in Survival Analysis*, Chapman and Hall/CRC, United States of America.

فاطمه زنجیران

فوق لیسانس آمار

تهران، شهر وحیدیه، خیابان اصغر جعفری، خیابان امام خمینی، مؤسسه مهرآفرینان راهیان سلامت.
رایانشانی: n.zanjiran@gmail.com

کیومرث مترجم

دکتری آمار

تهران، خیابان جلال آل احمد، دانشگاه تربیت مدرس، گروه آمار.
رایانشانی: k.motarjem@modares.ac.ir