کاربرد توابع مفصل نامتقارن در برآورد قابلیت اعتماد: مطالعه موردی خودروهای رانا و دنا

صدیقه شمس[۱]، صدیقه برزوئی

تاریخ دریافت: ۱۴۰۱/۰۵/۰۲
تاریخ پذیرش: ۱۴۰۱/۱۲/۰۲

چکیده:

توابع مفصل ابزار مفیدی در مدل‌سازی وابستگی بین متغیرهای تصادفی هستند اما بیشتر توابع مفصل موجود متقارن هستند درحالی‌که در بسیاری از کاربردها، توابع مفصل نامتقارن موردنیاز هستند. یکی از این کاربردها، مدل‌سازی قابلیت اعتماد است که در آن، توابع مفصل نامتقارن می‌توانند وابستگی‌های دمی متفاوت را تبیین کنند و مدل بهتری ارائه دهند؛ بنابراین نظریه ساختن توابع مفصل نامتقارن که می‌تواند دامنه وسیع‌تری از داده‌ها را مدل‌سازی کند، توسعه‌یافته است. در این پژوهش ضمن مرور روش‌های ساختن توابع مفصل نامتقارن که می‌توانند وابستگی‌های دمی مختلفی را تأمین نمایند، از این توابع برای برآورد قابلیت اعتماد دوبعدی داده‌های سن و میزان استفاده خودروهای رانا و دنا، استفاده می‌شود.

واژه‌های کلیدی: تابع مفصل، تابع مفصل نامتقارن، قابلیت اعتماد، وابستگی دمی

۱ مقدمه

برآورد تابع قابلیت اعتماد برای تولیدکنندگان و مدیران امری مهم است. ممکن است قابلیت اعتماد سیستم بر اساس یک متغیر مانند سن یا میزان استفاده برآورد شود و در بسیاری از موارد مانند پژوهش [۲]، این موضوع به دو یا چند متغیر مانند سن و میزان استفاده از سیستم وابسته است . برای مثال یک خودرو فروخته‌شده ممکن است برای ۳ سال یا مسافت طی شده ۳۰ هزار کیلومتری پوشش ضمانت داشته باشد که این بدان معنی است اگر خودرو در طول مدت ضمانت با مشکلی مواجه شود، تأمین‌کننده ضمانت آن را به‌رایگان تعمیر خواهد کرد. امروزه ارائه محصولاتی باقابلیت اعتماد بالاتر برای تولیدکنندگان، اصلی اساسی و لازم به شمار می‌آید. مصرف‌کنندگان انتظار دارند که محصولات موردنیاز، برای مدت‌زمان طولانی‌تر و کارکرد بیشتر ضمانت شوند. ازاین‌رو ضمانت در بحث قابلیت اعتماد، یکی از ویژگی‌های مهم است و در نتیجه برآورد هرچه دقیق‌تر قابلیت اعتماد به‌منظور بهبود برنامه‌ریزی مالی، یک رکن اساسی است.

در دنیای واقعی ساختار وابستگی نامتقارن بین متغیرها وجود دارد. به‌طور خاص وقتی‌که الگوی وابستگی در دم‌های توزیع توأم مهم است مانند زمینه‌های محیط‌زیست، علوم مهندسی و غیره. در پژوهش [۳] روش‌هایی برای تشخیص و اندازه‌گیری در الگوی وابستگی نامتقارن ارائه‌شده است. در مدل‌سازی قابلیت اعتماد نیز، روش‌هایی که قادر به تبیین ساختار وابستگی نامتقارن هستند، کاربرد بیشتری دارند. یکی از این روش‌ها استفاده از توابع مفصل نامتقارن است. توابع مفصل ابزار مفیدی در مدل‌سازی وابستگی بین متغیرهای تصادفی هستند (رجوع شود به [۴] و [۱])؛ اما بیشتر توابع مفصل موجود متقارن هستند. درحالی‌که در بسیاری از کاربردها از جمله در قابلیت اعتماد، از توابع مفصل نامتقارن استفاده می‌شود. بنابراین تعمیم مفصل‌های متقارن به نامتقارن که می‌تواند دامنه وسیع‌تری از داده‌ها را مدل‌سازی کند، لازم است. در [۷] دو روش برای ساختن توابع مفصل نامتقارن ذکرشده است. ولی این پژوهش بر اساس روشی است که در [۶] برای ساختن توابع مفصل نامتقارن ارائه‌شده و اهمیت راهکار ارائه‌شده آن است که توابع ساخته‌شده می‌توانند وابستگی‌های دمی متفاوت را در جهات مختلف تولید کنند. در بخش دوم مقاله، مروری بر مفهوم تابع مفصل و وابستگی دمی خواهیم داشت. در بخش سوم روش‌های ساخت توابع مفصل نامتقارن ارائه می‌شود. در بخش چهارم روش مدل‌سازی قابلیت اعتماد دوبعدی بر اساس تابع مفصل نامتقارن ارائه می‌شود و روش معرفی‌شده روی داده‌های خودروهای رانا و دنا به کار می‌رود.

[۱] گروه آمار، دانشکده علوم ریاضی، دانشگاه الزهرا (نویسنده مسئول: s.shams@alzahra.ac.ir)

۲ تابع مفصل و وابستگی دمی

فرض کنید دو متغیر تصادفی X و Y با توابع توزیع $F_X(x)$ و $G_Y(y)$ به ترتیب و با توزیع توأم $H(x,y)$ باشند. در این صورت مطابق قضیه اسکلار [۵] وجود دارد تابع مفصل C به‌صورتی که:

$$H(x,y) = C(F_X(x), G_Y(y)) \quad \forall x, y \in \mathbb{R}. \tag{۱}$$

و اگر X و Y پیوسته باشند آنگاه C منحصربه‌فرد است.

تعریف ۱.۲. *(تابع مفصل) اگر $I = [۰,۱]$، آنگاه تابع دومتغیره $C : I^۲ \to I$ را یک تابع مفصل گویند هرگاه در شرایط زیر صدق کند:*

۱) *برای هر* $u, v \in I$، $C(u,۰) = C(۰,v) = ۰$.

۲) $u = C(u,۱)$ و $v = C(۱,v)$.

۳) *برای هر* $u_۱ \le u_۲$ و $v_۱ \le v_۲$ *در* I *رابطه زیر برقرار باشد*

$$C(u_۱,v_۱) - C(u_۱,v_۲) - C(u_۲,v_۱) + C(u_۲,v_۲) \ge ۰.$$

سه تابع مفصل خاص که نشان‌دهنده عدم وابستگی، بیشترین وابستگی منفی و بیشترین وابستگی مثبت هستند، به ترتیب، عبارت‌اند از

$$\Pi(u,v) = uv, \tag{۲}$$

$$M(u,v) = \min(u,v), \tag{۳}$$

$$W(u,v) = \max(u+v-۱, ۰) \tag{۴}$$

توابع مفصل $M(u,v)$ و $W(u,v)$، کران‌های پائین و بالای فرشه-هافدینگ و تابع مفصل $\Pi(u,v)$ تابع مفصل حاصل‌ضرب نامیده می‌شوند.

تعریف ۲.۲. *(خانواده مفصل ارشمیدسی) تابع مفصل ارشمیدسی C عبارت است از:*

$$C(u,v) = \psi^{[-۱]}(\psi(u) + \psi(v)) \tag{۵}$$

که در آن ψ (تابع مولد C) تابعی پیوسته، اکیداً نزولی و محدب از I به $[۰,\infty]$ و $\psi(۱) = ۰$

و تابع $\psi^{[-۱]}$ به‌صورت زیر است

$$\psi^{[-۱]}(t) = \begin{cases} \psi^{[-۱]}(t) & t \in [۰, \psi(۰)] \\ ۰ & t \ge \psi(۰). \end{cases}$$

مثال ۳.۲. (۱) *اگر برای $t \in [۰,۱]$ قرار دهیم $\psi(t) = -\ln t$ در این صورت $\psi^{[-۱]}(t) = \exp(-t)$ و تابع مفصل C تولیدشده برابر است با*

$$C(u,v) = uv$$

یعنی $\Pi(u,v)$ یک تابع مفصل ارشمیدسی است.

(۲) *اگر به ازای $t \in [۰,۱]$ داشته باشیم $\psi(t) = (۱-t)$، در این صورت $\psi^{[-۱]}(t) = \max(۱-t, ۰)$ و بنابراین تابع مفصل C تولیدشده برابر است با*

$$C(u,v) = \max(۰, u+v-۱)$$

یعنی $W(u,v)$ یک تابع مفصل ارشمیدسی است.

(۳) *(خانواده کلایتون) اگر برای $t \in [۰,۱]$ و $\theta \ge -۱$، $\theta \ne ۰$، داشته باشیم $\psi(t) = \frac{t^{-\theta}-۱}{\theta}$ در این صورت خانواده‌ای از توابع مفصل ارشمیدسی به نام خانواده کلایتون به صوت زیر تولید می‌شود:*

$$C_\theta(u,v) = \left[\max(u^{-\theta} + v^{-\theta} - ۱, ۰)\right]^{\frac{-۱}{\theta}}. \tag{۶}$$

(۴) *(خانواده گامبل- هوگارد) با ضابطه*

$$C(u,v) = \exp\left[-\left((-\ln u)^\theta + (-\ln v)^\theta\right)^{\frac{۱}{\theta}}\right] \tag{۷}$$

است. این خانواده با تابع مولد $\psi(t) = (-\ln t)^\theta$ تولید می‌شود.

تعریف ۴.۲. *(تبادل‌پذیری و تقارن) فرض کنید X و Y متغیرهای تصادفی پیوسته و هم‌توزیع با تابع توزیع توأم H و تابع مفصل C باشند، بدین ترتیب X و Y تبادل‌پذیرند اگر و تنها اگر برای تمام مقادیر $(u,v) \in I^۲$*

$$C(u,v) = C(v,u) \tag{۸}$$

هم‌چنین هرگاه متغیرهای تصادفی U و V با تابع توزیع حاشیه‌ای یکنواخت و تابع مفصل C تبادل‌پذیر باشند آنگاه تابع مفصل C متقارن است.

تعریف ۵.۲. *(ضرایب وابستگی دمی) فرض کنیدX و Y متغیرهای تصادفی پیوسته با تابع توزیع F و G باشند. ضریب وابستگی دمی بالایی به‌صورت زیر تعریف می‌شود*

$$\lambda_u = \lim_{t \to ۱^-} P[Y > G^{-۱}(t) | X > F^{-۱}(t)] \tag{۹}$$

و به‌طور مشابه ضریب وابستگی دمی پایینی به‌صورت زیر است

$$\lambda_l = \lim_{t \to ۰^+} P[Y \le G^{-۱}(t) | X \le F^{-۱}(t)]. \tag{۱۰}$$

مفهوم وابستگی دمی وضعیت یک متغیر نسبت به اطلاع از متغیر دیگر را وقتی این اطلاع مربوط به مقادیر کرانگین دامنه می‌شود، بررسی می‌کند. به دلیل پایایی این ضرایب تحت تبدیلات صعودی، تعاریف فوق را می‌توان برحسب توابع مفصل نیز به‌صورت زیر بیان کرد:

$$\lambda_u(C) = \lim_{u\to 1^-} \frac{\bar{C}(u,u)}{1-u} \qquad (۱۱)$$

$$\lambda_l(C) = \lim_{u\to 0^+} \frac{C(u,u)}{u} \qquad (۱۲)$$

در [۶] ضرایب وابستگی دمی به‌صورت زیر تعمیم داده شده است.

$$\lambda_{u,u} = \lim_{t\to 1^-} P[Y > G^{-1}(t)|X > F^{-1}(t)] \qquad (۱۳)$$

$$\lambda_{l,l} = \lim_{t\to 0^+} P[Y \le G^{-1}(t)|X \le F^{-1}(t)]. \qquad (۱۴)$$

$$\lambda_{u,l} = \lim_{t\to 1^-} P[Y > G^{-1}(t)|X \le F^{-1}(t)] \qquad (۱۵)$$

$$\lambda_{l,u} = \lim_{t\to 0^+} P[Y \le G^{-1}(t)|X > F^{-1}(t)]. \qquad (۱۶)$$

که آن‌ها به ترتیب ضرایب وابستگی دمی بالایی-بالایی، پایینی-پایینی، بالایی-پایینی و پایینی-بالایی نامیده شده‌اند.

لم ۶.۲. *هرگاه متغیرهای تصادفی U و V با تابع مفصل C تبادل‌پذیر باشند آنگاه*

$$\lambda_{u,l}(C) = \lambda_{l,u}(C)$$

اثبات در [۶].

در حالت چندمتغیره اگر فرض کنیم بردار $X' = [X_1, X_2, ..., X_d]$ دارای تابع توزیع توأم H و هر یک از مؤلفه‌های آن دارای توزیع حاشیه‌ای F_i به ازای $i = 1, ..., d$ باشند در این صورت طبق قضیه اسکلار تابع مفصل به‌صورت زیر تعریف می‌شود

$$H(x_1, ...x_d) = C(F_1(x_1), ..., F_d(x_d))$$

فرض کنید $\{X_k \le F_k^{-1}(v)\}$، $\{X_k > F_k^{-1}(1-v)\}$ و $k \in \{1, \cdots, d\}$ به ترتیب با علامت $\lhd_k$ و $\rhd_k$ نشان داده شوند. آنگاه:

$$P(\lhd_1, \cdots, \lhd_d) = C(v_1, \cdots, v_d)$$

و ضرایب تعمیم‌یافته وابستگی دمی را به‌صورت زیر تعریف می‌شوند:

$$\lambda^{l,l}_{i|k}(C) = \lim_{v\to 0^+} P(\lhd_1, \cdots, \lhd_{k-1}, \lhd_{k+1}, \cdots, \lhd_{i-1}, \lhd_i, \lhd_{i+1}, \cdots, \lhd_d|\lhd_k) \qquad (۱۷)$$

$$\lambda^{l,u}_{i|k}(C) = \lim_{v\to 0^+} P(\lhd_1, \cdots, \lhd_{k-1}, \lhd_{k+1}, \cdots, \lhd_{i-1}, \rhd_i, \lhd_{i+1}, \cdots, \lhd_d|\lhd_k) \qquad (۱۸)$$

$$\lambda^{u,l}_{i|k}(C) = \lim_{v\to 0^+} P(\lhd_1, \cdots, \lhd_{k-1}, \lhd_{k+1}, \cdots, \lhd_{i-1}, \lhd_i, \lhd_{i+1}, \cdots, \lhd_d|\rhd_k) \qquad (۱۹)$$

$$\lambda^{u,u}_{i|k}(C) = \lim_{v\to 0^+} P(\lhd_1, \cdots, \lhd_{k-1}, \lhd_{k+1}, \cdots, \lhd_{i-1}, \rhd_i, \lhd_{i+1}, \cdots, \lhd_d|\rhd_k) \qquad (۲۰)$$

برای مثال $\lambda^{l,u}_{i|k}(C)$ نمایانگر احتمال مقادیر کرانگین در جهت پایین k امین متغیر و جهت بالای i امین متغیر می‌باشد؛ به‌عبارت‌دیگر احتمال این‌که i امین متغیر مقدار بزرگ داشته باشد به شرطی که k امین متغیر مقدار کوچک دارد.

بنابراین برای تابع مفصل $C(v_1, \cdots, v_d)$ اگر v_i و v_k تبادل‌پذیر باشند، آنگاه

$$\lambda^{u,l}_{i|k}(C) = \lambda^{l,u}_{i|k}(C).$$

مثال ۷.۲. *برای توابع مفصل کلایتون و گامبل دومتغیره داریم*

$$\lambda^{u,l}_{i|k}(C) = \lambda^{l,u}_{i|k}(C)$$

ولی $\lambda^{l,l}_{i|k}(C) \ne \lambda^{u,u}_{i|k}(C)$ *می‌باشد.*

این مثال نشان می‌دهد که اگر v_i و v_k در $C(v_1, \cdots, v_d)$ تبادل‌پذیر باشند، آنگاه ضرایب دمی در جهت پایین k امین متغیر و جهت بالای i امین متغیر و در جهت بالای k امین متغیر و در جهت پایین i امین متغیر، برابر هستند. در واقع لزوماً مفصل‌های متقارن نمی‌توانند در برازاندن داده‌ها، نابرابری وابستگی دمی در دو جهت را تأمین کنند. برای رفع این مشکل، روش ساخت مفصل‌های نامتقارن به شرح زیر مطرح می‌شود.

۳ ساختن توابع مفصل نامتقارن

روش اول: تعریف می‌کنیم

$$\breve{C}_k(v_1,\cdots,v_d)=C(v_1,\cdots,v_{k-1},1,v_{k+1},\cdots,v_d) - C(v_1,\cdots,v_{k-1},1-v_k,v_{k+1},\cdots,v_d) \quad (۲۱)$$

قضیهٔ ۱.۳. *اگر $C(v_1,\cdots,v_d)$ یک تابع مفصل باشد، آنگاه $\breve{C}_k(v_1,\cdots,v_d)$ نیز تابع مفصل است.*

تابع مفصل $C(v_1,\cdots,v_d)$ در عبارت (۲۱) به‌عنوان یک تابع مفصل پایه در نظر گرفته می‌شود.
اثبات در [۶].

مثال زیر نشان می‌دهد در قضیه ۱.۳، برابری

$$C(v_1,\cdots,v_{i-1},v_i,v_{i+1},\cdots,v_{k-1},v_k,v_{k+1},\cdots,v_d) = C(v_1,\cdots,v_{i-1},v_k,v_{i+1},\cdots,v_{k-1},v_i,v_{k+1},\cdots,v_d) \quad (۲۲)$$

نمی‌تواند تبادل پذیری v_i و v_k را در $\breve{C}_k$ نتیجه دهد؛ به‌عبارت‌دیگر، اگرچه v_i و v_k در مفصل $C(v_1,\cdots,v_d)$ تبادل‌پذیر هستند اما در مفصل $\breve{C}(v_1,\cdots,v_d)$ ممکن است تبادل‌پذیر نباشند.

مثال ۲.۳. *اگر از مفصل کلایتون به‌عنوان مفصل پایه استفاده کنیم*

$$\breve{C}_k(v_1,\cdots,v_d)=\left(\sum_{\substack{i\neq k\\ i=1}}^{d} v_i^{-\theta}\right)^{-\frac{1}{\theta}} - \left[(1-v_k)^{-\theta}+\sum_{\substack{i\neq k\\ i=1}}^{d} v_i^{-\theta}-1\right]^{-\frac{1}{\theta}} \quad (۲۳)$$

که در آن $i=1,\cdots,k-1,k+1,\cdots,d$ می‌باشد.

اگر تابع مفصل پایه تابع مفصل حاصل‌ضرب باشد، یعنی برای تمام $k=1,\cdots,d$ رابطه $C(v_1,\cdots,v_d)=\prod_{k=1}^{d} v_k$ برقرار باشد، آنگاه $\breve{C}(v_1,\cdots,v_d)$ یک تابع مفصل نامتقارن نخواهد بود.
یک مزیت استفاده از $\breve{C}_k$ این است که قابلیت به دست آوردن وابستگی دمی مختلف را دارد.

در روش دوم ساخت به شرح زیر، ویژگی نامتقارن بودن برای تمام متغیرها فراهم می‌شود.
روش دوم:

برای ساخت توابع مفصل با ویژگی نامتقارن بودن در چندین متغیر از لم زیر استفاده می‌شود.

لم ۳.۳. *اگر*

$$\bar{C}(v_1,v_2,\cdots,v_d)=\sum_{k=0}^{d} p_k \breve{C}_k(v_1,v_2,\cdots,v_d) \quad (۲۴)$$

که در آن P_k پارامتر تابع مفصل $\breve{C}$ و $\sum_{k=0}^{d} p_k=1$ و $0\leq p_k\leq 1$ می‌باشد و عبارت

$$\breve{C}_0(v_1,\cdots,v_d)=C(v_1,\cdots,v_d)$$

برقرار باشد، آنگاه $\bar{C}(v_1,\cdots,v_d)$، یک تابع مفصل است.

اثبات. با استفاده از قضیه ۱.۳ اثبات کردیم $\breve{C}_k(v_1,\cdots,v_d)$ یک تابع مفصل است. با فرض این‌که $\sum_{k=0}^{d} p_k=1$ و $\bar{C}(v_1,\cdots,v_d)$ یک ترکیب خطی از $\breve{C}_k$ است، در نتیجه $\bar{C}(v_1,\cdots,v_d)$ نیز تابع مفصل است و اثبات قضیه کامل است. □

ضرایب وابستگی دمی تعریف‌شده در روابط ۱۷ تا ۲۰ به‌صورت زیر نیز می‌توانند بیان شوند

$$\lambda_{i|k}^{l,l}(C)=\lim_{v\to 0^+}\frac{C(v,\cdots,v)}{v}, \quad (۲۵)$$

$$\lambda_{i|k}^{l,u}(C)=\lim_{v\to 0^+}\frac{\breve{C}_i(v,\cdots,v)}{v}, \quad (۲۶)$$

$$\lambda_{i|k}^{u,l}(C)=\lim_{v\to 0^+}\frac{\breve{C}_k(v,\cdots,v)}{v}, \quad (۲۷)$$

$$\lambda_{i|k}^{u,u}(C)=\lim_{v\to 0^+}\frac{\breve{C}_{k,i}(v,\cdots,v)}{v} \quad (۲۸)$$

که در آن

$$\breve{C}_{k,i}(v_1,\cdots,v_d)=\breve{C}_k(v_1,\cdots,v_{i-1},1,v_{i+1},\cdots,v_d) - \breve{C}_k(v_1,\cdots,v_{i-1},1-v_i,v_{i+1},\cdots,v_d)$$

می‌باشد.
بنابراین با در نظر گرفتن تابع مفصل پایه دومتغیره $C(u,v)$ از روابط (۲۱) و (۲۴)
خواهیم داشت

$$\breve{C}_1(u,v)=v-C(1-u,v), \quad (۲۹)$$

$$\breve{C}_2(u,v)=v_1-C(u,1-v), \quad (۳۰)$$

$$\bar{C}(u,v)=p_0C(u,v)+p_1\breve{C}_1(u,v)+p_2\breve{C}_2(u,v) \quad (۳۱)$$

که در آن $p_k\geq 0$ و $\sum_{k=0}^{2} p_k=1$ است.

لم ۴.۳. *برای تابع مفصل $\bar{C}(u,v)$ تعریف‌شده در رابطه (۳۱) داریم*

$$\lambda_{l,l}(\bar{C}) = p_0\lambda_{l,l}(C) + p_1\lambda_{u,l}(C) + p_2\lambda_{l,u}(C), \qquad (۳۲)$$

$$\lambda_{l,u}(\bar{C}) = p_0\lambda_{l,u}(C) + p_1\lambda_{u,u}(C) + p_2\lambda_{l,l}(C), \qquad (۳۳)$$

$$\lambda_{u,l}(\bar{C}) = p_0\lambda_{u,l}(C) + p_1\lambda_{l,l}(C) + p_2\lambda_{u,u}(C), \qquad (۳۴)$$

$$\lambda_{u,u}(\bar{C}) = p_0\lambda_{u,u}(C) + p_1\lambda_{l,u}(C) + p_2\lambda_{u,l}(C) \qquad (۳۵)$$

اثبات با استفاده از جایگذاری $\bar{C}$ در روابط (۲۵)، (۲۶)، (۲۷) و (۲۸) به‌سادگی انجام می‌شود.

مثال ۵.۳. *برای تابع مفصل پایه کلایتون داریم:* $\lambda_{l,l}(\bar{C}) = 2^{\frac{1}{\theta}}p_0$، $\lambda_{l,u}(\bar{C}) = 2^{\frac{1}{\theta}}p_2$، $\lambda_{u,l}(\bar{C}) = 2^{\frac{1}{\theta}}p_1$ و $\lambda_{u,u}(\bar{C}) = 0$.

۴ مدل‌سازی قابلیت اعتماد با تابع مفصل نامتقارن

در این بخش توابع مفصل نامتقارن در مدل‌سازی قابلیت اعتماد نشان داده می‌شود. فرض کنید متغیرهای تصادفی، X (سن) و Y (میزان استفاده) هستند و در شرایط زیر صدق می‌کنند

شرط ۱ (وابستگی مثبت): X و Y دارای وابستگی مثبت‌اند، به عبارتی تاو کندال و ρ اسپیرمن آن‌ها بزرگ‌تر از صفر است.

شرط ۲ (وابستگی دمی): اگر تابع مفصل متناظر با توزیع احتمال توأم آن‌ها C باشد آنگاه

$$\lambda_{u,l}(C) \geqslant \lambda_{l,u}(C).$$

در واقع شرط ۱، موردنیاز است چون معمولاً میزان استفاده یک سیستم، رابطه مثبت با سن سیستم دارد.

شرط ۲، بر رابطه بین مقادیر کرانگین سن و میزان استفاده تأکید دارد. با فرض $U = F_X(X)$ و $V = G_Y(Y)$ این شرط، به‌عنوان

$$\lim_{v\to 0^+} \frac{p(V \leqslant v | U > 1-v)}{p(V > 1-v | U \leqslant v)} > 1$$

نیز می‌تواند تفسیر شود؛ یعنی احتمال این‌که متغیر تصادفی میزان استفاده زیاد باشد وقتی متغیر تصادفی سن کم است بزرگ‌تر از احتمال این است که متغیر تصادفی سن کم باشد وقتی متغیر تصادفی میزان استفاده زیاد است. بسیاری از مدل‌های توزیع دومتغیره پیشنهادشده در مدل قابلیت اعتماد دوبعدی داده‌ها را می‌توان با در نظر گرفتن شرایط فوق تبیین نمود؛ اما توزیع‌های دومتغیره‌ای وجود دارند که نمی‌توانند شرایط فوق را محقق سازند برای مثال در [۲] توزیع بقای دومتغیره را برای برازاندن به داده‌های قابلیت اعتماد دوبعدی، پیشنهاد کردند

$$S(x,y) = \exp\left\{-\left[(\frac{x}{\beta_1})^{\alpha_1\theta} + (\frac{y}{\beta_2})^{\alpha_2\theta}\right]^{\frac{1}{\theta}}\right\} \qquad (۳۶)$$

اگر قرار دهیم $u = 1-\exp\{-(\frac{x}{\beta_1})^{\alpha_1}\}$ و $v = 1-\exp\{-(\frac{y}{\beta_2})^{\alpha_2}\}$ است؛ بنابراین تابع مفصل بقا به‌صورت زیر نوشته می‌شود

$$C(u,v) = u+v-1+\exp\{-[(-\ln(1-u))^\theta+(-\ln(1-v))^\theta]^{\frac{1}{\theta}}\} \qquad (۳۷)$$

که به‌صورت زیر می‌توان بازنویسی کرد

$$C(u,v) = u+v-1+C_0(1-u, 1-v) \qquad (۳۸)$$

که در آن $C_0(u,v) = \exp\{-[(-\ln(u))^\theta + (-\ln(v))^\theta]^{\frac{1}{\theta}}\}$ که همان تابع مفصل گامبل دومتغیره است. بنابراین $\lambda_{l,u}(C) = \lambda_{u,l}(C)$ درحالی‌که عبارت $\lambda_{l,u}(C) = \lambda_{u,l}(C)$ و $\lambda_{l,l}(C) \neq \lambda_{u,u}(C)$ به این دلالت دارد که این تابع مفصل می‌تواند نابرابری وابستگی دمی را در طول دم پایین- پایین و دم بالا- بالا را تأمین می‌کند ، اما نابرابری وابستگی دمی را در طول دم پایین- بالا یا دم بالا- پایین را ایجاب نمی‌کند.

در مدل‌سازی از تابع مفصل بقا به‌صورت $C(u,v) = u+v-1+C_0(1-u, 1-v)$ استفاده می‌کنیم و در آن تابع مفصل پایه عبارت است از:

$$C_0(u,v) = \exp\{-[(-\ln(u))^\theta + (-\ln(v))^\theta]^{\frac{1}{\theta}}\}$$

با استفاده از روش دوم و رابطه (۳۱) تابع مفصل نامتقارن ساخته‌شده عبارت است از:

$$\bar{C}(u,v) = p_0C(u,v;\theta_1) + p_1\breve{C}_1(u,v;\theta) + p_1\breve{C}_2(u,v;\theta) \qquad (۳۹)$$

و به ازای تابع مفصل رابطه (۳۹) داریم:

$$\lambda_{l,l}(\bar{C}) = p_0(2-2^{\frac{1}{\theta}}), \lambda_{u,l}(C_1) = p_1(2-2^{\frac{1}{\theta}})$$

و $\lambda_{l,u}(C_1) = 0, \lambda_{u,u}(C_1) = 0$ همچنین به ازای $p_0 > p_1$، $\lambda_{u,l}(C_1) > \lambda_{l,u}(C_1)$ بنابراین شرط ۲ برقرار است.

۱.۴ مدل‌سازی بر اساس داده‌های سن و میزان استفاده از خودروهای رانا و دنا

در مورد خودرو دنا از تعداد ۱۲۶۹۹ خودرو تولیدشده در بازه زمانی موردنظر، تعداد ۴۷۲۹ مورد در محدوده ضمانت (۷۳۰ روز و ۴۰ هزار کیلومتر) به دلیل نقص فنی مراجعه داشته‌اند و تعداد ۷۹۷۰ خودرو مراجعه نکرده‌اند. در ارتباط با خودرو رانا از تعداد ۱۳۵۲۳ خودرو تولیدشده در بازه زمانی موردنظر، تعداد ۲۶۷۸ مورد در محدوده‌ی ضمانت (۷۳۰ روز و ۴۰ هزار کیلومتر) به دلیل نقص فنی مراجعه داشته‌اند و تعداد ۱۰۸۴۵ خودرو گزارش مراجعه نکرده‌اند.
آمار توصیفی از اطلاعات به‌دست‌آمده در جدول (۱) نشان داده است. برای هر یک از متغیرها به تفکیک نوع خودرو توزیع حاشیه‌ای برآورد شده و نتایج در جدول (۲) نشان داده شده است. برای داده‌های هر یک از خودروها با استفاده از آزمون نیکویی برازش اندرسون دارلینگ، توابع مفصل مختلف برازش داده شد و نتایج زیر در جدول (۳) نشان داده شده است. با توجه به این‌که مقادیر احتمال در جدول (۳) برای هر دو نوع خودرو رانا و دنا، فرض صفر در سطح ۰۵۰۰ رد می‌شود یعنی برای هر دو نوع خودرو هیچ‌یک از توابع مفصل متقارن معرفی‌شده مناسب نیستند، در نتیجه توابع مفصل نامتقارن را از روی توابع مفصل متقارن مذکور می‌سازیم.

برای خودرو دنا توابع مفصل نامتقارن با استفاده از مدل‌های ذکرشده در بخش ۴ با توابع مفصل پایه کلایتون، فرانک و گامبل ساخته‌شده‌اند. برآورد پارامترها انجام‌شده و بر اساس آزمون نیکویی برازش اندرسون دارلینگ و مقدار آکائیکه امکان مقایسه فراهم‌شده است. نتایج محاسبات در جدول (۴) نشان داده شده است.

برای خودرو دنا توابع مفصل نامتقارن با استفاده از مدل‌های ذکرشده در بخش ۴ با توابع مفصل پایه کلایتون، فرانک و گامبل ساخته‌شده‌اند. برآورد پارامترها انجام‌شده و بر اساس آزمون نیکویی برازش اندرسون دارلینگ و مقدار آکائیکه امکان مقایسه فراهم‌شده است. نتایج محاسبات در جدول (۵) نشان داده شده است.

با توجه به مقادیر آکائیکه در جداول (۴) و (۵)، اگر از تابع مفصل نامتقارن با تابع مفصل پایه گامبل برای خودرو دنا و با تابع مفصل فرانک برای خودرو رانا، برای محاسبه تابع بقا بر مبنای ۷۳۰ روز و ۴۰۰۰۰ کیلومتر، استفاده کنیم، نتایج جدول (۶) به دست می‌آید.

جدول ۱. آمار توصیفی مربوط به داده‌های خودرو دنا و رانا

نوع خودرو	متغیر	میانگین	انحراف معیار	چولگی	برجستگی
خودرو دنا	سن	۱۰۰۷/۱۶۶	۸۵۹۳۳/۳۵	−۰۵۲۷۵۰۱/۰	۸۶۲۹۱۴/۱
	میزان استفاده	۷۲/۸۵۷۴	۵۴۳۴۱۵/۴	۰۹۰۲۰۵/۱	۵۴۳۴۱۵/۴
خودرو رانا	سن	۱۹۹۴/۱۴۲	۳۱۱۲۱/۳۰	۶۶۸۶۰۴۸/۰	۶۳۵۱۴۶/۲
	میزان استفاده	۵۶۵/۹۷۳۹	۹۱۴/۵۵۶۲	۳۴۳۲۷۹/۱	۸۰۰۹۱۴/۵

جدول ۲. برآورد توزیع‌های حاشیه‌ای مربوط به داده‌های خودرو دنا و رانا

نوع خودرو	متغیر	توزیع برآورد شده	برآورد پارامترها		
			شکل	مقیاس	مکان
خودرو رانا	سن	مقادیر غایی تعمیم‌یافته	۰۹۱۷۵۱۴/۰	۶۲۳۵/۲۲	۹۸۲۵/۱۲۶
			(۰۲۳۹۹۰۰۲/۰)	(۴۳۲۹۲۷۸/۰)	(۵۴۵۴۴۹۱/۰)
	میزان استفاده	مقادیر غایی تعمیم‌یافته	۰۴۹۴۸۱۱۹/۰	۵۵۷/۴۱۴۴	۳۸۸/۱۷۴۰۶
			(۰۱۴۹۵۹۵۰۵/۰)	(۰۶۲۰۳/۹۶)	(۰۵۰۶۵/۹۱)
خودرو دنا	سن	مقادیر غایی تعمیم‌یافته	۶۹۴/۱۵۵	۳۱۸۳۵/۳۷	−۴۱۱۵۳۹۶/۰
			(۰۱۰۶۱۱۲۴/۰)	(۴۶۴۰۸۴۶/۰)	(۵۹۸۷۷۴۸/۰)
	میزان استفاده	مقادیر غایی تعمیم‌یافته	۵۶۹۳۵۷۴/۰	۱۰۴۹۷۹۰۹۸	
			(۰۱۲۷۵۱۲۶/۰)		

جدول ۳. برازش تابع مفصل

نوع خودرو	تابع مفصل	آماره آزمون	مقدار احتمال	معیار آکائیکه
رانا	کلایتون	۷۶۱۳۲/۲۱	۰۴۵۴۵۴۵۵/۰	−۲۱۵/۱۵۳
	گامبل	۷۶۱۰۶/۲۱	۰۴۵۴۶۳/۰	−۵۱۰۸/۱۵۲
	فرانک	۷۶۱۳۲/۲۱	۰۴۵۴۵۴۵۵/۰	−۴۲۶۳/۱۹۵
دنا	کلایتون	۹۷۴۸۵۵/۸	۰۴۵۴۵۴۵۵/۰	−۶۲۱۲/۳۴۸
	گامبل	۹۷۴۸۰۷/۸	۰۴۵۴۵۳۴۸/۰	−۱۳۰۰/۳۰۲
	فرانک	۹۷۴۸۵۵/۸	۰۴۵۴۵۴۵۵/۰	−۳۳۸۱/۳۹۲

جدول ۴. پارامترها و عملکرد تابع مفصل نامتقارن برای خودرو دنا

تابع مفصل پایه	آماره اندرسون–دارلینگ	مقدار احتمال	آکائیکه	$p_۱$	$p_۰$
کلایتون	۸۹۸/۱۶	۰۵۶۸۰۳۵۷/۰	−۳۷۲۹/۲۷۱	۲/۰	۵/۰
فرانک	۱۳۵۸۴/۱۴	۱۳۷۵۷۴/۰	−۲۸۷/۴۰۰	۲/۰	۵/۰
گامبل	۹۴۹۴۷/۱۷	۰۹۵۷۶۴۳۵/۰	−۰۹۹۹/۴۰۱	۲/۰	۵/۰

جدول ۵. پارامترها و عملکرد تابع مفصل نامتقارن برای خودرو رانا

توابع مفصل پایه	آماره اندرسون–دارلینگ	مقدار احتمال	آکائیکه	$p_۱$	$p_۰$
کلایتون	۶۱۵۷۲۸/۸	۴۷۳۴۷۴۴/۰	−۵۰۷۷/۱۹۳	۲/۰	۵/۰
فرانک	۶۱۵۷۲۸/۸	۴۷۳۴۷۴۵/۰	−۵۰۷۷/۱۹۷	۲/۰	۵/۰
گامبل	۱۹۷۲۷/۱۰	۳۳۴۷۵۲۸/۰	−۰۱۶۶/۱۶۳	۲/۰	۵/۰

جدول ۶. برآورد تابع بقا بر مبنای سن ۷۳۰ روز و میزان استفاده ۴۰۰۰۰ کیلومتر

نوع خودرو	برآورد تابع بقا
دنا	۶۷۶۰۸۴۷/۰ ۷
رانا	۸۱۰۶۹۷۲/۰

۵ بحث و نتیجه‌گیری

هنگامی‌که تابع قابلیت به دو یا چند متغیر وابسته باشد لازم است تابع بقای چندمتغیره با خطای کمتر و مقید به شرایط موجود، برآورد شود. یکی از روش‌های برآورد تابع بقا در این حالت روش ذکرشده در بخش سوم این مقاله بود که علاوه بر نامتقارن بودن و تبیین ضرایب وابستگی دمی متفاوت، قادر بود در صورت لزوم شرایط خاص مسئله را نیز تأمین نماید.

نوآوری این پژوهش در مدل‌سازی وابستگی توأم سن و میزان استفاده خودروهای رانا و دنا، با استفاده از توابع مفصل نامتقارن ساخته‌شده است. در مطالعه موردی با استفاده از اطلاعات مربوط به خودروهای دنا و رانا دریافتیم که استفاده از تابع مفصل بقای نامتقارن با توابع مفصل پایه‌ی، به ترتیب گامبل و فرانک، مناسب بوده و منجر به برآورد تابع بقای مندرج در جدول (۶) شد.

مراجع

[1] Joe, H. (1996). *Multivariate models and dependence concepts*, Monographs on statistics and applied probability (Vol. 73), Chapman and Hall, London.

[2] Jung, M., and Bai, D. S. (2007). Analysis of field data under two-dimensional warranty. *Reliability Engineering and System Safety*, **92(2)**, 135–143.

[3] Hua, L., Polansky, A., and Pramanik, P. (2019). Assessing bivariate tail non-exchangeable dependence. *Statistics & Probability Letters*, **155(2)**, 108556

[4] Nelsen, R. (2006), *An Introduction to Copulas*, Springer, New York.

[5] Sklar, A. (1959). Fonctions de répartition àn dimensions et leurs marges. *Publications de l'Institut de Statistique de l'Université de Paris*, **8**, 229–231.

[6] Wu, Sh. (2014). Construction of asymmetric copulas and its application in two-dimensional reliability modelling. *European Journal of Operational Research*, **238**, 476–485.

[7] Zhang, L., and Singh, V. P., (2019). *Copulas and their applications in water resources engineering*, Cambridge University Press.

Asymmetric Copula Functions Applicaton in Reliablity Estimation: Case study for Rana and Dena cars

Sedigheh Shams[1] Sedigheh Borzoee

Abstract:

Copula functions are useful tools in modelling the dependence between random variables, but most existing copula functions are symmetric, while in many applications, asymmetric copula functions are required. One of these applications is reliability modeling, where asymmetric copula functions can explain different tail dependencies and provide a better model. Therefore, the theory of constructing asymmetric copula functions that can model a wider range of data has been developed. In this research, while reviewing the methods of constructing asymmetric copula functions that can provide various tail dependencies. These functions are used to estimate the two dimensional reliability of data on the age and usage of Rana and Dena cars.

Keywords: Asymmetric copula function, Copula function, Reliability, Tail dependence.

[1]**Department of Statistics, Faculty of Mathematical Sciences, Alzahra University**