

باسمه تعالی



مجموعه مقالات

دومین کارگاه آموزشی

نظریه مفصل و انواع وابستگی ها

قطب علمی داده های ترتیبی و فضایی دانشگاه فردوسی مشهد

با همکاری

گروه آمار دانشگاه فردوسی مشهد

و انجمن آمار ایران

آبان ماه ۱۳۹۱

دبیر کارگاه:

آقای دکتر محمد امینی عضو هیأت علمی گروه آمار دانشگاه فردوسی مشهد
اعضای کمیته علمی (به ترتیب حروف الفبا):

- دکتر محمد امینی، دانشگاه فردوسی مشهد
- دکتر هادی جباری نوقایی، دانشگاه فردوسی مشهد
- دکتر علی دولتی، دانشگاه یزد
- دکتر محمدحسین علامت ساز، دانشگاه اصفهان

اعضای کمیته اجرایی (به ترتیب حروف الفبا):

- دکتر جعفر احمدی، دانشگاه فردوسی مشهد
- دکتر محمد امینی، دانشگاه فردوسی مشهد
- دکتر هادی جباری نوقایی، دانشگاه فردوسی مشهد
- دکتر غلامرضا محتشمی برزادران، دانشگاه فردوسی مشهد

اعضای کمیته دانشجویی (به ترتیب حروف الفبا):

- حامد احمدزاده، دانشگاه فردوسی مشهد
- خیراله اخلی، دانشگاه فردوسی مشهد
- حسینعلی محتشمی برزادران، دانشگاه فردوسی مشهد

محورها:

- تاریخچه تابع مفصل
- روش‌های ساخت مفصل
- توابع مفصل و مفاهیم وابستگی
- مدل‌سازی با استفاده از توابع مفصل
- کاربردهای توابع مفصل
- تابع مفصل و آمار فضایی
- ارزیابی مدل‌سازی با تابع مفصل به کمک نرم افزار

پیشگفتار

به منظور تبادل نظر اعضای هیأت علمی و ارتقای سطح علمی دانشجویان تحصیلات تکمیلی دانشگاه ها و مراکز آموزش عالی کشور و کاربران علمی تابع مفصل و انواع وابستگی ها، دومین کارگاه آموزشی تابع مفصل و انواع وابستگی ها روزهای ۹ و ۱۰ خردادماه سال ۱۳۹۱ توسط قطب علمی داده های ترتیبی و فضایی دانشگاه فردوسی مشهد و با همکاری گروه های آمار دانشگاه های فردوسی مشهد، اصفهان، تربیت مدرس، تهران و رازی کرمانشاه و انجمن آمار ایران در محل دانشکده علوم ریاضی دانشگاه فردوسی مشهد برگزار گردید. در این کارگاه دو روزه بیش از ۵۰ نفر از اعضای هیأت علمی، دانشجویان تحصیلات تکمیلی و کاربران تابع مفصل و انواع وابستگی ها از دانشگاه ها و مراکز علمی شرکت نمودند. پس از فراخوان کارگاه، کمیته علمی اقدام به بررسی مقالات دریافتی نموده و تعداد چهارده مقاله را به صورت سخنرانی مورد موافقت قرار داد. همچنین تعداد پنج سخنرانی عمومی که شامل بحث های کاربردی تابع مفصل بود در سه نیم روز برگزاری کارگاه ارائه گردید. سخنرانان عمومی این کارگاه به شرح زیر بودند:

- آقای دکتر ابوالقاسم بزرگ نیا، استاد گروه آمار دانشگاه فردوسی مشهد
- آقای دکتر محمد امینی، استاد گروه آمار دانشگاه فردوسی مشهد
- آقای دکتر غلامرضا محتشمی برزادران، استاد گروه آمار دانشگاه فردوسی مشهد
- آقای دکتر علی دولتی، استادیار گروه آمار دانشگاه یزد
- آقای دکتر هادی جباری نوقایی، استادیار گروه آمار دانشگاه فردوسی مشهد

متن سخنرانی های عمومی این کارگاه در پایگاه اطلاعاتی قطب علمی داده های ترتیبی و فضایی دانشگاه فردوسی مشهد به نشانی <http://osdce.um.ac.ir> قابل دسترسی است. متن کامل تعداد هفت مقاله ی فارسی و شش مقاله ی انگلیسی پذیرفته شده پس از تأیید کمیته علمی در این مجموعه تدوین شده است که بخش اول شامل مقالات انگلیسی و بخش دوم شامل مقالات فارسی است.

برگزاری این کارگاه علاوه بر تلاش خستگی ناپذیر کمیته های علمی و اجرایی، مرهون پشتیبانی گروه آمار دانشگاه فردوسی مشهد و مسئولین محترم این دانشگاه بوده است. بدین وسیله مراتب سپاس و قدرانی قطب علمی داده های ترتیبی و فضایی گروه آمار دانشگاه فردوسی مشهد را ابراز می نمایم.

با آرزوی توفیق الهی

قطب علمی داده های ترتیبی و فضایی

فهرست

مروری بر روابط نرم‌های مثلثی و مفصل

احمدزاده، ح. و امینی، م. ۱

خوشه‌بندی مدل محور با استفاده از توابع مفصل

افسریه، ا. ۷

تحلیل داده‌های فضایی با استفاده از توابع مفصل

امیدی، م. و محمدزاده، م. ۱۲

بررسی ساختار وابستگی خانواده مفصل‌های مقدار فرین

امینی، م. و زنده باد، س.و. ۲۵

بررسی وابستگی دمی بین حجم تجارت و سود بازگشتی بورس اوراق بهادار با استفاده از تابع مفصل ترکیبی

حیدری بهنوییه، ا.، کارگر برزی، ع. و پرهام، غ.ع. ۳۳

مروری بر روش‌های برآورد مفصل و پارامترهای آن

راسخی نژاد، س.، جمهوری، س. و نیلی ثانی، ح.ر. ۳۸

آیا اطلاعات نامتقارن در بازار محصولات بیمه ایران وجود دارد؟

داریوش همدانی ح. و صفری کته سری، ه. ۴۹

مروری بر روابط نرم‌های مثلثی و مفصل

احمدزاده، ح. ا. و امینی، م.

گروه آمار دانشگاه فردوسی مشهد

چکیده

در این مقاله نرم‌های مثلثی و شباهت آن‌ها با تابع مفصل بیان می‌شود. سپس کلاس بزرگی از نرم‌های مثلثی موسوم به نرم‌های مثلثی ارشمیدسی معرفی شده و به ارتباط بین نرم‌های مثلثی ارشمیدسی و مفصل ارشمیدسی پرداخته می‌شود و در پایان نکاتی در مورد تعمیم نرم‌های مثلثی بیان شده است. **واژه‌های کلیدی:** نرم‌های مثلثی، مفصل.

۱ مفصل و نرم‌های مثلثی

X_1, X_2, \dots, X_k متغیر تصادفی k F_1, F_2, \dots, F_k و توزیع های حاشیه ای آن‌ها را در نظر بگیرید آن‌گاه توزیع توام آن‌ها بدین صورت خواهد بود:

$$F(y_1, y_2, \dots, y_k) = C(F_1(y_1), \dots, F_k(y_k))$$

اکنون k مجموعه‌ی فازی A_1, \dots, A_k و تابع عضویت آن‌ها یعنی $m_{A_1}(x), \dots, m_{A_k}(x)$ را در نظر بگیرید حال می‌خواهیم تابع عضویت $\cap_{i=1}^k A_i$ را بدست آوریم. زاده این تابع عضویت را بدین صورت پیشنهاد داد:

$$m_{\cap_{i=1}^k A_i}(x) = \min(m_{A_1}(x), \dots, m_{A_k}(x))$$

که \min برخی از اصول را بوشش نمی‌دهد. بدین منظور تابع عضویت اشتراک با به کارگیری T -نرم بدین فرم نمایش داده می‌شود.

$$m_{\cap_{i=1}^k A_i}(x) = T(m_{A_1}(x), \dots, m_{A_k}(x))$$

نرم‌های مثلثی نخستین بار توسط منجر معرفی شدند که به صورت زیر تعریف می‌شوند. تابع دو متغیره $T(x, y) : I \times I \rightarrow I$ را یک T -نرم گوئیم اگر در شرایط زیر صدق کند:

$$T(x, 1) = x, T(x, 0) = T(0, x) = 0 \quad (1)$$

(۲) یکنوایی: $x_1 \leq x_2, y_1 \leq y_2 \Rightarrow T(x_1, y_1) \leq T(x_2, y_2)$

(۳) جا به جایی: $T(x, y) = T(y, x)$

(۴) شرکت پذیری: $T(x, T(y, z)) = T(T(x, y), z)$

- تعریف: تابع دو متغیره $S(x, y) : I \times I \rightarrow I$ را یک T هم نرم یا S -نرم گوئیم:

(۱) $S(x, 1) = 1, S(0, x) = S(x, 0) = x$

(۲) یکنوایی: $x_1 \leq x_2, y_1 \leq y_2 \Rightarrow S(x_1, y_1) \leq S(x_2, y_2)$

(۳) جا به جایی: $S(x, y) = S(y, x)$

(۴) شرکت پذیری: $S(x, S(y, z)) = S(S(x, y), z)$

با توجه به تعاریف فوق T -نرم و S -نرم دوگان هم می باشند. از طرف دیگر برای هر T -نرم می توان یک و فقط یک S -نرم تعریف کرد به قسمی که

$$T(x, y) = 1 - S(1 - x, 1 - y)$$

اگر T یک T -نرم باشد آن گاه

$$Z \leq T \leq \text{Min}$$

که Z عملگر دراستیک (ضعیف ترین T -نرم) می باشد

$$Z(x, y) = \begin{cases} x & y = 1 \\ y & x = 1 \\ 0 & \text{در غیر اینصورت} \end{cases}$$

۱.۱ - نرم های ارشمیدسی

[۱] T را یک T -نرم ارشمیدسی گوئند اگر برای هر $u, v \in (0, 1)$ n ای متعلق به N موجود باشد به قسمی که

$$T(u, \dots, u) < v$$

۲ مفصل

تابع دو متغیره $C : I \times I \rightarrow I$ را یک مفصل گوئند اگر در شرایط زیر صدق کند:

(۱) اگر برای هر $u, v \in I$: $C(u, 1) = u, C(1, v) = v$ و $C(u, 0) = C(0, v) = 0$

(۲) برای هر u_1, u_2, v_1, v_2 که $u_1 < u_2$ و $v_1 < v_2$ داشته باشیم

$$C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0.$$

اگر C یک مفصل باشد آن گاه در شرط لیب شیتس صدق می کند. همانگونه که می دانید $W < C < Min$ که $W(x, y) = \max(x + y - 1, 0)$ نرم T یک مفصل می باشد اگر و تنها اگر در شرط لیب شیتس صدق کند.

۱.۲ برخی مفصل ها

مفصل هایی که متقارن هستند اما شرکت پذیر نمی باشند:

$$C_\alpha(x, y) = \frac{1}{1+\alpha} \max(x + y - 1 + \alpha - \alpha|x - y|, 0), \quad 0 < \alpha < 1. \quad (1)$$

$$C_\alpha(x, y) = (1 - \alpha) \max(x + y - 1, 0) + \alpha \min(x, y), \quad 0 < \alpha < 1. \quad (2)$$

(۳) $C_\alpha(x, y) = \frac{\alpha^2(1-\alpha)}{1-\alpha^2} \max(x + y - 1, 0) + (1 - \alpha^2)xy + \frac{\alpha^2(1+\alpha)}{1-\alpha^2} \min(x, y)$ یک مفصل T نرم می باشد اگر و تنها اگر شرکت پذیر باشد. فرض کنید T یک نرم باشد در این صورت

$$d_T = \begin{cases} S_T(x, y) - T(x, y) & x \neq y \\ 0 & x = y \end{cases}$$

که S_T هم نرم T می باشد را در نظر می گیریم. اگر T -نرم T مفصل باشد آن گاه d_T متر می باشد. عکس قضیه فوق برقرار نیست.

یعنی اگر T یک نرم T باشد و d_T نیز یک متر باشد در این صورت لزوماً T مفصل نمی باشد. اگر C یک مفصل (T) باشد در این صورت

$$C_\phi(x, y) = \phi(C(\phi^{-1}(x), \phi^{-1}(y)))$$

نیز یک مفصل (T) نرم می باشد. [۴] نرم T مفصل نمی باشد اما d_T متر می باشد.

$$t(x) = \begin{cases} 1 - \log_3 x & x \in [0, \frac{1}{3}] \\ 3(1 - x) & x \in (\frac{1}{3}, 1] \end{cases}$$

و نرم تولید شده توسط این مولد جمعی به صورت زیر می باشد:

$$T(x, y) = \begin{cases} \frac{xy}{3} & (x, y) \in [0, \frac{1}{3}]^2 \\ x3^{-2(1-y)} & (x, y) \in [0, \frac{1}{3}] \times (\frac{1}{3}, 1] \\ y3^{-2(1-x)} & (x, y) \in (\frac{1}{3}, 1] \times [0, \frac{1}{3}] \\ 3^3x+3y-5 & (x, y) \in (\frac{1}{3}, 1]^2, x+y < \frac{4}{3} \\ x+y-1 & (x, y) \in (\frac{1}{3}, 1]^2, x+y \geq \frac{4}{3} \end{cases}$$

که با توجه به این که مولد جمعی t محدب نمی باشد بنابراین نرم ارشمیدسی T مفصل نمی باشد. مفصل C ارشمیدسی است اگر و تنها اگر شرکت پذیر باشد و $\delta(x) < x$. یک نرم ارشمیدسی مفصل ارشمیدسی می باشد اگر و تنها اگر مولد ϕ محدب باشد. یک نرم ارشمیدسی است اگر و تنها اگر $\forall 0 < x < 1$ $\lim_{n \rightarrow \infty} \delta_T^n(x) = 0$

که $\delta_T(x) = T(x, x)$ با توجه بدین قضیه کلاس T نرم‌های ارشمیدسی یک کلاس بزرگ می‌باشد که با توجه به محدب بودن مولد مفصل ارشمیدسی حاصل می‌شود. به عنوان مثال T نرم هاماخر یعنی $\frac{xy}{1-\alpha(1-x)(1-y)}$ یک T نرم ارشمیدسی می‌باشد با مولد $\log \frac{1-\alpha+\alpha x}{x}$ می‌باشد که به ازای $\alpha \in [-1, 1]$ مفصل ارشمیدسی حاصل می‌شود که موسوم به مفصل AMH می‌باشد.

۲.۲ T نرم‌ها و نکاتی در مورد تعمیم آن‌ها:

(۱) تعمیم T نرم هاماخر
 ما مفصل AMH را بر حسب مفصل بقا بر اساس کار دورانت [۳] تعمیم دادیم. در این جا نیز بر اساس T نرم و هم‌نرم این مفصل را تعمیم داد یعنی:

$$\frac{T_1(x, y)}{1 - \alpha(1 - S_1(x, y))}.$$

تنها باید شرط شرکت‌پذیری را بررسی نمود.

تعمیم T نرم دوبوا و براد:

$$T_\alpha = \frac{xy}{\max(x, y, \alpha)}$$

برای این مفصل تعمیمی بدین صورت بر اساس روش دورانت پیشنهاد می‌شود:

$$\frac{\min(x, y)f(\max(x, y))}{f(\max(x, y, \alpha))}$$

اگر $f(x) = x$ آن‌گاه T نرم دوبوا و براد حاصل می‌شود.

اگر $f(x) = 1$ باشد T نرم Min .

در حالت کلی باید شرط شرکت‌پذیری این تعمیم بررسی شود.

البته در مقاله‌ای یک T نرم بدین فرم معرفی شده است:

$$T_{\alpha, \beta}(x, y) = \frac{\max(xy - \beta(1-x)(1-y), 0)}{1 - \alpha(1-x)(1-y)}$$

که این T نرم را می‌توان بدین صورت تعمیم داد:

$$T_{\alpha, \beta}(x, y) = \frac{\max(T_1(x, y) - \beta(1 - S_1(x, y)), 0)}{1 - \alpha(1 - S_1(x, y))}.$$

که باید شرط شرکت‌پذیری بررسی شود. این سؤال مطرح می‌شود که اگر T نرم $superadditive$ باشد در این صورت آیا T مفصل است؟ [۲] فرض کنید t یک مولد جمعی باشد

$$t(x) = \begin{cases} 1 - \log_2 x & x \in [0, \frac{1}{2}] \\ 4(1-x) & \text{ow} \end{cases}$$

در این صورت

$$T(x, y) = \begin{cases} \frac{xy}{\sqrt[4]{1-xy}} & (x, y) \in [0, \frac{1}{4}]^2, \\ x\sqrt[4]{1-y} & (x, y) \in [0, \frac{1}{4}] \times (\frac{1}{4}, 1], \\ y\sqrt[4]{1-x} & (x, y) \in (\frac{1}{4}, 1] \times [0, \frac{1}{4}], \\ \sqrt[4]{4x+4y-4} & (x, y) \in (\frac{1}{4}, 1]^2, x+y \leq \frac{3}{4}, \\ x+y-1 & (x, y) \in (\frac{1}{4}, 1]^2, x+y \geq \frac{3}{4}, \end{cases}$$

که T superadditive می باشد. که با توجه به این که مولد t محدب نمی باشد لذا T نرم ارشمیدسی T مفصل نخواهد بود. عملگرهایی که در کلیه شرایط T نرم به جز تقارن صدق می کند:

$$T(x, y) = \begin{cases} x+y-1, & x+y > \frac{3}{4}, \\ \frac{1}{4}, & x+y = \frac{3}{4} \text{ and } x < y, \\ Z(x, y), & o.w. \end{cases}$$

و

$$T(x, y) = \begin{cases} 0, & (x, y) \in [0, \frac{1}{4}] \times [0, 1], \\ \min(x, y), & otherwise. \\ Z(x, y), & o.w. \end{cases}$$

عملگرهایی که شرکت پذیر می باشند اما در شرط یکنوایی و تقارن صدق نمی کنند: الف) $T(x, y) = x$ (ب) به ازای $\frac{1}{4} \leq \alpha \leq 1$

$$T(x, y) = \begin{cases} \min[\max(x, y), \alpha], & 0 \leq x \leq \alpha, \\ x, & \alpha \leq x \leq 1-\alpha, \\ \max[\min(x, y), 1-\alpha], & 1-\alpha \leq x \leq 1, \end{cases}$$

بحث و نتیجه گیری

بسیاری از مفصل ها را می توان بر اساس یک مفصل و مفصل بقای آن تعمیم داد. در این مقاله نیز می توان بسیاری از نرم های مثلثی را بر اساس یک نرم مثلثی و هم نرم آن تعمیم داد. سپس نرم مثلثی بدست آمده را از این منظر مورد بررسی قرار داد که آیا برخی از اصول مانند اصل شمول و اصل طرد و خاصیت جبران پذیری را پوشش می دهد؟

مراجع

- [1] Alsina, A. (2006), Associative functions: Traingular norms and copulas, World Scientific.
- [2] Durante, F. (2005), Solution of an open problem for associative copulas, *Fuzzy Sets and Systems*, **152**, 411-415.

- [3] Durante, F. (2007), A new family of bivariate copulas, *comptes Rendus Mathematique*, **344**, 195–198.
- [4] Kelment, E., Mesiar, R. and Pap, E. (2000), Traingular norms, Springer.
- [5] Nelsen, R. (2006), An introduction to Copulas, Springer.
- [6] Ouyang, Y. (2012), A note on metrics induced by copulas, *Fuzzy Sets and Systems*, **191**, 122–125.

خوشه‌بندی مدل محور با استفاده از توابع مفصل

افسریه، ا.ا.

گروه آمار دانشگاه یزد

چکیده

خوشه بندی، یکی از روش‌های تحلیل داده‌های چند متغیره است. یکی از روش‌های خوشه‌بندی، یافتن الگوی احتمالی مناسبی برای داده‌هاست. به این راهبرد، خوشه‌بندی مدل محور گفته می‌شود. در عمل معمولاً آمیخته‌ای از توزیع‌های نرمال چندمتغیره به عنوان توزیع خوشه‌ها استفاده می‌شود. در این مقاله راهبرد استفاده از توزیع‌های آمیخته دیگر و توابع مفصل برای خوشه‌بندی مورد بحث قرار می‌گیرد. **واژه‌های کلیدی:** خوشه بندی مدل محور، توزیع‌های آمیخته، مفصل.

۱ مقدمه

خوشه به مجموعه‌ای از داده‌ها گفته می‌شود که به هم شباهت داشته باشند. در خوشه‌بندی سعی می‌شود تا داده‌ها به خوشه‌هایی تقسیم شوند که شباهت بین داده‌های درون هر خوشه حداکثر و شباهت بین داده‌های درون خوشه‌های متفاوت حداقل شود. تحلیل خوشه‌ای یک روش قدیمی است که در آن هیچ فرضی در مورد تعداد گروه‌ها یا ساختمان آن‌ها در نظر گرفته نمی‌شود. تحلیل خوشه‌ای یکی از پرکاربردترین روش‌ها برای آنالیز داده‌های چندمتغیره است و در آمار خوشه‌بندی به عنوان یک روش برای برازش یک مدل احتمالی مناسب در مکانیزم تولید داده‌هاست. بنابر این یکی از اهداف تحلیل خوشه‌ای، یافتن الگوی مناسب (یک مدل) برای داده‌هاست. یکی از روش‌های خوشه‌بندی، استفاده از مدل‌های آمیخته است. به این نوع خوشه‌بندی، خوشه‌بندی مدل محور گفته می‌شود. برای جزئیات این موضوع می‌توان به مقاله رفتی (۱۹۹۳) اشاره کرد. در این روش داده‌های $X = (x_1, \dots, x_p)$ به عنوان مشاهده‌ای از یک تابع چگالی آمیخته به فرم

$$\mathbf{h}(X, \theta) = \sum_{k=1}^K \pi_k f_k(X, \theta_k)$$

در نظر گرفته می‌شود که در آن f_k تابع چگالی توأم مشاهدات خوشه k ام و θ_k پارامتر آن است. خوشه k ام دارای یک احتمال پیشین π_k است، که بیانگر احتمال تعلق مشاهده x به این خوشه است و $\sum_{k=1}^K \pi_k = 1$.

مربوط به $f_k(\cdot; \theta_k)$ که بر روی R^{d_k} تعریف شده است، هستند. در این مدل پارامترهای θ_k و π_k مجهول هستند و نیاز به برآورد دارند. برآورد پارامترهای مدل آمیخته با استفاده از الگوریتم EM انجام می شود. معمولاً فرض می شود که f_k تابع چگالی نرمال چند متغیره با میانگین μ_k و ماتریس کواریانس Σ_k باشد. فرض نرمال بودن تا اندازه ای محدود کننده است و معمولاً برای سادگی کار در نظر گرفته می شود. در عمل ممکن است توزیع حاشیه ای هر کدام از متغیرهای X_1, \dots, X_p نرمال باشد، اما توأم نرمال نباشند، یا حتی ممکن است توزیع های حاشیه ای و توأم هیچکدام نرمال نباشند. در این موارد با استفاده از قضیه اسکالر، می توان یک ساختار همبستگی برای توزیع توأم متغیرها در نظر گرفت.

تعریف: تابع $C: [0, 1]^2 \rightarrow [0, 1]$ را تابع مفصل گویم هرگاه

$$(1) \quad C(u, 1) = C(1, u) = u \quad \text{و} \quad C(u, 0) = C(0, u) = 0, \quad u \in [0, 1]$$

$$(2) \quad \text{برای هر } 0 \leq v_1 \leq v_2 \leq 1 \quad \text{و} \quad 0 \leq u_1 \leq u_2 \leq 1$$

$$V_C([u_1, u_2] \times [v_1, v_2]) := C(u_2, v_2) - C(u_1, v_2) - C(u_2, v_1) + C(u_1, v_1) \geq 0.$$

اگر X و Y متغیرهای تصادفی پیوسته با تابع توزیع توأم $H(x, y)$ و توابع توزیع حاشیه ای $F(x)$ و $G(y)$ باشند، تابع مفصل C در واقع تابع توزیع توأم $(U, V) = (F(x), G(y))$ است که بنا به قضیه تبدیل انتگرال، U و V متغیرهای تصادفی یکنواخت بر فاصله $[0, 1]$ هستند.

قضیه اسکالر: فرض کنید H یک تابع توزیع توأم با توابع توزیع حاشیه ای F و G باشد، در اینصورت تابع مفصل C وجود دارد به طوری که برای هر $x, y \in \mathbb{R}$ $H(x, y) = C(F(x), G(y))$ اگر F و G پیوسته باشند، آنگاه C یکتاست؛ در غیر اینصورت C بطور یکتا بر روی $\text{Range}F \times \text{Range}G$ تعیین می شود. برعکس اگر C یک تابع مفصل باشد و F و G توابع توزیع باشند در اینصورت تابع H که در عبارت بالا تعریف شد یک تابع توزیع توأم با توزیع های حاشیه ای F و G است.

بنابراین تابع مفصل C ، مربوط به تابع توزیع توأم H را می توان بصورت

$$C(u, v) = H(F^{-1}(u), G^{-1}(v))$$

بدست آورد.

در این مقاله، خوشه بندی مدل محور را با در نظر گرفتن یک ساختار همبستگی دلخواه به جای توزیع نرمال توأم بیان می کنیم. حال فرض کنید X_1, \dots, X_p یک بردار تصادفی p بعدی که بر روی فضای \mathbb{R}^p تعریف شده است، باشد. همچنین F^j ، $j = 1, \dots, p$ تابع توزیع مربوط به X_j است. بنابراین یک نمونه $(F_1, \dots, F_N) \in \mathbb{F}$ از N توزیع p بعدی داریم که $F_u = (F_u^1, \dots, F_u^p)$ ، هر F_u متعلق به فضای $\Omega_F = \Omega_F^1 \times \dots \times \Omega_F^p$ است.

هدف، خوشه‌بندی این نمونه N تایی توزیع‌ها به K خوشه و پس از آن بدست آوردن برآوردهای توزیع هر خوشه است. برای تحقق این هدف به توابع توزیع مشاهدات و تابع توزیع توام مشاهدات هر خوشه نیازمندیم. فرض کنید $\mathbb{F} = (F_1, \dots, F_N)$ یک نمونه N تایی از توزیع‌ها از جامعه Ω_F باشد. یک تابع توزیع در نقطه Z ، تابعی است که بصورت $G_Z(x) = G_Z(x)$ ، $x \mapsto G_Z(x)$ ، $x \in [0, 1] \rightarrow [0, 1]$ تعریف می‌شود که برای هر $x \in \mathbb{R}$ ، $G_Z(x) = G$ ، $F(Z)$ که تابع توزیع است و دامنه Z مشابه دامنه F است. تابع توزیع توام در نقطه $Z = ((Z_1^1, \dots, Z_{n_1}^1), \dots, (Z_1^p, \dots, Z_{n_p}^p))$ تابعی است که به فرم زیر تعریف می‌شود:

$$H_Z : [0, 1]^n \rightarrow [0, 1], n = \sum_{j=1}^p n_j, x = ((x_1^1, \dots, x_{n_1}^1), \dots, (x_1^p, \dots, x_{n_p}^p)) \mapsto H_Z(x)$$

$$H_Z(x) = P(F_u \in \mathbb{F}; F_u^1(Z_1^1) \leq x_1^1, \dots, F_u^p(Z_{n_p}^p) \leq x_{n_p}^p, u = 1, \dots, N)$$

برای سادگی کار، با $p = 1$ کار می‌کنیم.

فرض کنید H تابع توزیع n بعدی با توابع توزیع حاشیه‌ای یک بعدی F_1, \dots, F_N باشد، آنگاه با استفاده از قضیه اسکالر (۱۹۵۹)، یک تابع مفصل C وجود دارد که برای هر $(x_1, \dots, x_n) \in \mathbb{R}^n$ ،

$$H(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n))$$

اگر F_1, \dots, F_N پیوسته باشند، آنگاه C یکتاست؛ در غیر اینصورت C بطور یکتا بر روی $\text{Range}F_1 \times \dots \times \text{Range}F_N$ تعیین می‌شود که $\text{Range}F_u = [0, 1]$ دامنه F_u است.

فرض کنید G_{Z_1}, \dots, G_{Z_n} توابع توزیع در نقاط Z_1, \dots, Z_n باشند و H_{Z_1, \dots, Z_n} تابع توزیع توام این توابع توزیع است. بنابراین یک تابع مفصل C وجود دارد بطوریکه برای هر $(x_1, \dots, x_n) \in \mathbb{R}^n$ ،

$$H_{Z_1, \dots, Z_n}(x_1, \dots, x_n) = C(G_{Z_1}(x_1), \dots, G_{Z_n}(x_n))$$

حال $H_{Z_1, \dots, Z_n}(x_1, \dots, x_n)$ را می‌توان بصورت ترکیبی از توزیع‌ها نوشت.

$$H(x_1, \dots, x_n; \theta) = \sum_{k=1}^K \pi_k H_k(x_1, \dots, x_n; \theta_k)$$

که $0 \leq \pi_k \leq 1$ و $\sum_{k=1}^K \pi_k = 1$. $H_k(\cdot; \theta_k)$ توزیع توام خوشه k ام با پارامتر $\theta_k = (\theta_{k1}, \dots, \theta_{kd_k})$ که بر روی \mathbb{R}^{d_k} تعریف شده است می‌باشد. π_k احتمال پیشین تعلق (x_1, \dots, x_n) به خوشه k ام است. تابع H_k ، توزیع توام در

نقطه $Z = (Z_1, \dots, Z_n)$ برای خوشه k ام با توزیع حاشیه‌ای $G_{Z_1}^k, \dots, G_{Z_n}^k$ است.

با استفاده از قضیه اسکالر، یک تابع مفصل C_k ، $k = 1, \dots, K$ وجود دارد بطوریکه:

$$H(x_1, \dots, x_n; \theta) = \sum_{k=1}^K \pi_k C_k(G_{Z_1}^k(x_1; b_1^k), \dots, G_{Z_n}^k(x_n; b_n^k); \beta_k)$$

که در آن $\theta = \{b_i^k, i = 1, \dots, n; \beta_k, \pi_k, k = 1, \dots, K\}$ پارامتر تابع مفصل مربوط به خوشه k ام است. تابع $G_{Z_i}^k(\cdot; b_i^k)$ توزیع با پارامتر b_i^k در نقطه Z_i در خوشه k ام است. ما با استفاده از عبارات فوق می توانیم به نتایج زیر برسیم. فرض کنید $h_k(\cdot) \equiv h_k(x_1, \dots, x_n; \theta_k) = \partial^n H_k \partial x_1 \dots \partial x_n$ نشانگر تابع چگالی احتمال مربوط به تابع توزیع $H_k(\cdot)$ باشد. آنگاه $h_k(\cdot)$ به فرم زیر در می آید:

$$h_k(x_1, \dots, x_n; \theta_k) = \left\{ \prod_{i=1}^n \frac{d}{dx_i} G_{Z_i}^k(x_i; b_i^k) \right\} \frac{\partial^n}{\partial x_1 \dots \partial x_n} C_k(G_{Z_1}^k(x_1; b_1^k), \dots, G_{Z_n}^k(x_n; b_n^k); \beta_k)$$

$$h(\cdot) \equiv h(x_1, \dots, x_n; \theta) = \frac{\partial^n H}{\partial x_1 \dots \partial x_n}$$

$$h(x_1, \dots, x_n; \theta) = \sum_{k=1}^K \pi_k \left\{ \prod_{i=1}^n \frac{d}{dx_i} G_{Z_i}^k(x_i; b_i^k) \right\} \frac{\partial^n}{\partial x_1 \dots \partial x_n} C_k(G_{Z_1}^k(x_1; b_1^k), \dots, G_{Z_n}^k(x_n; b_n^k); \beta_k)$$

حال برای برآورد پارامترهای مدل آمیخته فوق، به برآورد پارامتر تابع مفصل، برآورد پارامتر تابع $G(Z)$ و برآورد توابع احتمال پیشین نیاز داریم.

۲ برآورد تابع $G(Z)$

اگر تابع $G(Z)$ از فضای \mathbb{F} بیاید، تابع تجربی توزیع $G(Z)$ به فرم زیر به دست می آید،

$$G_Z^e(x) = P(F_u \in \mathbb{F}; F_u(Z) \leq x, u = 1, \dots, N) = \frac{\text{card}(F_u \in \mathbb{F}; F_u(Z) \leq x, u = 1, \dots, N)}{\text{card}(\mathbb{F})}$$

با تعمیم دادن تابع تجربی فوق، می توانیم به برآورد $G(Z)$ دست یابیم.

۳ برآورد تابع مفصل

برای برآورد پارامتر تابع مفصل از روش ماکزیمم درستنمایی استفاده می کنیم. اگر هر مشاهده u بوسیله F_u مشخص شود، فرض کنید $\{F_u(Z_i), i = 1, \dots, n\}$ بوسیله $\{x_i, i = 1, \dots, n\}$ تعریف می شود. بنابراین پارامتر $\beta_k, k = 1, \dots, K$ زمانی که عبارت زیر برای تابع مفصل مشخصی ماکزیمم گردد، برآورد می شود.

$$L = \sum_{k=1}^K \sum_{u \in P_k} \ln \left[\left\{ \prod_{i=1}^n \frac{d}{dx_i} G_{Z_i}^k(x_i; b_i^k) \right\} \frac{\partial^n}{\partial x_1 \dots \partial x_n} C_k(G_{Z_1}^k(x_1; b_1^k), \dots, G_{Z_n}^k(x_n; b_n^k); \beta_k) \right]$$

به عنوان مثال، برای تابع مفصل فرانک می‌توانیم برای حالت $n = 2$ نشان دهیم،

$$C(x_1, x_2) = (\ln \beta)^{-1} \ln \left\{ 1 + \frac{[(\beta^{x_1} - 1)(\beta^{x_2} - 1)]}{(\beta - 1)} \right\}$$

$$\frac{\partial^2 C}{\partial x_1 \partial x_2} = \frac{(\beta - 1) \beta^{x_1 + x_2} \ln \beta}{[(\beta - 1) + (\beta^{x_1} - 1)(\beta^{x_2} - 1)]^2}$$

تابع L با استفاده از روش‌های عددی می‌تواند ماکزیمم شود. بنابراین، تابع L را می‌توانیم به فرم زیر بنویسیم،

$$L = \sum_{k=1}^K \sum_{u \in P_k} \ln \left\{ \prod_{i=1}^n \frac{d}{dx_i} G_{Z_i}^k(x_i; b_i^k) \right\} \times l_k(\beta_k)$$

$$l_k(\beta_k) = \frac{\partial^n}{\partial x_1 \dots \partial x_n} C_k(G_{Z_1}^k(x_1; b_1^k), \dots, G_{Z_n}^k(x_n; b_n^k); \beta_k)$$

برای برآورد تابع مفصل $C_k(\cdot)$ باید بردار β_k را با ماکزیمم کردن $l_k(\beta_k)$ برای یک تابع مفصل مشخص، بیابیم که این هدف با استفاده از روش‌های عددی تحقق می‌یابد.

۴ برآورد π_k

احتمال‌های پیشین در مدل آمیخته به صورت $\hat{\pi}_k = \frac{\text{card}(P_k)}{\text{card}(\mathbb{F})}$ برآورد می‌شود. که در آن P_k نماد خوشه k ام است. پس از به دست آمدن برآوردها، با جایگذاری آن‌ها در فرم توزیع آمیخته، می‌توان توزیع آمیخته خوشه‌ها را با استفاده از تابع مفصل بدست آورد.

مراجع

- [1] Banfield, J.D. and Raftery, A.E. (1993), Model-based Gaussian and non-Gaussian clustering, *Bio-metrics*, **49**, 803–812.

تحلیل داده‌های فضایی با استفاده از توابع مفصل

امیدی، م. ۱ و محمدزاده، م.

گروه آمار دانشگاه تربیت مدرس

چکیده

یکی از ابزارهای قوی برای ساخت توزیع توأم متغیرهای وابسته بر اساس توزیع‌های کناری متغیرها توابع مفصل هستند. توابع مفصل مدلی را برای بررسی ساختار ارتباط بین متغیرها ارائه می‌دهند، که بر اساس آن تمام خصوصیات وابستگی میان متغیرها قابل بیان است. از طرفی در تحلیل داده‌های فضایی لازم است توزیع چندمتغیره تحقق‌های میدان تصادفی، ساختار همبستگی داده‌ها مشخص گردد. در تحلیل داده‌های فضایی-زمانی گاهی برای راحتی از تابع کواریانس تفکیک‌ناپذیر استفاده می‌شود در حالی که این ویژگی در برخی از مثال‌های کاربردی واقع‌گرایانه نیست در این‌گونه موارد استفاده از توابع کواریانس فضایی-زمانی تفکیک‌ناپذیر ضرورت پیدا می‌کند. در این مقاله ابتدا نقش توابع مفصل در تعیین توزیع توأم تحقق‌های میدان تصادفی بررسی می‌شود، سپس با استفاده از توابع مفصل اقدام به ساخت کواریانس‌های فضایی و فضایی-زمانی معتبر خواهد شد.

واژه‌های کلیدی: تابع مفصل، تابع کواریانس فضایی-زمانی، داده‌های فضایی.

۱ مقدمه

یکی از مسائل مهم در آمار فضایی پیشگویی مقدار میدان تصادفی در مکان‌های مشخص بر اساس مشاهدات است که معمولاً با فرض گاوسی بودن میدان تصادفی انجام می‌شود. گاهی این فرض واقع‌گرایانه نیست و باید توزیع میدان تعیین شود. یکی از روش‌های تعیین توزیع میدان استفاده از توابع مفصل^۲ است. این توابع اولین بار توسط اسکالر (۱۹۵۹) معرفی شدند، جو (۱۹۹۷) و نلسن (۲۰۰۶) خانواده‌های مختلف این توابع را که در علوم مختلف به کار می‌رود جمع‌آوری نمودند. امیدی و همکاران (۱۳۸۹) خانواده‌های مفصل مناسب برای مدل‌بندی خشکسالی در استان تهران را بررسی کردند. امیدی و محمدزاده (۱۳۹۰) با رویکرد بیز تجربی برآورد پارامتر تابع مفصل را بدست آوردند. باردوسی (۲۰۰۶) برای اولین بار از مفصل گاوسی و مفصل خی-دو غیر مرکزی در تحلیل داده‌های فضایی استفاده کرد. باردوسی و لای (۲۰۰۸) از مفصل V -تبدیل یافته نرمال چند متغیره برای پیشگویی خصوصیات مکانی آب‌های زیرزمینی استفاده کردند. گرایلر و پیسما (۲۰۱۱) با استفاده از ترکیب خطی از توابع

^۱email: omidi_280@yahoo.com

^۲Copula Function

مفصل، تابع مفصل فضایی معتبر ساخته و از آن در پیشگویی فلزات سنگین در رودخانه میوز استفاده کردند. تعیین توابع کواریانس فضایی و فضایی-زمانی^۳ مساله‌ای مهم در تحلیل داده‌های همبسته فضایی و زمانی است. کرسی و هوانگ (۱۹۹۹) از روش تابع طیفی برای ساخت این توابع استفاده کردند. این روش توسط گنیتینگ (۲۰۰۲) با استفاده از توابع کاملاً یکنوا و برنشتاین تعمیم داده شد. ما (۲۰۰۲ و ۲۰۰۵) بر اساس ترکیبات محدب از توابع کواریانس معتبر و استفاده از تبدیلات لاپلاس رده جدیدی از توابع کواریانس فضایی-زمانی تفکیک‌ناپذیر را معرفی کرد. کنت و همکاران (۲۰۱۱) نشان دادند که مدل گنیتینگ (۲۰۰۲) برای برخی از توابع کاملاً یکنوا و برنشتاین دلای گودالی^۴ است که استفاده از آن‌ها را به‌گونه‌ای محدود می‌کند. پورسو و همکاران نشان دادند که با استفاده از توابع مفصل ساخت تابع کواریانس فضایی-زمانی تفکیک‌ناپذیر می‌تواند قابل توجه باشد. در این مقاله ابتدا نقش توابع مفصل در تعیین توزیع توأم تحقق‌های یک میدان تصادفی بررسی می‌شود، سپس با استفاده از توابع مفصل روش‌هایی برای ساخت توابع کواریانس فضایی و فضایی-زمانی معتبر ارائه و اعتبار آن‌ها مورد ارزیابی قرار می‌گیرد.

۲ توابع مفصل

برای دو متغیر تصادفی پیوسته X و Y به ترتیب با توابع توزیع $F_X(x)$ و $G_Y(y)$ و تابع توزیع توأم $H(x, y)$ ، توابع مفصلی مانند $K(\cdot, \cdot)$ وجود دارند که برای هر x و y در $\bar{R} = R \cup \{-\infty, \infty\}$ داریم:

$$H(x, y) = K(F_X(x), G_Y(y))$$

اگر F_X و G_Y مطلقاً پیوسته باشند، تابع مفصل K یکتاست. اگر $f_X(x)$ و $f_Y(y)$ توابع چگالی کناری X و Y باشند، آنگاه تابع چگالی توأم آن‌ها به صورت

$$f_{X,Y}(x, y) = k(F_X(x), F_Y(y))f_X(x)f_Y(y)$$

است، که در آن k تابع مفصل چگالی توأم بوده و از رابطه $k(u, v) = \frac{\partial^2 K(u, v)}{\partial u \partial v}$ به دست می‌آید. همچنین می‌توان نشان داد که برای هر تابع مفصل حدهای بالا و پایین به صورت

$$W(u, v) = \max(u + v - 1, 0) \leq K(u, v) \leq \min(u, v) = M(u, v) \quad (1)$$

وجود دارد، که با M و W نشان داده می‌شوند و به ترتیب کران‌های بالا و پایین فرشه-هافدینگ هستند (جو، ۱۹۹۷). بنا بر (1) اگر دو متغیر دارای حداکثر وابستگی باشند آنگاه تابع مفصل آن‌ها M و اگر دارای حداقل

^۳Spatio-Temporal

^۴Dimple

وابستگی باشند، تابع مفصل آن‌ها W است. همچنین برای دو متغیر تصادفی مستقل تابع مفصل حاصل ضرب به صورت $\Pi(u, v) = uv$ تعریف می‌شود.

قضیه ۱.۲. (نلسن، ۲۰۰۶) اگر $\varphi(t)$ یک تابع پیوسته اکیداً نزولی از $[0, 1]$ به $[0, \infty]$ با شرط $\varphi(1) = 0$ و وارون $\varphi^{-1}(t)$ باشد، آنگاه

$$K(F_X(x), G_Y(y)) = \varphi^{-1}(\varphi(F_X(x)) + \varphi(G_Y(y))) \quad (2)$$

یک تابع مفصل است، اگر و فقط اگر $\varphi(t)$ محدب باشد. مجموعه توابع مفصل ساخته شده بر اساس رابطه (2) خانواده ارشمیدسی است و $\varphi(t)$ تابع مولد مفصل ارشمیدسی نامیده می‌شود.

۳ توابع مفصل و آمار فضایی

برای پیشگویی در آمار فضایی میدان تصادفی $Z = \{Z(s); s \in D \subset R^d\}$ در موقعیت مشخص s بر اساس تحقیقات میدان تصادفی یعنی $z = (z(s_1), \dots, z(s_n))$ در موقعیت‌های s_1, \dots, s_n است. پیشگویی فضایی معمولاً با فرض گاوسی بودن میدان تصادفی انجام می‌شود. گاهی این فرض واقع‌گرایانه نیست و لازم است توزیع $Z = (Z(s_1), \dots, Z(s_n))$ تعیین شود. با استفاده از توابع مفصل می‌توان توزیع توأم $Z(s)$ و $Z(s+h)$ را به صورت

$$K_h(F_Z(z_1), F_Z(z_2)) = P(Z(s) \leq z_1, Z(s+h) \leq z_2)$$

به دست آورد. با توجه به خواص وابستگی داده‌های فضایی تابع مفصل $K_h(\cdot, \cdot)$ معتبر است هرگاه وابستگی مثبت میان متغیرها را نتیجه دهد، به علاوه وقتی $\|h\|$ به بینهایت میل کند تابع مفصل حاصل ضرب و موقعی که $\|h\|$ به صفر میل کند باید کران بالای فرشه-هافدینگ را نتیجه دهد، یعنی

$$\lim_{\|h\| \rightarrow \infty} K_h(u, v) = uv \quad , \quad \lim_{\|h\| \rightarrow 0} K_h(u, v) = M(u, v)$$

بدیهی است که تمام توابع مفصل ارائه شده در منابع لزوماً واجد شرایط یک مفصل فضایی معتبر نیستند. در ادامه توابع مفصل فضایی معتبر معرفی می‌شوند. برخی از مفصل‌های فضایی معتبر موجود در منابع بر اساس تابع مفصل گاوسی و تبدیل‌هایی از توزیع نرمال چندمتغیره به دست آمده‌اند. مفصل گاوسی به صورت

$$K(u_1, \dots, u_n) = \Phi_{\cdot, \Lambda}(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_n))$$

تعریف می‌شود. از آنجا که این مفصل برای همبستگی صفر تابع مفصل استقلال و برای همبستگی یک کران بالای فرشه-هافدینگ را نتیجه می‌دهد، یک تابع مفصل فضایی معتبر است. اکنون تبدیل‌هایی از توزیع نرمال چندمتغیره که منجر به ساخت مفصل‌های فضایی معتبر می‌شوند، مورد بررسی قرار می‌گیرند. **مفصل خی- دو غیرمرکزی:**

باردوسی (۲۰۰۶) با استفاده از تبدیل $g(Y) = Y^2$ ، که در آن Y دارای توزیع نرمال چندمتغیره است، مفصل
خی-دوی غیرمرکزی را با تابع مفصل چگالی به صورت

$$k(u, v) = K_{\nu} \frac{e^{-\frac{\sqrt{xy}r + (x+y)r^{\nu}}{\nu(r^{\nu}-1)}} (1 + e^{\frac{\nu r \sqrt{xy}}{r^{\nu}-1}})}{\sqrt{4\pi xy}}$$

به دست آورد. این مفصل برای همبستگی صفر استقلال و برای همبستگی یک کران بالای فرشه-هافدینگ را نتیجه
می دهد. **توزیع V -تبدیل یافته تابع مفصل نرمال چندمتغیره:** این مفصل توسط باردوسی و لای (۲۰۰۸) و
با استفاده از تبدیل

$$X_j = g(Y_j) = \begin{cases} k(Y_j - m) & Y_j \geq m \\ m - Y_j & Y_j < m \end{cases}$$

معرفی شد، که در آن توزیع های یک و چندمتغیره X به ترتیب به صورت

$$H_1(x) = P(X \leq x) = \Phi\left(\frac{k}{m} + 1\right) - \Phi(-x - m)$$

و

$$\begin{aligned} H_n(x_1, \dots, x_n) &= P(X_1 \leq x_1, \dots, X_n \leq x_n) \\ &= \Phi\left(\frac{k}{m} + 1\right) - \Phi(-x - m) \\ &= \sum_{i=0}^{\nu^n - 1} (-1)^i \Phi(\xi_i + m) \end{aligned}$$

به دست می آیند، که در آن $\xi_i^T = \{b(-1)^{i_1} x_1, \dots, b(-1)^{i_n} x_n\}$

$$b(x) = \begin{cases} -1 & x < 0 \\ \frac{1}{k} & Y_j < m \end{cases}$$

تابع مفصل فضایی معتبر متناظر با این تبدیل به صورت

$$K(u_1, \dots, u_n) = \sum_{i=0}^{\nu^n - 1} (-1)^i \Phi(\xi_i^T + m)$$

حاصل می شود، که در آن $\xi_i^T = \{b(-1)^{i_1} H_1^{-1}(u_1), \dots, b(-1)^{i_n} H_n^{-1}(u_n)\}$.

مفصل ترکیبی: گرایلر و پیسما (۲۰۱۱) با استفاده از ترکیب خطی از توابع مفصل، تابع مفصل فضایی معتبر
را بر اساس فواصل بین موقعیت ها به صورت زیر تعریف کردند.

$$K(u, v) = \begin{cases} \lambda_1 M(u, v) + (1 - \lambda_1) K_{1,h}(u, v) & 0 \leq h \leq h_1 \\ \vdots & \vdots \\ \lambda_i M(u, v) + (1 - \lambda_i) K_{1,h}(u, v) & h_{i-1} \leq h \leq h_i \\ \vdots & \vdots \\ \lambda_l M(u, v) + (1 - \lambda_l) K_{1,h}(u, v) & h_{l-1} \leq h \leq h_l \end{cases}$$

این تابع ویژگی تابع مفصل فضایی معتبر را بر اساس فاصله‌ها حفظ می‌کند، اما انتخاب خانواده مفصل مناسب در هر بازه بایستی به گونه‌ای باشد که همواره همبستگی مثبت را نتیجه دهد و با افزایش فاصلهٔ موقعیت‌ها همبستگی کاهش یابد. به عنوان مثال تابع مفصل گامبل-بارنت که به ازای $\theta \in (0, 1]$ به صورت $K(u, v) = uv \exp(-\theta \log u \log v)$ است، همواره وابستگی منفی را نتیجه می‌دهد و نمی‌تواند یک تابع مفصل فضایی معتبر باشد.

۴ توابع مفصل و همبستگی فضایی

مفصل‌ها ابزاری برای مدل‌بندی ساختار همبستگی متغیرهای تصادفی وابسته هستند. بسیاری از خانواده‌های این توابع می‌توانند وابستگی کامل مثبت و منفی را برای داده‌های وابسته پوشش دهند. داده‌های فضایی وابسته مثبت هستند و ساختار همبستگی آن‌ها با توابع کواریانس همیشه مثبت بیان می‌شود. در این بخش نقش توابع مفصل در ساخت توابع کواریانس فضایی معتبر بیان می‌شود.

تعریف ۱.۴. تابع پیوستهٔ مثبت $\phi(t)$ به ازای $t \geq 0$ ، یک تابع کاملاً یکنوا^۵ است، اگر مشتق‌های تمام مراتب آن وجود داشته و

$$(-1)^n \phi^{(n)}(t) \geq 0, \quad n = 0, 1, \dots$$

طبق قضیه برنشتاین (فلر، ۱۹۶۶) تابع مثبت $\phi(t)$ با شرط $\phi(0) = 1$ کاملاً یکنواست اگر و تنها اگر تبدیل لاپلاسی از یک تابع کراندار نامنفی باشد، یعنی

$$\phi(t) = \int_0^\infty e^{-\alpha t} dM(r), \quad t > 0 \quad (۳)$$

که در آن M یک اندازهٔ کراندار نامنفی است.

تعریف ۲.۴. تابع پیوستهٔ مثبت $\psi(t)$ به ازای $t \geq 0$ ، یک تابع برنشتاین^۶ است، هرگاه مشتق اول آن یک تابع کاملاً یکنوا باشد.

قضیه ۱.۴. (میلر و سامکو، ۲۰۰۱) اگر $\phi_1(t)$ و $\phi_2(t)$ توابعی کاملاً یکنوا باشند، آنگاه $\phi_1(t) + \phi_2(t)$ و $\phi_1(t)\phi_2(t)$ نیز کاملاً یکنوا هستند.

قضیه ۲.۴. (میلر و سامکو، ۲۰۰۱) اگر $\phi(t)$ تابعی کاملاً یکنوا و $\psi(t)$ برنشتاین غیر منفی باشد، آنگاه $\phi(\psi(t))$ نیز کاملاً یکنوا است.

^۵Competely Monotone

^۶Bernestine

قضیه ۳.۴. اگر $\psi_1(t)$ و $\psi_2(t)$ دو تابع برنشتاین باشند، آنگاه $\psi(t) = \psi_2(\psi_1(t))$ برنشتاین است.

برهان. چون

$$\psi'(t) = \psi_1'(\psi_1(t))\psi_2'(\psi_1(t))$$

بنا به فرض $\psi_1(t)$ و $\psi_2(t)$ برنشتاین هستند، بنابراین طبق تعریف مشتق‌های $\psi_1'(t)$ و $\psi_2'(t)$ توابعی کاملاً یکنوا بوده و بنا بر قضایای 1.4 و 2.4 تابع $\psi(t)$ برنشتاین است. \square

قضیه ۴.۴. (ما، ۲۰۰۵) فرض کنید $\phi(t)$ تابعی کاملاً یکنوا روی $[0, \infty)$ یا به‌طور معادل تبدیل لاپلاسی از یک تابع غیر نزولی کراندار باشد، در این صورت

$$C_s(h) = \phi(\|h\|^2)$$

یک تابع کواریانس فضایی پیوسته همسان‌گرد است.

نتیجه ۱.۴. چون t^ρ به‌ازای $\rho \in [0, 1]$ تابعی برنشتاین است، برای هر تبدیل لاپلاس $\phi(t)$ ، بنا بر قضیه 2.4 تابع $\phi(t^\rho)$ کاملاً یکنوا و تبدیل لاپلاس است. بنا بر این $C_s(h) = \phi(\|h\|^\rho)$ به‌ازای هر $\rho \in [0, 2]$ یک تابع کواریانس پیوسته فضایی همسان‌گرد معتبر است.

نتیجه ۲.۴. چون وارون تابع مولد یک مفصل ارشمیدسی تابعی کاملاً یکنوا است، می‌توان بر اساس آن توابع کواریانس فضایی یا زمانی معتبر ساخت.

مثال ۱.۴. تابع مولد مفصل ارشمیدسی علی-میخیایل-حق به‌ازای $\theta \in [-1, 1)$ به‌صورت $\phi(t) = \ln \frac{1-\theta(1-t)}{t}$ و وارون آن به‌صورت $\frac{1-\theta}{\exp(t)-\theta}$ است. حال بنا بر نتیجه 1.4 تابع $\frac{1-\theta}{\exp(t)-\theta}$ به‌ازای $0 \leq \rho \leq 2$ یک کواریانس فضایی معتبر است.

در جدول ۱ برخی از توابع مفصل خانواده ارشمیدسی به همراه تابع کواریانس فضایی معتبر متناظر آن‌ها ارائه شده است. لازم به ذکر است که نتیجه 2.4 لزوماً برای همه توابع مفصل از جمله موارد زیر برقرار نیست.

مثال ۲.۴. تابع مولد مفصل ارشمیدسی ۲ از نلسن (۲۰۰۶) به‌ازای $\theta \in [1, \infty)$ به‌صورت $(1-t)^\theta$ و وارون $1-t^{\frac{1}{\theta}}$ است. با وجود اینکه این تابع کاملاً یکنواست اما تابع $1-\|h\|^{\frac{1}{\theta}}$ برای بعضی از $\|h\|$ منفی است، که مغایر با همبستگی مثبت داده‌های فضایی است.

مثال ۳.۴. تابع مولد مفصل ارشمیدسی ۸ از نلسن (۲۰۰۶) به‌ازای $\theta \in [1, \infty)$ به‌صورت $\frac{1-t}{1+(\theta-1)t}$ و وارون $\frac{t-1}{(\theta-1)t-1}$ است. با وجود اینکه این تابع کاملاً یکنواست اما تابع $\frac{\|h\|-1}{(\theta-1)\|h\|-1}$ برای بعضی از $\|h\|$ منفی است، که مغایر با همبستگی مثبت داده‌های فضایی است.

۵ ساخت تابع کواریانس فضایی- زمانی

داده‌های فضایی که در طول زمان وابسته باشند، داده‌های فضایی- زمانی نامیده می‌شوند و با میدان تصادفی $\{Z(s, t); (s, t) \in D \times T\}$ مدل‌بندی می‌شوند، که در آن $T \subset R$ ، $D \subset R^d$ ، موقعیت فضایی s و موقعیت زمانی مشاهده را نشان می‌دهند. توابع میانگین و کواریانس فضایی- زمانی میدان تصادفی به ترتیب به صورت

$$\mu(s, t) = E(Z(s, t)); (s, t) \in D \times T$$

$$C(s, s', t, t') = Cov(Z(s, t), Z(s', t')); (s, t), (s', t') \in D \times T$$

تعریف می‌شوند.

تعریف ۱.۵. میدان تصادفی فضایی- زمانی $Z(\cdot, \cdot)$ مانای مرتبه دوم است هرگاه میانگین آن نسبت به (s, t) ثابت، یعنی $\mu(s, t) = \mu$ و کواریانس فضایی- زمانی تابعی از فاصله موقعیت‌های فضایی- زمانی به صورت

$$Cov(Z(s, t), Z(s', t')) = C(\mathbf{h}_s, \mathbf{h}_t); (\mathbf{s}, \mathbf{t}), (\mathbf{s}', \mathbf{t}') \in \mathbf{D} \times \mathbf{T}$$

باشد، که در آن $h_t = t - t'$ و $h_s = s - s'$ به ترتیب تاخیرهای فضایی^۷ و زمانی^۸ هستند. تابع کواریانسی که فقط تابعی از اندازه تاخیرها بوده و به جهت آن‌ها بستگی ندارد همسان‌گرد^۹ نامیده می‌شود.

ما (۲۰۰۳) دو ترکیب از توابع کواریانس فضایی و زمانی معتبر را برای ساخت توابع کواریانس فضایی- زمانی به صورت زیر تعریف کرد.

اگر $C_s(s)$ و $C_t(t)$ به ترتیب دو تابع کواریانس فضایی و زمانی در فضای R^d و R باشند آنگاه توابع

$$C_{s,t}(\mathbf{h}_s, \mathbf{h}_t) = \int C_s(uh_s)C_t(vh_t)dW(u, v) \quad (۴)$$

$$C_{s,t}(\mathbf{h}_s, \mathbf{h}_t) = \int [C_s(h_s)]^u [C_t(h_t)]^v dW(u, v) \quad (۵)$$

کواریانس‌های فضایی- زمانی تفکیک‌ناپذیر معتبر هستند، که در آن‌ها $W(u, v)$ اندازه غیر منفی و کراندار در $R^d \times R$ است. چون توابع مفصل کراندار و مثبت هستند، برای هر تابع مفصل $K(u, v)$ بنا بر روابط (۴) و (۵) توابع

$$C_{s,t}(h_s, h_t) = \int C_s(uh_s)C_t(vh_t)dK(u, v)$$

$$C_{s,t}(h_s, h_t) = \int [C_s(h_s)]^u [C_t(h_t)]^v dK(u, v)$$

تابع کواریانس‌های فضایی- زمانی معتبر هستند.

^۷Spatial Lag

^۸Time Lag

^۹Isotropic

نتیجه ۱.۵. هر تابع مفصل به جای $W(u, v)$ در روابط (4) و (5) خانواده‌ای جدیدی از توابع کواریانس فضایی-زمانی را فرع می‌دهد.

نتیجه ۲.۵. برخی از خانواده‌های توابع مفصل فقط تابع کواریانس فضایی-زمانی تفکیک‌پذیر را فرع می‌دهند. به‌عنوان مثال از جایگذاری

$$K(u, v) = pW + (1 - p - q)\Pi + qM$$

یا

$$K(u, v) = \max\{\theta\Pi + (1 - \theta)W, 0\}$$

در روابط (4) و (5) کواریانس فضایی-زمانی تفکیک‌پذیر حاصل می‌شود.

برهان. برای هر α مثبت داریم $dK(u, v) = \alpha dudv$ ، که با قرار دادن آن در روابط (4) و (5) تابع کواریانس به‌صورت حاصل ضرب دو تابع کواریانس فضایی و زمانی به‌دست می‌آید.

جو (۱۹۹۷) برای چند خانواده از تبدیلات لاپلاس نشان داد که اگر $\phi_{\theta_1}(t)$ تابعی کاملاً یکنوا باشد، آنگاه $\phi_{\theta_2}^{-1}(\phi_{\theta_1}(t))$ برای $\theta_1 \geq \theta_2$ متعلق به ζ_{∞}^* است، که در آن کلاسی از توابع برنشتاین بوده و برای $\omega \in \zeta_{\infty}^*$ روابط $\omega(\infty) = \infty$ و $\omega(0) = 0$ برقرار است. □

نتیجه ۳.۵. اگر $\varphi_{\theta}(t)$ تابع مولد مفصل ارشمیدسی باشد، آنگاه وارون آن $\phi_{\theta}(t)$ تابعی کاملاً یکنوا است و $\varphi_{\theta_2}(\phi_{\theta_1}(t))$ یا به‌طور معادل $\phi_{\theta_2}^{-1}(\phi_{\theta_1}(t))$ تابعی برنشتاین در کلاس ζ_{∞}^* است.

گنیتینگ (۲۰۰۲) نشان داد اگر $\phi(t)$ تابعی کاملاً یکنوا و $\psi(t)$ تابعی برنشتاین با شرایط $\psi(0) = 1$ و $\psi(\infty) = \infty$ باشد، آنگاه

$$C_{s,t}(\mathbf{h}_s, \mathbf{h}_t) = \frac{1}{\psi(\mathbf{h}_t^2)^{d/2}} \phi\left(\frac{\|\mathbf{h}_s\|^2}{\psi(\mathbf{h}_t^2)}\right) \quad (6)$$

تابع کواریانس فضایی-زمانی مانا در $R^d \times R$ است.

نتیجه ۴.۵. با توجه به فرع 3.5 از تابع مولد مفصل ارشمیدسی می‌توان توابع کاملاً یکنوا و برنشتاین ساخت. چون تابع برنشتاین در کلاس توابع کواریانس گنیتینگ باید در شرط $\psi(0) = 1$ صدق کند، برای صدق کردن کلاس توابع ζ_{∞}^* در تابع کواریانس گنیتینگ کافی است مقدار توابع این کلاس در نقطه صفر با عدد یک جمع شود. بر این اساس با توجه به تابع مولد مفصل ارشمیدسی تابع برنشتاینی که در تابع کواریانس گنیتینگ صدق کند به ازای $\theta_1 \geq \theta_2$ به‌صورت $\phi_{\theta_2}^{-1}(\phi_{\theta_1}(t)) + 1$ به‌دست می‌آید.

مثال ۱.۵. برای تابع مفصل کلایتون تابع کاملاً یکنوا به صورت $\phi_{\theta_1}(t) = (1+t)^{\frac{-1}{\theta_1}}$ و تابع برنشتاین برای استفاده در تابع کواریانس گنیتینگ به صورت

$$\begin{aligned}\phi_{\theta_2}^{-1}(\phi_{\theta_1}(t)) + 1 &= (1+t)^{\frac{-\theta_2}{\theta_1}} - 1 + 1 \\ &= (1+t)^\rho\end{aligned}$$

به دست می آید، که در آن $\rho = \frac{\theta_2}{\theta_1}$. چون طبق فرض $\theta_1 \geq \theta_2$ ، بنابراین $0 < \rho \leq 1$.

مثال ۲.۵. برای تابع مفصل گامبل-هوگارد تابع کاملاً یکنوا به صورت $\phi_{\theta_1}(t) = \exp\{-t^{1/\theta_1}\}$ و تابع برنشتاین برای استفاده در تابع کواریانس گنیتینگ به صورت

$$\begin{aligned}\phi_{\theta_2}^{-1}(\phi_{\theta_1}(t)) + 1 &= (-\ln(\exp\{-t^{1/\theta_1}\}))^{\theta_2} + 1 \\ &= t^\rho + 1\end{aligned}$$

به دست می آید، که در آن $\rho = \frac{\theta_2}{\theta_1}$. چون طبق فرض $\theta_1 \geq \theta_2$ بنا بر این $0 < \rho \leq 1$.

مثال ۳.۵. برای تابع مفصل گامبل-بارنت تابع کاملاً یکنوا به صورت $\phi_{\theta_1}(t) = \exp\{\frac{1-e^t}{\theta_1}\}$ و تابع برنشتاین برای استفاده در تابع گنیتینگ به صورت

$$\begin{aligned}\phi_{\theta_2}^{-1}(\phi_{\theta_1}(t)) + 1 &= \ln\{1 - \frac{\theta_2}{\theta_1}(1 - e^t)\} + 1 \\ &= \ln\{e - e\rho(1 - e^t)\}\end{aligned}$$

به دست می آید، که در آن $\rho = \frac{\theta_2}{\theta_1}$. چون طبق فرض $\theta_1 \geq \theta_2$ بنابراین $0 < \rho \leq 1$.

گنیتینگ (۲۰۰۲) نشان داد اگر $C_s^r(h)$ و $C_t^r(u)$ به ترتیب توابع کواریانس مانای فضایی و زمانی در R و R^d باشند، آنگاه برای اندازه متناهی و نامنفی μ عبارت

$$C_{s,t}(\mathbf{h}_s, \mathbf{h}_t) = \int C_s^r(\mathbf{h}) C_t^r(\mathbf{u}) d\mu(\mathbf{r}) \quad (7)$$

یک تابع کواریانس فضایی-زمانی مانا در $R^d \times R$ است.

قضیه ۱.۵. اگر تابع مولد ارشمیدسی و $\varphi(C_s(h_s))$ و $\varphi(C_t(h_t))$ توابعی برنشتاین باشند، آنگاه با جایگذاری $\exp(-\varphi_1(C_s(h_s)))$ و $\exp(-\varphi_2(C_t(h_t)))$ به ترتیب به جای $C_s^r(h_s)$ و $C_t^r(h_t)$ در (7) یک تابع کواریانس فضایی-زمانی مانای معتبر در $R^d \times R$ به صورت

$$C_{s,t}(\mathbf{h}_s, \mathbf{h}_t) = \int e^{-r(\varphi_1(C_s(\mathbf{h}_s))) + \varphi_2(C_t(\mathbf{h}_t))} d\mu(\mathbf{r}) \quad (8)$$

حاصل می شود.

برهان. طبق قضیه 4.4 چون $\exp(-t)$ کاملاً یکنواست و $\varphi(C_s(h))$ و $\varphi(C_t(u))$ برنشتاین هستند، پس $\exp(-\varphi_1(C_s(h_s)))$ و $\exp(-\varphi_2(C_t(h_t)))$ کاملاً یکنوا است و بنا بر رابطه (7)، رابطه (8) یک تابع کواریانس فضایی-زمانی مانا در $R^d \times R$ است. \square

نتیجه ۵.۵. با در نظر گرفتن (3) مشاهده می‌شود که (7) انتگرال مربوط به تبدیل لاپلاس است. بنا بر این اگر $\phi(t)$ تبدیل لاپلاس متناظر آن باشد، آنگاه

$$\begin{aligned} \int e^{-r(\varphi_1(C_s(\mathbf{h}_s)) + \varphi_2(C_t(\mathbf{h}_t)))} d\mu(r) &= \varphi(\varphi_1(C_s(\mathbf{h}_s)) + \varphi_2(C_t(\mathbf{h}_t))) \\ &= \phi(\varphi_1^{-1}(C_s(\mathbf{h}_s)) + \varphi_2^{-1}(C_t(\mathbf{h}_t))) \end{aligned}$$

یک تابع کواریانس فضایی-زمانی مانا در $\mathbb{R}^d \times \mathbb{R}$ است. بر اساس نتیجه 2.44 و قضیه 4.4 به ازای $\alpha \in (0, 2]$ ، $\beta \in (0, 2]$ ، $\theta_1 \geq \theta_2$ ، $\theta_3 \geq \theta_4$ و a و b مثبت عبارت

$$C_{s,t}(\mathbf{h}, \mathbf{t}(\mathbf{h}_s, \mathbf{h}_t)) = \varphi_\theta(\varphi_{1\theta_2}^{-1}(\varphi_{1\theta_1}(|\mathbf{a}||\mathbf{h}_s|^\alpha)) + \varphi_{2\theta_4}^{-1}(\varphi_{2\theta_3}(|\mathbf{b}||\mathbf{h}_t|^\beta))) \quad (9)$$

یک تابع کواریانس فضایی-زمانی مانای معتبر در $R^d \times R$ است.

نتیجه ۶.۵. رابطه (9) نشان می‌دهد برای هر n ($n > 1$) تابع مولد مفصل ارشمیدسی می‌توان 3^n تابع کواریانس فضایی-زمانی مانای معتبر در $R^d \times R$ ساخت.

مثال ۴.۵. برای تابع مفصل کلایتون با توجه به (9)

$$\begin{aligned} C_{s,t}(\mathbf{h}, \mathbf{t}) &= \{(1 + a||h_s|^\alpha)^{r_1} - 1 + (1 + b|h_t|^\beta)^{r_2} - 1 + 1\}^{\frac{1}{\theta_1}} \\ &= \{(1 + a||h_s|^\alpha)^{r_1} + (1 + b|h_t|^\beta)^{r_2} - 1\}^{\frac{1}{\theta_1}} \end{aligned} \quad (10)$$

به ازای $1 < r_1 = \frac{\theta_1}{\theta_2} < \infty$ ، $1 < r_2 = \frac{\theta_1}{\theta_3} < \infty$ ، $\alpha \in (0, 2]$ ، $\beta \in (0, 2]$ ، $\theta_1 > 0$ ، a و b مثبت یک تابع کواریانس فضایی-زمانی مانای معتبر در $\mathbb{R}^d \times \mathbb{R}$ است.

لازم به ذکر است که با فرض $r_1 = r_2 = 1$ رابطه (10) تابع کواریانس فضایی-زمانی دیاکو و همکاران (۲۰۰۲) را فرع می‌دهد.

مثال ۵.۵. برای تابع مفصل گامبل-هوگارد با توجه به رابطه (9) تابع

$$C_{s,t}(\mathbf{h}, \mathbf{t}) = \exp\{-a||\mathbf{h}|^\alpha - \mathbf{b}|\mathbf{t}|^\beta\}^{\frac{1}{\theta_1}}$$

به ازای $\alpha \in (0, 2]$ ، $\beta \in (0, 2]$ ، $\theta_1 > 1$ و a و b مثبت یک تابع کواریانس فضایی-زمانی مانا در $\mathbb{R}^d \times \mathbb{R}$ است.

مثال ۶.۵. برای توابع مفصل گامبل- هوگارد و کلایتون تابع

$$C_{s,t}(\mathbf{h}, \mathbf{t}) = \exp\{-(\mathbf{1} + \mathbf{a}\|\mathbf{h}_s\|^\alpha)^{r_2} - (\mathbf{1} + \mathbf{b}\|\mathbf{h}_t\|^\beta)^{r_3} + 2\}\}^{\frac{1}{\theta_1}}$$

به ازای $0 < r_2 \leq 1, 0 < r_3 \leq 1, \alpha \in (0, 2], \beta \in (0, 2], \theta_1 > 1$ و a, b مثبت یک تابع کواریانس فضایی-زمانی مانای معتبر در $R^d \times R$ است. جدول ۲ توابع کواریانس فضایی-زمانی معتبر برای این دو خانواده را نشان می‌دهد، که در آن‌ها $0 < r \leq 1, \alpha \in (0, 2], \beta \in (0, 2], \theta > 1$ و a, b مثبت هستند.

جدول ۱: خانواده توابع مفصل ارشمیدسی و تابع کواریانس معتبر متناظر آن‌ها

تابع مفصل	تابع مولد	تابع کاملاً یکنوا	فضای پارامتر	تابع کواریانس
جو	$-\ln(1 - (1 - t)^\theta)$	$1 - (1 - e^{-t})^{\frac{1}{\theta}}$	$[1, \infty)$	$1 - (1 - e^{-\ h\ ^\rho})^{\frac{1}{\theta}}$
فرانک	$-\ln\{\frac{\exp(-\theta t) - 1}{\exp(-\theta) - 1}\}$	$-\frac{1}{\theta} \ln\{(\exp(-\theta) - 1)e^{-t}\}$	$\mathbb{R} - \{0\}$	$-\frac{1}{\theta} \ln\{(\exp(-\theta) - 1)e^{-\ h\ ^\rho}\}$
گامبل-بارنت	$\ln(1 - \theta \ln t)$	$\exp\{\frac{1-e^{-t}}{\theta}\}$	$[0, 1)$	$\exp\{\frac{1-e^{-\ h\ ^\rho}}{\theta}\}$
گامبل-هوگارد	$(-\ln t)^\theta$	$\exp(-t^{1/\theta})$	$[1, \infty)$	$\exp(-\ h\ ^\rho)$

جدول ۲: خانواده توابع مفصل و تابع کواریانس فضایی-زمانی متناظر آن‌ها

تابع کواریانس فضایی-زمانی	فضای پارامتر	تابع مولد		
		φ_2	φ_2	φ_1
$\{a\ h_s\ ^\alpha + b\ h_t\ ^\beta\}^{\frac{-1}{\theta}}$	$[0, \infty)$	گامبل-هوگارد	گامبل-هوگارد	کلایتون
$\{[1 + a\ h_s\ ^\alpha]^r + b\ h_t\ ^\beta\}^{\frac{-1}{\theta}}$	$[0, \infty)$	گامبل-هوگارد	کلایتون	کلایتون
$\exp\{-a\ h_s\ ^\alpha - (1 + b\ h_t\ ^\beta)^r + 1\}\}^{\frac{1}{\theta}}$	$[1, \infty)$	گامبل-هوگارد	کلایتون	گامبل-هوگارد

بحث و نتیجه‌گیری

مفصل‌ها ابزاری مناسب برای مدل‌بندی تابع کواریانس داده‌های فضایی و فضایی-زمانی هستند. گستردگی و انعطاف این توابع امکان ساخت توزیع توأم تحقق‌های میدان تصادفی، ساخت توابع کواریانس فضایی و فضایی-زمانی تفکیک‌ناپذیر را میسر می‌سازد. اما بررسی معتبر بودن توابع مفصل فضایی و توابع کواریانس فضایی و تابع کواریانس فضایی-زمانی تفکیک‌ناپذیر ممکن است فقط برای توابع مفصل خاص میسر باشد.

مراجع

- [۱] امیدی، م.، محمدزاده، م. و مرید، س.، (۱۳۸۸)، تحلیل احتمالاتی مدت - شدت خشکسالی در استان تهران با استفاده از توابع مفصل، مجله تحقیقات آب و خاک ایران، ۴۱، ۹۵-۱۰۱.
- [۲] امیدی، م. و محمدزاده، م.، (۱۳۸۸)، مدل‌بندی درستیابی و بیز تجربی تابع مفصل برای مدل‌بندی خشکسالی، مجله علوم دانشگاه تربیت معلم، در حال چاپ.
- [3] Bardossy, A. (2006), Copula-based Geostatistical Models for Groundwater Quality Parameters, *Water Resources Research*, **42**, W11416.
- [4] Bardossy, A. and Li, J. (2008), Geostatistical Interpolation Using Copulas, *Water Resources Research*, **44**, W07412.
- [5] Cressie, N. and Huang, H.C. (1999). Classes of Non-Separable, Spatio-Temporal Stationary Covariance Functions, *Journal of the American Statistical Association*, **96**, 1330-1340.
- [6] Feller, W. (1966), *An Introduction to Probability Theory and Its Applications*, **2**, John Wiley, New York.
- [7] Gneiting, T. (2002), Nonseparable, Stationary Covariance Functions for Space-Time Data, *Journal of the American Statistical Association*, **97**, 590-600.
- [8] Graler, B. and Pebesma, E. (2011), The Pair Copula Construction for Spatial Data, a New Approach to Model Spatial Dependency. *Procedia Environmental Sciences*, **7**, 206-211.
- [9] Joe, H. (1997), *Multivariate Models and Dependence Concepts*, Chapman and Hall.

- [10] Kent, J., Mohammadzadeh, M. and Mosammam, A. (2011), The Dimple in Gneiting's Spatio Temporal Covariance Model, *Biometrika*, **98**, 489–494.
- [11] Ma, C. (2002), Spatio-Temporal Covariance Functions Generated by Mixtures, *Mathematical Geology*, **34**, 965–974.
- [12] Ma, C. (2005), Spatio-Temporal Variograms and Covariance Models, *Advances in Applied Probability*, **37**, 706–725.
- [13] Miller, S. and Samko, G. (2001), Completely Monotonic Functions, *Integral Transformation and Special Function*, **12**, 389–402.
- [14] Nelsen, R.B. (2006), *An Introduction to Copulas*, Lecture Notes in Statistics, New York, Springer.
- [15] Sklar, A. (1959), Fonctions de Repartition an Dimensions et Leurs Marges, *Publications de l'Institut de Statistique de l'Universite de Paris*, **8** , 229–231.

بررسی ساختار وابستگی خانواده مفصل‌های مقدار فرین

امینی، م.^۱ و زنده باد، س.و.

گروه آمار دانشگاه فردوسی مشهد

چکیده

در این مقاله به معرفی خانواده‌ای از مفصل‌ها تحت عنوان مفصل‌های مقدار فرین و بررسی ساختار وابستگی آن پرداخته‌ایم. در این خانواده کمیت‌هایی مثل تاو کندال و روی اسپیرمن که در تعیین میزان وابستگی کاربرد وسیعی دارند محاسبه گردیده است. و در انتها با استفاده از نمودار کندال (K-Plot) نوع وابستگی مورد انتظار در این خانواده را بررسی کرده‌ایم.

واژه‌های کلیدی: تابع وابستگی پیکند، تاو کندال، توزیع کندال، روی اسپیرمن، مفصل، مقدار فرین، وابستگی.

۱ مقدمه

در دهه‌ی اخیر استفاده از مفصل‌ها و کاربردهایشان در علوم مختلف رشد چشمگیری داشته است. یکی از اولین تحقیقات انجام شده بر روی مقادیر فرین توسط گامبل و گلدستاین (۱۹۶۴) بوده است. محققان بسیار دیگری چون راشندورف (۱۹۹۶)، بنزو استپن (۱۹۹۷) و نلسن (۱۹۹۰) در رابطه با مفصل‌های مقدار فرین و کاربردهای آن مطالعه کرده‌اند. علاوه بر این هاتچینسون و لی (۲۰۰۲) و گادندورف و سیگرز (۲۰۰۹) برخی از ویژگی‌های مفید این مفصل‌ها را بررسی نموده‌اند. مفصل‌های مقدار فرین در زمینه‌های مختلف از جمله سرمایه‌گذاری، بیمه و علوم محیطی، پیشامدهای توام فرین نقش به‌سزایی دارند. به عنوان مثال پدیده‌ی سیل را در نظر بگیرید. سیل زمانی رخ می‌دهد که سطح آب ساکن و موج از یک سطحی تجاوز کند. بنابراین احتمال وقوع سیل نیاز به معلوم بودن توزیع توام در نقاط ماکزیمم دارد. در چنین شرایطی مفصل‌های مقدار فرین در تعیین میزان وابستگی بین مقادیر فرین مفید خواهند بود.

فرض کنید $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$ نمونه‌ای تصادفی و هم‌توزیع با تابع توزیع توام H و مفصل C باشند و X_1 و Y_1 به ترتیب دارای توزیع حاشیه‌ای F و G باشند. اگر

$$X_{(n)} = \max\{x_i\}_{i=1, \dots, n}, \quad Y_{(n)} = \max\{y_i\}_{i=1, \dots, n} \quad (1)$$

^۱email: m-amini@um.ac.ir

آنگاه با انجام محاسبات ساده ریاضی داریم:

$$H_{(n)}(x, y) = C \left[[F_{(n)}(x)]^{\frac{1}{n}}, [G_{(n)}(y)]^{\frac{1}{n}} \right]^n$$

در نتیجه اگر قرار دهیم: $F_{(n)}(x) = u$ و $G_{(n)}(y) = v$ آن گاه مفصل متناظر با $H_{(n)}(x, y)$ بنا به قضیه ۲.۳.۳ نلسن (۲۰۰۶) برابر است با $C_{(n)}(u, v) = C^n(u^{\frac{1}{n}}, v^{\frac{1}{n}})$. در حالت کلی $\lim_{n \rightarrow \infty} C_{(n)}(u, v)$ ممکن است مفصل نباشد ولی اگر مفصل C_E وجود داشته باشد به قسمی که $\lim_{n \rightarrow \infty} C_E^n(u^{\frac{1}{n}}, v^{\frac{1}{n}}) = C(u, v)$ آن گاه مفصل C را یک مفصل مقدار فرین نامند.

تعاریف زیر که بیان کننده‌ی برخی از مفاهیم وابستگی می‌باشند در ادامه مورد نیاز است.

تعریف ۱.۱.۱. X و Y به طور ربعی وابسته‌ی مثبت ${}^{\vee}PQD$ هستند، هرگاه:

$$\forall x, y \in R \quad H(x, y) \geq F(x)G(y)$$

نکته ۱.۱.۱. اگر C مفصل متناظر با توزیع توام $H(x, y)$ باشد آن گاه شرط PQD معادل است با این که $C(u, v) \geq uv$ به ازای هر $u, v \in [0, 1]$.

تعریف ۲.۱.۱. Y به طور تصادفی نسبت به X صعودی است هرگاه $P(Y \geq y | X = x)$ به ازای هر y نسبت به x صعودی باشد. همچنین X به طور تصادفی نسبت به Y صعودی است هرگاه $P(X \geq x | Y = y)$ به ازای هر x نسبت به y صعودی باشد. اگر دو ویژگی فوق برقرار باشد X و Y را وابسته‌ی رگرسیونی مثبت نامند و با نماد ${}^{\vee}PRD(X, Y)$ نمایش می‌دهند.

نکته ۲.۱.۱. اگر C مفصل متناظر با توزیع توام متغیرهای تصادفی X و Y باشد، آن گاه:

$$PRD(Y|X) \Leftrightarrow \frac{\partial C(u, v)}{\partial u} \text{ Decreasing in } u,$$

$$PRD(X|Y) \Leftrightarrow \frac{\partial C(u, v)}{\partial v} \text{ Decreasing in } v$$

قضیه ۱.۱.۱. پیکند (۱۹۸۱) مفصل C ، مفصل مقدار فرین است اگر و تنها اگر

$$C(u, v) = (uv)^{A\left(\frac{\log v}{\log uv}\right)} \quad \forall (u, v) \in (0, 1]^2 - \{(1, 1)\} \quad (2)$$

که در آن $A: [0, 1] \rightarrow [\frac{1}{4}, 1]$ تابع وابستگی پیکند نام دارد که محدب است و

$$\forall t \in [0, 1] \quad \max\{t, 1-t\} \leq A(t) \leq 1$$

[∨] Positive Quadrant Dependence

[∩] Positive Regression Dependence

کران بالا و پایین A معنای خاصی دارند، کران بالا یعنی $A(t) = 1$ متناظر با مفصل $C(u, v) = uv$ می‌باشد که استقلال را نتیجه می‌دهد، و کران پایین یعنی $A(t) = \max\{t, 1-t\}$ متناظر با مفصل $C(u, v) = \min\{u, v\}$ می‌باشد که وابستگی کامل را نتیجه می‌دهد.

نکته ۳.۱ الف) نامساوی $A(t) \leq 1$ نتیجه می‌دهد که $C(u, v) \geq uv$ یعنی مفصل‌های مقدار فرین لزوماً دارای وابستگی ربعی مثبت هستند (PQD).

ب) مفصل مقدار فرین دارای خاصیت وابستگی رگرسیون مثبت است.

۲ نتایج اصلی

در این بخش به معرفی و محاسبه برخی اندازه‌های وابستگی نظیر λ_u, λ_L روی اسپیرمن و توی کندال در این خانواده از مفصل‌ها پرداخته‌ایم و در انتها از نمودار کندال (k-Plot) برای بررسی نوع وابستگی مفصل مقدار فرین با استفاده از نرم‌افزار R، استفاده می‌کنیم.

۱.۲ اندازه‌های وابستگی

یکی از مزایای خانواده مفصل‌های مقدار فرین این است که اندازه‌های وابستگی مانند روی اسپیرمن، تاو کندال و ضرائب وابستگی دمی بالا و پایین برحسب تابع وابستگی $A(\cdot)$ بیان می‌شود.

ضرائب وابستگی دمی

اندازه‌های وابستگی دمی، وابستگی بین متغیرهای تصادفی را در گوشه‌ی یک چهارم بالایی سمت راست مربع I^2 و در گوشه‌ی یک چهارم پایین سمت چپ آن اندازه می‌گیرند. این اندازه‌ها در سال ۱۹۶۰ توسط ماساکی سبوتیا معرفی شده و شکل متداول و امروزی‌شان را می‌توان در کتاب نلسن (۲۰۰۶) یافت.

اگر $(X, Y) \sim H(x, y)$ باشند با حاشیه‌های $X \sim F(x)$ و $Y \sim G(y)$ و مفصل متناظر C ، آنگاه تعریف می‌کنیم:

$$\text{(Lower Tail Dependence)} \quad \lambda_L = \lim_{t \rightarrow 0^+} P(Y \leq G^{-1}(t) | X \leq F^{-1}(t)),$$

$$\text{(Upper Tail Dependence)} \quad \lambda_u = \lim_{t \rightarrow 1^-} P(Y > G^{-1}(t) | X > F^{-1}(t)).$$

مفصل C ، مفصل متناظر با توزیع H باشد، آنگاه

$$\lambda_L = \lim_{t \rightarrow 0^+} \frac{C(t, t)}{t}, \quad \lambda_u = 2 - \lim_{t \rightarrow 1^-} \frac{1 - C(t, t)}{1 - t}$$

گزاره: اگر C مفصل مقدار فرین باشد آنگاه

(الف)

$$\begin{aligned} \lambda_u &= \lim_{t \rightarrow 0} \frac{1 + C(1-t, 1-t) - 2 - 2t}{t} \\ &= \lim_{t \rightarrow 0} t^{-1} (2t - 1 + C(1-t, 1-t)) \\ &= 2 - \lim_{t \rightarrow 0} \frac{1}{t} (1 - C(1-t, 1-t)) \\ &= 2 - l(1, 1) \\ &= 2 \left(1 - A\left(\frac{1}{2}\right) \right) \in [0, 1] \end{aligned}$$

(ب) ضریب λ_u از صفر (استقلال، $A = 1$) تا ۱ (وابستگی کامل) تغییر می کند.

$$\begin{aligned} \lambda_L &= \lim P(U \leq u, V \leq u) \\ &= \lim u^{(2A(\frac{1}{2})-1)} \\ &= \begin{cases} 0 & A\left(\frac{1}{2}\right) > \frac{1}{2} \\ 1 & A\left(\frac{1}{2}\right) = \frac{1}{2} \end{cases} \end{aligned}$$

۲.۲ ضریب تاو کندال

ضریب وابستگی تاو کندال براساس قضیه ۵.۲.۲۰ نلسن (۲۰۰۶) متوسط میزان وابستگی درست‌نمایی یک توزیع توام با مفصل متناظر آن را اندازه‌گیری و بیان می کند. حال اگر $(X, Y) \sim H$ و مفصل متناظر C باشد آنگاه:

$$\tau_{X,Y} = 4 \int_0^1 \int_0^1 C(u,v) d(u,v) - 1 \quad (۳)$$

گزاره: اگر C ، مفصل مقدار فرین باشد در این حالت τ_C برابر است با:

$$\tau = 1 - \int_0^1 \left[1 + (1-t) \frac{A'(t)}{A(t)} \right] \left[1 - t \frac{A'(t)}{A(t)} \right] dt$$

روی اسپیرمن

ضریب وابستگی روی اسپیرمن براساس رابطه‌ی $\rho_C = 12 \int_0^1 \int_0^1 (C-uv) dudv$ متوسط فاصله بین مفصل C و مفصل استقلال $\pi(u, v) = uv$ را اندازه‌گیری و بیان می کند.

گزاره: اگر C مفصل مقدار فرین باشد آنگاه

$$\rho_s = 12 \int_0^1 \frac{dt}{(1+A(t))^2} - 3$$

توزیع کندال

تابع توزیع کندال که آن را با نماد $K(w)$ نشان می‌دهیم. تابع توزیع تجمعی متغیر تصادفی $W = H(X, Y)$ می‌باشد. در واقع $K(w)$ بیانگر احتمال پیشامد $\{H(X, Y) \leq w\}$ می‌باشد.

$K(w)$ تنها از طریق $H(x, y)$ به مفصل C وابسته است و به توزیع‌های حاشیه‌ای $F(x)$ و $G(y)$ بستگی ندارد اگر $H(x, y) = C\{F(x), G(y)\}$ و $I(A)$ نشان دهنده تابع نشانگر مجموعه A باشد، آنگاه

$$\begin{aligned} K(w) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I\{H(x, y) \leq w\} dH(x, y) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I\{C(F(x), G(y)) \leq w\} dC(F(x)G(y)) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I\{C(u, v) \leq w\} dC(u, v) \\ &= P\{C(U, V) \leq w\} \end{aligned}$$

کپرا، موگرس و جنست در سال ۱۹۹۷ یک رابطه تصادفی را به صورت زیر تعریف کردند:

زوج (X, Y) با تابع توزیع H از زوج (X^*, Y^*) با تابع توزیع H^* وابستگی مثبت کمتری دارد اگر و تنها اگر $K(w) \geq K^*(w)$ و آن را با نماد $(X, Y) <_K (X^*, Y^*)$ نشان می‌دهند، که در آن K و K^* به ترتیب تابع توزیع تجمعی $H(X, Y)$ و $H^*(X^*, Y^*)$ می‌باشد.

علت نام‌گذاری تابع توزیع K به نام تابع توزیع کندال را بطه این توزیع با تاو کندال است که به صورت زیر بیان می‌شود.

$$\tau(X, Y) = 4E\{H(X, Y)\} - 1 = 3 - 4 \int_0^1 K(w) dw \quad (۴)$$

در واقع از این رابطه نتیجه می‌شود که

$$K(w) \geq K^*(w) \Leftrightarrow \tau_K \geq \tau_K^*$$

قضیه ۱.۰۲. الف) اگر $C = uv$ مفصل استقلال باشد در این صورت $K(w) = w - w \ln w$
 ب) اگر C مفصل مقدار فرین باشد در این صورت

$$K_C(w) = w - (1 - \tau)w \log w$$

نمودار کندال (K-Plot)

نمودار کندال یک ابزار بصری برای برآورد وابستگی در نمونه تصادفی $(X_1, Y_1), \dots, (X_n, Y_n)$ می‌باشد. این نمودار بر اساس مفهوم نمودار Q-Q می‌باشد. با این وجه تمایز که با مفصل‌ها ارتباط دارد و برای حالت چندمتغیره به راحتی

قابل تعمیم است. روش ساخت این نمودار برای یک نمونه تصادفی $(X_1, Y_1), \dots, (X_n, Y_n)$ به صورت زیر می باشد.

۱. برای هر $i, 1 \leq i \leq n$ را به صورت زیر تعریف می کنیم.

$$H_i = \frac{1}{n-1} \#\{j \neq i : X_j \leq X_i, Y_j \leq Y_i\}$$

۲. مقادیر H_i را به صورت $H_{(1)} \leq \dots \leq H_{(n)}$ مرتب کنید.

۳. زوج های $(w_{i:n}, H_{(i)})$ را رسم کنید که در آن $w_{i:n}$ بنانگر امید i امین آماره مرتب در یک نمونه تصادفی به اندازه n از توزیع K تحت فرضیه صفر استقلال است.

گراف حاصل نمودار کندال با $K - Plot$ نامیده می شود. در این نمودار عدم وابستگی در نزدیک بودن نقاط به خط مستقیم تفسیر می شود.

$$\begin{aligned} w_{i:n} &= E[w_{(i)}] \\ &= n \binom{n-1}{i-1} \int_0^1 w \{K_*(w)\}^{i-1} \{1 - K_*(w)\}^{n-1} dK_*(w) \quad \forall 1 \leq i \leq n \end{aligned}$$

نمودار کندال دارای خواص زیر می باشد.

۱. براساس مقاله جنسیت و ریوست (۱۹۹۳)

$$K_{n \rightarrow \infty} \rightarrow K(w) \quad \forall 0 \leq w \leq 1$$

و از این رو

$$K_n^{-1}(p) \rightarrow K^{-1}(p) \quad 0 \leq p \leq 1$$

که در آن $K_n(w)$ تابع توزیع تجربی می باشد و به صورت زیر تعریف می شود.

$$K_n(w) = \sum_{i=1}^n \delta(w - H_i) / n$$

که در آن

$$\delta(t) = \begin{cases} 1 & t = 0 \\ \cdot & o.w \end{cases}$$

چون K_n و معکوس آن کراندارند. از این رو این همگرایی به امیدهایشان نیز قابل تعمیم است.

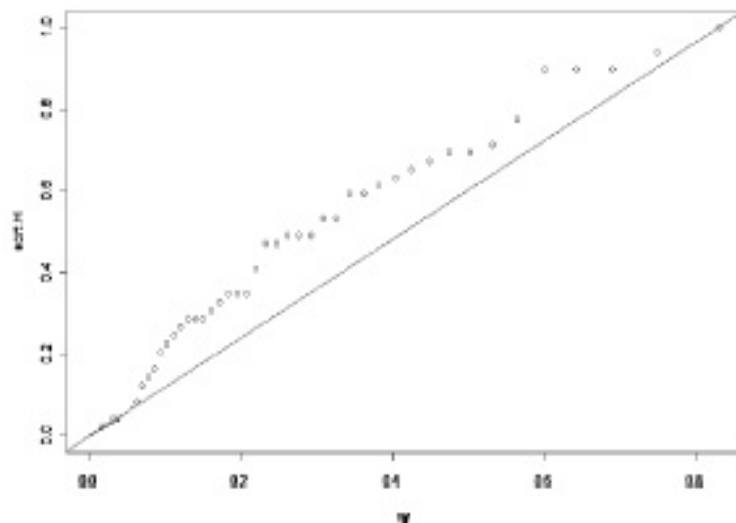
۲. برای هر $n \geq 1$ و $0 \leq p \leq 1$ فرض کنید $[np]$ کوچکترین عدد صحیح بزرگتر از np باشد. آنگاه

$$H_{[np]} = K_n^{-1}(p) \rightarrow K^{-1}(p)$$

و از این رو

$$\lim E(H_{[np]}) = \lim w_{[np];n} = K^{-1}$$

تحت فرضیه صفر (استقلال)

شکل ۱: نمودار k

۳. برای n های بزرگ زوج $(w_{i:n}, H(i))$ متمایل به قرار گرفتن بر روی منحنی $(K^{-1}(p), K(p))$ است.

۴. زمانی که $K = K$ نمودار به خطی بودن متمایل است.

۵. زمانی $\tau(X, Y)$ تمام نقاط در نمودار بر روی محور افقی E قرار می گیرند. $(p \equiv 0)$ زیرا در این صورت $K_{-1}(p) = 0 \quad \forall \quad 0 \leq p \leq 1$.

۶. زمانی که $T(X, Y) = 1$ تمام نقاط بر روی منحنی $K(p)$ قرار می گیرند. زیرا در این صورت $K_{-1}(p) \equiv p$.

۳.۲ رسم $K - Plot$ بر اساس مشاهدات مقدار فرین

همان طور که اشاره کردیم، مفصل مقادیر فرین دارای وابستگی مثبت است. حال با استفاده از نرم افزار مشاهداتی از مفصل فوق تولید کرده ایم و نمودار $K - Plot$ آن را رسم نموده ایم. همان طور که انتظار داشتیم نقاط در بالای نیمساز قرار گرفته اند و این نشان دهنده مثبت بودن وابستگی است.

مراجع

- [1] Beans, V. and Stepen, J. (1997), Distributions with Given and Moment Problems, Kluwer Academic Publishers, Dordrecht.

- [2] Caperaa, P., Fougères, A.L. and Genest, C.A. (1997), Stochastic ordering Based on a Decomposition of Kendall's Tau, in Distributions with Biven Marginal and Moment Problems, eds. V. Benes and Stephan, Dordrecht Kluwer, pp. 81–86.
- [3] Genest, C. and Jean-Claud B. (2003), Detecting Dependence with Kendall Plots, *The American Statistician*, **57(4)**, 275–284.
- [4] Genest, C. and Rivest, L.P. (1993), Statistical inference procedures for Bivariate Archimedean Copula, *Journal of the American Statistical Association*, .
- [5] Ghoudi, K., Khoudrayi, A. and Rivest, L.P. (1998), Propriétés statistiques des copules de valeurs extrêmes bidimensionnelles, *Canada. J. Statis.*, **26(1)**, 187–197.
- [6] Gudendorf F., G. and Segers, J. (2009), Extreme-Value copulas, *Math. Stat.*, **7**.
- [7] Murlimann, W. (2003), Hutchinson-Lai's conjecture for bivariate extreme value copulas, *Statistics & Probability Letters*, **61**, 191–198.
- [8] Nelsen, R. (2006), An Introduction to copulas. Springer series in statistics.
- [9] Rushendorf, L., Schweizer, B. and Taylor, M.D. (1996), Distributions with Fixed Marginals and related Topics. Institute of Mathematical Statistics, Hayward.
- [10] Sibuya, M. (1960), Bivariate extreme statistics, *Annals of the Institute of statistical Mathematics*, **11(2)**, 195–210.

بررسی وابستگی دمی بین حجم تجارت و سود بازگشتی بورس اوراق بهادار با استفاده از تابع مفصل ترکیبی

حیدری بهنوییه، ا.ا^۱،

گروه آمار دانشگاه شهید چمران اهواز

کارگر برزی، ع.ع

گروه اقتصاد دانشگاه شهید چمران اهواز

پرهام، غ.ع.ع

گروه آمار دانشگاه شهید چمران اهواز

چکیده

در این مقاله با استفاده از تابع مفصل ترکیبی به بررسی وابستگی دمی بین دو متغیر حجم تجارت و بازگشت های سود پرداخته شده است و از داده های روزانه سال های ۲۰۰۵ - ۲۰۱۱ مربوط به حجم تجارت و سود بازگشتی در بورس اوراق بهادار ایران استفاده شده است. نتایج نشان می دهد که در حالت خاص بازگشت سود های بسیار زیاد به حجم های تجارت خیلی زیاد تمایل دارند اما بازگشت سود های بسیار کم تمایلی به هیچ کدام از حجم تجارت های کم یا زیاد ندارند.
واژه های کلیدی: تابع مفصل، وابستگی دمی.

۱ مقدمه

نلسن (۱۹۹۹) برای اولین بار بیان کرد که هر ترکیب محدودی از توابع مفصل خود یک تابع مفصل است هو (۲۰۰۶) و رودریگز (۲۰۰۷) از تابع مفصل ترکیبی برای بیان وابستگی استفاده کردند همچنین کتی (۲۰۰۹) از توابع مفصل ترکیبی برای بیان وابستگی دمی بین حجم تجارت و سود در بازارهای شرق آسیا استفاده کرد.

^۱email: ahmadheydari37@yahoo.com

۲ بخش اول

۱.۲ تابع مفصل ترکیبی

نلسن (۱۹۹۹) برای اولین بار بیان کرد که هر ترکیب خطی محدب از توابع مفصل ها یک تابع مفصل است و تابع مفصل ترکیبی را به صورت زیر معرفی کرد.

$$C_{mix}(u, v) = \sum_{i=1}^m W_i C_i(u, v)$$

که در آن W_i وزن های تعریف شده برای توابع مفصل C_i با شرط $\sum_{i=1}^m W_i = 1$ می باشند. پنا و همکاران (۲۰۰۴) به این نکته اشاره کردند که توابع مفصل ترکیبی انعطاف پذیری بیشتری نسبت به توابع مفصل ساده دارند و استفاده از توابع مفصل ترکیبی روش مناسبی برای به دست آوردن وابستگی دمی بالایی و پایینی به طور همزمان است.

۲.۲ ضرایب وابستگی دمی

جوی (۱۹۹۷) ضرایب وابستگی دمی بالایی λ_u و پایینی λ_l برای دو متغیر تصادفی X و Y با توابع توزیع حاشیه ای $F(X)$ ، $G(Y)$ و تابع مفصل C با فرض $u = F(X)$ و $v = G(Y)$ از روابط زیر به دست می آیند.

$$\lambda_u = 2 - \lim_{t \rightarrow 1^-} \frac{1 - C(t, t)}{1 - t}$$

$$\lambda_l = \lim_{t \rightarrow 1^-} \frac{C(t, t)}{1 - t}$$

۳.۲ ضرایب وابستگی دمی برای توابع مفصل ترکیبی

ضرایب وابستگی دمی برای توابع مفصل ترکیبی از روابط زیر به دست می آیند.

$$\lambda_u = \sum_{i=1}^m W_i \lambda_u(i)$$

$$\lambda_l = \sum_{i=1}^m W_i \lambda_l(i)$$

برای انتخاب ضرایب تابع مفصل ترکیبی و هم چنین انتخاب تابع مفصل مناسب باید ضرایب و توابع مفصلی را انتخاب کرد که معیارهای AIC و BIC را که توسط هیروتگو آکاییک (۱۹۷۴) و اسپوارز (۱۹۷۸) به صورت زیر معرفی شده اند برای آن ها مینیمم باشند.

$$AIC = -2 \ln L(\hat{\theta}) + 2K$$

$$BIC = -2\ln L(\hat{\theta}) + 2nK$$

در اینجای L ماکزیمم درستنمایی و k تعداد پارامترهای برآورد شده n حجم نمونه است. در این مقاله از توابع مفصل کلایتون (وابستگی دمی پایینی را به خوبی نشان می دهد)، گامبل (وابستگی دمی بالایی را به خوبی نشان می دهد) و فرانک (استقلال دمی را به خوبی نشان می دهد) برای بیان وابستگی دمی استفاده شده است.

۳ نتایج تجربی

۱.۳ آزمون دیکی فولر تعمیم یافته

با توجه به این که سری های زمانی معمولاً ایستا نمی باشند، از این رو ایستایی متغیرها با استفاده از آزمون ریشه واحد دیکی فولر تعمیم یافته مورد آزمون قرار می گیرد. با توجه به مقدار p -value برابر 0.4353 مربوط به آزمون دیکی فولر برای ایستایی داده های حجم تجارت نتیجه می گیریم که این داده ها ایستا نمی باشند و این مقدار برای داده های مربوط به سود ب سود بازگشتی برابر 0.2373 می باشد که نشان می دهد ایستا نمی باشند. با توجه به نمودار مربوط به حجم تجارت و بازگشت سود مشخص می شود این داده ها ناهمسانی واریانس دارند بنابراین از فیلتر $GHARCH$ برای ایستایی داده ها استفاده می کنیم. حال با استفاده از آزمون دیکی فولر تعمیم یافته ایستایی باقیمانده ها را آزمون می کنیم که با توجه به مقادیر p -value به دست آمده 0.011 و 0.002 داده ها ایستا می باشند.

۲.۳ آزمون تصادفی بودن داده ها

برای آزمون تصادفی بودن داده ها از آزمون چرخش^۲ استفاده شده است که نتایج آن در جدول (۱) آورده شده است که با توجه به مقادیر بحرانی به دست آمده داده های حجم تجارت و سود بازگشتی تصادفی می باشند

جدول ۳: نتایج آزمون تصادفی بودن داده ها

متغیر	مقدار بحرانی
حجم تجارت	0.231
سود بازگشتی	0.653

^۲Run Test

۳.۳ برآورد پارامترهای تابع مفصل

برای به دست آوردن پارامترهای تابع مفصل برآورد شده به داده ها از روش ماکزیمم درستنمایی کانونی (CML)^۳ استفاده شده است. پارامترهای برآورد شده و لگاریتم ماکزیمم درستنمایی مربوط به توابع مفصل کلایتون، گامبل و فرانک در جدول (۳) و همچنین این مقادیر برای تابع مفصل ترکیبی در جدول (۴) به همراه ضرایب W_1 برای تابع مفصل کلایتون و W_2 برای تابع مفصل گامبل آورده شده است.

جدول ۴: پارامتر تابع مفصل، ماکزیمم درستنمایی و معیارهای AIC و BIC

تابع مفصل	مقدار بحرانی $\hat{\theta}$	ماکزیمم درستنمایی	BIC	AIC
کلایتون	0.0827	0.0034	365.74	-553.43
گامبل	1.0653	1.234	365.39	-692.16
فرانک	0.5356	1.344	106.62	-175.321

جدول ۵: پارامتر تابع مفصل، ماکزیمم درستنمایی، ضرایب و معیارهای AIC و BIC

تابع مفصل ترکیبی	$\hat{\theta}_1$	$\hat{\theta}_2$	W_1	W_2	AIC	BIC
تابع مفصل ترکیبی	1.152	0.0731	0.24	0.76	-824.46	-783.13

ضرایب وابستگی دمی بین دو متغیر تصادفی شوک قیمت نفت به دست آمده با استفاده از توابع مفصل کلایتون، گامبل و تابع مفصل ترکیبی در جدول (۵) آورده شده است.

جدول ۶: ضرایب وابستگی دمی تابع مفصل کلایتون، گامبل و تابع مفصل ترکیبی

تابع مفصل	ضریب وابستگی دمی پایینی	ضریب وابستگی دمی بالایی
کلایتون	0.00214	0
گامبل	0	0.2460
ترکیبی	0.002033	0.28696

^۳Caninocal Maximun Likelihood

بحث و نتیجه‌گیری

با توجه به مقادیر به دست آمده برای معیارهای AIC و BIC از آنجا که این معیارها برای تابع مفصل ترکیبی کم تر هستند بنابراین تابع مفصل ترکیبی برآزش بهتری به داده‌ها داشته است. بنابراین برای بیان وابستگی دمی بین حجم تجارت و سود بازگشتی از ضرایب وابستگی بدست آمده برای تابع مفصل ترکیبی استفاده می کنیم که با توجه به مقدار برای ضریب وابستگی دمی بالایی و مقدار برای ضریب وابستگی دمی پایینی مشخص می شود فرضیه دوم یعنی وجود وابستگی دمی بالایی بین حجم تجارت و سود بازگشتی تایید می شود بنابراین سودهای خیلی زیاد تمایل به همبستگی با حجم تجارت های خیلی زیاد دارند.

مراجع

- [1] Ansley, C.F. and Kohn, R. (1983), Exact Likelihood of Vector Autoregressive Moving Average Process with Missing or Aggregated Data, *Biometrika*, **70**, 275–278.
- [2] Cressie, N. (1993), Statistics for Spatial Data, Revised edition, John Wiley, New York.

مروری بر روش‌های برآورد مفصل و پارامترهای آن

راسخی نژاد، س. ۱، جمهوری، س. و نیلی ثانی، ح. ر.

گروه آمار دانشگاه بیرجند

چکیده

در این مقاله، مفصل و برخی از ویژگی‌های اساسی، نحوه برآورد مفصل و پارامترهای آن به روشهای مختلف پارامتری، نیم پارامتری و ناپارامتری تشریح و با استفاده از روش‌های شبیه‌سازی کارایی برخی از این برآوردگرها را با یکدیگر مقایسه می‌کنیم. در بخشی از این مقاله با بهبود و اصلاح نتایج مربوط به قضیه ۲ وو و همکاران (۲۰۰۶) قضیه ۳.۱ جنست (۱۹۹۳) تعمیم داده می‌شود. در پایان مفصل ارشمیدسی مناسب به قیمت سهام شرکت‌های نفت بهران و نفت پارس برازش و پارامترهای آن‌ها را برآورد و به وسیله مفاصل قیمت سهام هر یک از این دو متغیر را پیش بینی می‌کنیم.

واژه‌های کلیدی: برآورد پارامتری، برآورد ناپارامتری، برآورد نیم پارامتری، قضیه اسکالر، مفصل ارشمیدسی.

۱ مقدمه و مفاهیم مقدماتی

اکثر مدل‌های آماری بر اساس استقلال بین داده‌ها انجام می‌پذیرد. با این وجود در بسیاری از موارد متغیرهای مورد مطالعه به نوعی وابسته به یکدیگرند. ساختارهای متنوعی برای توصیف و تحلیل چنین وابستگی‌هایی معرفی شده‌اند. مدل‌های ساخته شده بر مبنای مفاصل از مهم‌ترین و پر کاربردترین این مدل‌ها می‌باشند. مفاصل که اولین بار توسط اسکالر (۱۹۵۹) معرفی شده است، ساختارهای وابستگی مورد مطالعه را از توزیع‌های حاشیه‌ای منفک می‌سازد. در اقتصاد این توابع به خوبی توانسته‌اند وابستگی متغیرهای ریسک را مدل‌سازی نمایند و از این منظر مورد توجه اقتصاددانان می‌باشند. یک رده مهم از مفاصل به نام مفاصل ارشمیدسی به دلیل انعطاف پذیری و سادگی محاسبات مورد توجه می‌باشند. در بخش دوم و سوم ضمن معرفی مفصل، برآورد مفصل و پارامترهای آن بررسی خواهند شد. در بخش چهارم داده‌های مربوط به قیمت سهام شرکت نفت بهران و نفت پارس را تجزیه و تحلیل می‌کنیم. در یک دوره زمانی ممکن است به نرخ یکی از این دو متغیر دسترسی نداشته باشیم ولی فعالان بازار سرمایه علاقه‌مند به پیش‌بینی قیمت سهام می‌باشند. روش‌های متعددی برای پیش‌بینی قیمت سهام به ویژه در سری‌های زمانی وجود دارند. در این مقاله با استفاده از مفاصل برازش شده قیمت سهام را پیش‌بینی و با استفاده از آزمون نیکویی برازش دقت آن را مورد بررسی قرار می‌دهیم.

۲ معرفی مفصل

$$1 - \forall t \in [0, 1] \quad C(0, t) = C(t, 0) = 0, \quad C(1, t) = C(t, 1) = t$$

$$2 - \forall u_1, u_2, v_1, v_2 \in I; \quad u_1 \leq u_2, \quad v_1 \leq v_2, \quad \text{تابع } C: I^2 \rightarrow I \text{ را یک مفصل نامیم اگر}$$

$$C(u_1, v_1) - C(u_1, v_2) - C(u_2, v_1) + C(u_2, v_2) \geq 0.$$

قضیه ۱.۲. (قضیه اسکالر): اگر H یک تابع توزیع توأم با توزیع‌های حاشیه‌ای F و G باشد، آنگاه مفصلی مانند C وجود دارد به قسمی که

$$\forall x, y \in \bar{R}, \quad H(x, y) = C(F(x), G(y)).$$

اگر F و G پیوسته باشند، C یکتاست در غیر این صورت C به طور یکتا روی $\text{Ran}F \times \text{Ran}G$ تعیین می‌شود.

مثال ۱.۲. فرض کنید متغیرهای X و Y با تکیه‌گاه I^2 دارای تابع توزیع توأم $F(x, y) = \frac{1}{4}(x+y)$ می‌باشند. توزیع‌های حاشیه‌ای عبارتند از $F_1(x) = \frac{1}{4}(x+1)$ و $F_2(y) = \frac{1}{4}(y+1)$. متغیرهای X و Y پیوسته نیستند. با این وجود و بر اساس قسمت دوم قضیه اسکالر می‌توان بر روی ناحیه $[\frac{1}{4}, 1] \times [\frac{1}{4}, 1]$ مفصل منحصر به فرد C را به شکل زیر تعریف کرد

$$F(x, y) = \max\left(\frac{1}{4}(x+1) + \frac{1}{4}(y+1) - 1, 0\right) = C(F_1(x), F_2(y)).$$

اما در خارج این ناحیه مفصل C منحصر به فرد نمی‌باشد.

اگر دامنه C زیر مجموعه‌ای از I^2 باشد C را یک زیرمفصل می‌نامیم.

لازم به ذکر است که در این مقاله ما خود را محدود به متغیرهای پیوسته کرده‌ایم.

اگر F^{-1} و G^{-1} توابع معکوس F و G باشند آنگاه برای هر (u, v) در دامنه C داریم

$$C(u, v) = H(F^{-1}(u), G^{-1}(v)).$$

برای مطالعه بیشتر مفصل و ویژگی‌های اساسی آن علاقه‌مندان را به نلسن (۲۰۰۶) و جو (۱۹۹۷) ارجاع می‌دهیم. چریوبینی و همکاران (۲۰۰۴) و پانجار (۲۰۰۶) دو مرجع مناسب برای آشنایی با کاربردهای مفاصل در مسائل اقتصادی و مالی می‌باشند. یک کلاس بزرگ از مفاصل، مفاصل ارشمیدسی می‌باشند. از آنجا که اعضای آن به راحتی ساخته می‌شوند و ساختار وابستگی‌هایشان تنها توسط یک تابع حقیقی مقدار یک متغیره مشخص می‌شود، بنابراین مورد توجه بسیاری قرار گرفته‌اند.

تعریف ۱.۲. فرض کنید φ تابعی پیوسته، اکیداً نزولی از I به $[0, \infty]$ باشد به طوری که $\varphi(1) = 0$. مفصلی به فرم $C(u, v) = \varphi^{-1}(\varphi(u) + \varphi(v))$ را مفصل ارشمیدسی و φ را تابع مولد مفصل ارشمیدسی گویند. زمانیکه $\varphi(0) = \infty$ تابع φ را مولد اکید گویند.

سه زیر کلاس مهم از مفصل ارشمیدسی عبارتند از

(الف) مفصل کلیتون. این مفصل توسط کلیتون (۱۹۸۷) معرفی و بعدها توسط کوک و جانسون (۱۹۸۱) و آکاس (۱۹۸۲) مورد مطالعه و بررسی قرار گرفت. در این خانواده $\alpha \in [-1, \infty) \setminus \{0\}$ و $\varphi(t) = \frac{t^{-\alpha}-1}{\alpha}$ و لذا $C_\alpha(u, v) = (u^{-\alpha} + v^{-\alpha} - 1)^{-\frac{1}{\alpha}}$.

(ب) مفصل فرانک. این مفصل توسط فرانک (۱۹۷۹) مطرح و ویژگی های آن توسط نلسن (۱۹۸۶) و جنست (۱۹۸۷) مورد ارزیابی قرار گرفته است. در این خانواده

$$\varphi(t) = \ln\left(\frac{1-e^{-\alpha t}}{1-e^{-\alpha}}\right), \quad \alpha \in (-\infty, \infty) \setminus \{0\}$$

$$C_\alpha(u, v) = \frac{-1}{\alpha} \ln\left(1 + \frac{(e^{-\alpha u}-1)(e^{-\alpha v}-1)}{e^{-\alpha}-1}\right).$$

(ج) مفصل گامبل. این مفصل توسط هوگارد (۱۹۸۶) مورد بررسی قرار گرفته است. در این خانواده $\varphi(v) = (-\ln(v))^{\alpha+1}$ که در آن $\alpha \geq 0$ می باشد. بنابراین

$$C_\alpha(u, v) = \exp\left\{-\left[(-\ln u)^{\alpha+1} + (-\ln v)^{\alpha+1}\right]^{\frac{1}{\alpha+1}}\right\}.$$

۳ برآورد مفصل و پارامترهای آن

فرض کنید متغیرهای تصادفی پیوسته X_1, \dots, X_m دارای تابع توزیع توأم

$$F(\underline{x}, \alpha_1, \dots, \alpha_m, \theta) = C(F_1(x_1, \alpha_1), \dots, F_m(x_m, \alpha_m), \theta), \quad \underline{x} = (x_1, \dots, x_m).$$

که در آن $\alpha_1, \dots, \alpha_m$ پارامترهای توزیع حاشیه ای و θ پارامتر وابستگی می باشند. چگالی توأم این متغیرها عبارت است از

$$f(\underline{x}, \alpha_1, \dots, \alpha_m, \theta) = c(F_1(x_1, \alpha_1), \dots, F_m(x_m, \alpha_m), \theta) \prod_{j=1}^m f_j(x_j).$$

که در این رابطه c تابع چگالی مفصل می باشد $(c(u, v) = \frac{\partial^2 C(u, v)}{\partial u \partial v})$. در این بخش مفصل و پارامترهای آن را به روش های پارامتری، نیم پارامتری و ناپارامتری برآورد می کنیم.

۱.۳ برآورد پارامتری

از جمله روش های پارامتری می توان به روش های MLE و IFM اشاره کرد. برآورد به روش MLE از ماکسیم کردن تابع چگالی توأم نسبت به همه پارامترها به دست می آید. در روش IFM ابتدا پارامترهای $\alpha_1, \dots, \alpha_m$ به روش ماکسیم درستنمایی توزیع های حاشیه ای برآورد گردیده و سپس پارامتر وابستگی θ از ماکسیم کردن تابع چگالی توأم هنگامی که به جای مقادیر α_i ها برآورد آن ها قرار داده شده است، به دست می آید. با استفاده از ماتریس های کواریانس مجانبی MLE و IFM و همچنین مقایسه توان دوم خطاها به کمک شبیه سازی مونت کارلو می توان کارایی برآوردگرهای حاصل از این دو روش را با هم مقایسه کرد. به وسیله شبیه سازی در نرم افزار R میانگین مربعات خطای

دو برآوردگر مربوط به پارامتر مفصل را برای مفصل کلیتون با توزیع‌های حاشیه‌ای نرمال استاندارد و $N(0, 2)$ ، به دست آورده و سپس با محاسبه نسبت کارایی‌ها، به مقایسه این دو برآوردگر می‌پردازیم. این شبیه‌سازی شامل نمونه‌ای به حجم ۱۰۰۰ با ۲۰ بار تکرار بوده است. با توجه به این شبیه‌سازی $\frac{MSE_{MLE}}{MSE_{IFM}} = 1,25357$ می‌باشد که بیانگر این است که روش IFM برای برآورد پارامتر مفصل نسبت به روش MLE کارتر می‌باشد. (جزئیات این شبیه‌سازی در راسخی نژاد (۱۳۹۰) آمده است.)

۲.۳ برآورد نیم‌پارامتری

از جمله روش‌های نیم‌پارامتری روش CML می‌باشد. در این روش توابع توزیع حاشیه‌ای را توسط توزیع تجربی برآورد کرده و سپس پارامتر مفصل را به روش ماکسیمم درست‌نمایی برآورد می‌کنیم. ویژگی‌های این روش توسط چریوبینی و همکاران (۲۰۰۴) بررسی شده است. تقریب رتبه Z -برآوردگر روشی است که در آن برآورد پارامتر مفصل از حل معادله $\sum_{i=1}^n w\{\hat{F}_{n1}(X_{1i}), \dots, \hat{F}_{nm}(X_{mi}), \theta\} = 0$ به دست می‌آید که در این رابطه \hat{F}_{nj} توابع توزیع تجربی و $w(\cdot, \theta)$ یک تابع R^d مقدار d بعد فضای پارامتر روی $(0, 1)^m$ می‌باشد به طوری که اعضای آن در رابطه $\int w_k(u, \theta) dC(u) = 0$ صدق کنند. شرح بیشتری در خصوص روش Z -برآوردگر را می‌توانید در تسوکاها را (۲۰۰۵) بیابید. یک روش دیگر برای برآورد مفصل، مینیمم فاصله می‌باشد. اگر D مفصل متناظر با توزیع توأم و $\{C_\theta, \theta \in \Theta \subset R^d\}$ خانواده پارامتری برازش داده شده به داده‌ها باشد، $\hat{\theta}$ مقداری است که باعث مینیمم شدن فاصله بین D و C_θ می‌شود.

۳.۳ برآورد ناپارامتری

می‌توان مفصل را به روش ناپارامتری نیز برآورد کرد. دهیولز (۱۹۷۹) برآوردگر مفصل تجربی را به صورت

$$C_n(u, v) = \frac{1}{n} \sum_{i=1}^n I(\hat{U}_i \leq u, \hat{V}_i \leq v)$$

بیان کرده است که در این رابطه \hat{U}_i و \hat{V}_i توابع توزیع تجربی توزیع‌های حاشیه‌ای می‌باشند. این برآوردگر به طور مجانبی با برآوردگری که به طور مستقیم از قضیه اسکالر به دست می‌آید، معادل است. (املکا و همکاران (۲۰۰۹))

در قضیه زیر ضمن اصلاح نتایج وو و همکاران (۲۰۰۶) قضیه (۳.۱) جنست (۱۹۹۳) را تعمیم می‌دهیم. این قضیه می‌تواند نقش کلیدی در استنباط پارامتری مفاصل ارشمیدی ایفا کند.

قضیه ۱.۳. فرض کنید (U_1, \dots, U_n) یک بردار تصادفی n بعدی با توزیع‌های حاشیه‌ای یکنواخت و تابع توزیع توأم تعریف شده به صورت مفصل ارشمیدی

$$C(u_1, \dots, u_n) = \varphi^{-1}(\varphi(u_1) + \dots + \varphi(u_n))$$

باشد. همچنین فرض کنید

$$S_k = \sum_{j=1}^k \varphi(U_j) / \sum_{j=1}^{k+1} \varphi(U_j), \quad k = 1, \dots, n-1, \quad (1)$$

$$T = C(U_1, \dots, U_n) = \varphi^{-1} \left(\sum_{k=1}^n \varphi(U_k) \right). \quad (2)$$

آنگاه تابع توزیع توأم $(S_1, S_2, \dots, S_{n-1}, T)$ به صورت زیر است

$$\begin{aligned} h(s_1, \dots, s_{n-1}, t) &= \frac{\partial^n P(S_1 \leq s_1, \dots, S_{n-1} \leq s_{n-1}, T \leq t)}{\partial s_1 \dots \partial s_{n-1} \partial t} \\ &= s_1 s_2 s_3 \dots s_{n-1}^{-1} \varphi^{-1(n)}(\varphi(t)) [\varphi(t)]^{n-1} \varphi'(t), \end{aligned}$$

که در آن $\varphi^{-1(n)}$ مشتق مرتبه n معکوس مولد مفصل ارشمیدسی، φ^{-1} می باشد.

برهان. با توجه به روابط (۱) و (۲) داریم

$$\varphi(U_1) = S_1 \dots S_{n-1} \varphi(T),$$

$$\varphi(U_2) = (1 - S_1) S_2 \dots S_{n-1} \varphi(T),$$

$$\varphi(U_n) = (1 - S_{n-1}) \varphi(T),$$

ماتریس ژاکوبین $n \times n$ تبدیلات، $J = \left(\frac{\partial(u_1, \dots, u_n)}{\partial(s_1, \dots, s_{n-1}, t)} \right)$ ، دارای درایه های زیر است:

اعضای سطر اول ماتریس ژاکوبین به صورت

$$\partial u_1 / \partial s_k = \prod_{j=1, j \neq k}^{n-1} s_j \times \varphi(t) / \varphi'(u_1), \quad k = 1, \dots, n-1,$$

$$\partial u_1 / \partial t = \prod_{j=1}^{n-1} s_j \times \varphi'(t) / \varphi'(u_1).$$

طور مشابه اعضای سطر دوم به صورت زیر به دست می آیند:

$$\partial u_2 / \partial s_k = (1 - s_1) \prod_{j=2, j \neq k}^{n-1} s_j \times \varphi(t) / \varphi'(u_2), \quad k = 2, \dots, n-1,$$

$$\partial u_2 / \partial t = (1 - s_1) \prod_{j=2}^{n-1} s_j \times \varphi'(t) / \varphi'(u_2).$$

به طور کلی اعضای m امین سطر ($m \geq 3$) به فرم زیر هستند:

$$\partial u_m / \partial s_k = 0, \quad k \leq m-2,$$

$$\partial u_m / \partial s_k = (1 - s_{m-1}) \prod_{j>m, j \neq k}^{n-1} s_j \times \varphi(t) / \varphi'(u_m),$$

$$m-2 < k \leq n-1,$$

$$\partial u_m / \partial t = (1 - s_{m-1}) \prod_{j=m}^{n-1} s_j \varphi'(t) / \varphi'(u_m).$$

همه اعضای سطر آخر ماتریس ژاکوبین به جز دو عضو آخر آن صفر می باشند و دو عضو آخر عبارتند از

$$\partial u_n / \partial s_{n-1} = -\varphi(t) / \varphi'(u_n), \quad \partial u_n / \partial t = (1 - s_{n-1}) \varphi'(t) / \varphi'(u_n).$$

فرص کنید $n = 2$ باشد. بنابراین

$$S_1 = \frac{\varphi(U_1)}{\varphi(U_1) + \varphi(U_2)}, \quad T = \varphi^{-1}(\varphi(U_1) + \varphi(U_2)),$$

و در نتیجه

$$\varphi(U_1) = S_1 \varphi(T), \quad \varphi(U_2) = (1 - S_1) \varphi(T),$$

لذا دترمینال ماتریس ژاکوبین به صورت زیر به دست می آید

$$\begin{aligned} |J| &= \frac{\varphi(t)}{\varphi(u_1)} \cdot \frac{(1-s_1)\varphi'(t)}{\varphi'(u_2)} + \frac{s_1\varphi'(t)}{\varphi'(u_1)} \cdot \frac{\varphi(t)}{\varphi'(u_2)} \\ &= \frac{\varphi(t)\varphi'(t) - s_1\varphi(t)\varphi'(t) + s_1\varphi(t)\varphi'(t)}{\varphi'(u_1)\varphi'(u_2)} \\ &= \frac{s_1\varphi(t)\varphi'(t)}{\varphi'(u_1)\varphi'(u_2)}. \end{aligned}$$

ادامه مبحث به استقرا ثابت می شود. چگالی مفصل ارشمیدی عبارت است از:

$$c(u_1, \dots, u_n) = \frac{\partial^n C(u_1, \dots, u_n)}{\partial u_1 \dots \partial u_n} = \varphi^{-1(n)}(\varphi(C(u_1, \dots, u_n))) \prod_{k=1}^n \varphi'(u_k).$$

ژاکوبین محاسبه شده، تابع چگالی توأم به صورت زیر حاصل می شود

$$h(s_1, \dots, s_{n-1}, t) = |J| \times \varphi^{-1(n)}(\varphi(t)) \prod_{k=1}^n \varphi'(u_k)$$

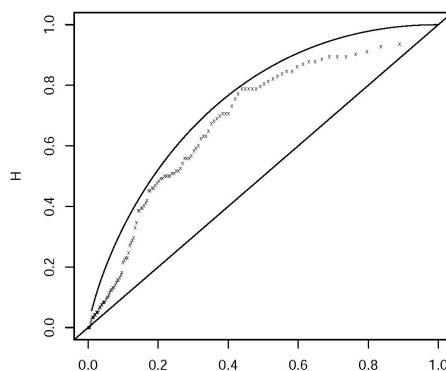
□

$$= s_1 s_2 s_3 \dots s_{n-1} \varphi^{-1(n)}(\varphi(t)) [\varphi(t)]^{n-1} \varphi'(t).$$

۴ مثال کاربردی

در این بخش داده‌های مربوط به نرخ سهام نفت بهران و نفت پارس، دو شرکت پذیرفته شده در بورس اوراق بهادار را مورد بررسی قرار می دهیم. این داده‌ها مربوط به دوره زمانی ۶ فروردین تا ۶ مهر سال ۱۳۹۰ می باشند. در یک دوره زمانی به دلایل متعددی ممکن است به نرخ یکی از این دو متغیر دسترسی نداشته باشیم. اما فعالان بازار سرمایه علاقه مند به پیش‌بینی قیمت سهام می باشند. روش‌های مختلفی برای پیش‌بینی قیمت سهام به ویژه در سری‌های زمانی وجود دارد. ولی در این قسمت علاقه‌مندیم که با استفاده از مفاصل این پیش‌بینی را انجام دهیم. بنابراین برای انتخاب مناسب‌ترین مفصل لازم است وابستگی بین این دو متغیر را مورد بررسی قرار دهیم. یک روش گرافیکی برای بررسی وابستگی بین متغیرها توسط جنست و بیوس (۲۰۰۳) مطرح شده است که $K-Plot$ نامیده شده است. این روش گرافیکی بر اساس جفت‌های $(W_{i:n}, H_{(i)})$ می باشد. $H_{(i)}$ آماره‌های مرتب متناظر با

$$H_i = \frac{1}{n-1} \#\{j \neq i : X_j \leq X_i, Y_j \leq Y_i\}$$



شکل ۱: $K - Plot$ مربوط به نرخ نفت بهران و نفت پارس

و امید ریاضی i امین آماره مربوط به یک نمونه تصادفی به سائز n از متغیر تصادفی $W = C(u, v) = H(X, Y)$ تحت فرضیه استقلال بین U و V (یا X و Y) به شرح زیر می باشد:

$$W_{i:n} = n \binom{n-1}{i-1} \int_0^1 w k \cdot (w) \{K \cdot (w)\}^{i-1} \{1 - K \cdot (w)\}^{n-i} dw,$$

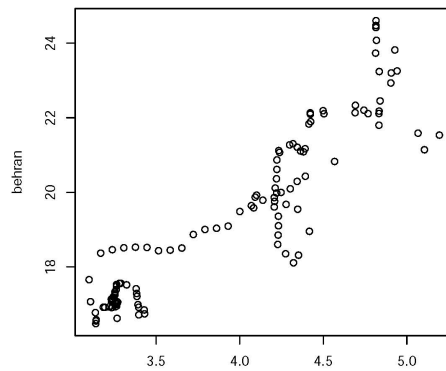
که در آن

$$K \cdot (w) = P(UV \leq w) = \int_0^1 P(U \leq \frac{w}{v}) dv = \int_0^w dv + \int_w^1 \frac{w}{v} dv = w - w \log(w)$$

و $k \cdot (.)$ تابع چگالی متناظر می باشد.

با استفاده از نرم افزار R مقادیر $W_{i:n}$ و $H_{(i)}$ را محاسبه و $K - Plot$ را رسم می کنیم. در $K - Plot$ هر انحرافی از قطر اصلی بیانگر وابستگی بین متغیرها می باشد. زمانی که مجموعه نقاط روی محور $W_{i:n}$ قرار گیرد وابستگی منفی کامل و زمانی که مجموعه نقاط روی منحنی $K \cdot (w)$ قرار گیرد وابستگی مثبت کامل رخ می دهد. همان طور که در شکل ۱ ملاحظه می شود، بین دو متغیر نرخ سهام نفت بهران و نفت پارس وابستگی مثبت وجود دارد. لذا به دنبال برازش مناسب ترین مفصل ارشمیدسی از میان سه مفصل کلیتون، فرانک و گامبل به داده های نرخ نفت بهران و نرخ نفت پارس می باشیم. یکی از روش های بررسی مناسب بودن مفصل برازش داده شده به داده ها استفاده از آزمون نیکویی برازش است. در این بخش آزمون نیکویی برازش را بر اساس مفصل تجربی با توجه به آماره کلموگروف اسمیرنوف انجام می دهیم. نتایج این آزمون در جدول ۱ آمده است. با مقایسه مقادیر آماره T_n با آماره جدول کلموگروف، مفصل کلیتون و فرانک برازش های مناسبی به داده ها می باشند.

جدول ۱: آماره های کلموگروف مربوط به مدل های برازش داده شده به مجموعه داده های بورس



شکل ۲: نمودار پراکنش نرخ نفت پارس در مقابل بهران

خانواده	$\hat{\alpha}$	T_n
کلیتون	۵/۲۶۳	۰/۰۸۱۳
فرانک	۱۲/۶۳۵	۰/۰۷۳۲
گامبل	۲/۶۳۱	۰/۱۰۵۷

ضریب وابستگی دم بالایی و پایینی برای مفصل کلیتون به ترتیب $\lambda_U = 0$ و $\lambda_L = 2/\alpha = 0/87$ و برای مفصل فرانک $\lambda_U = \lambda_L = 0$ می باشد. یعنی در مورد مفصل کلیتون ازدحام داده‌ها در گوشه پایین سمت چپ بیشتر می باشد ولی در مفصل فرانک ازدحام داده‌ها در گوشه‌ها وجود ندارد. با توجه به شکل ۲ و وابستگی‌های دمی، مفصل کلیتون برازش مناسبی برای داده‌های بورس به نظر می رسد. حال به دنبال پیش بینی نرخ متغیر مجهول می باشیم. برای این کار در ابتدا مفصل شرطی را با توجه به مشخص بودن یکی از این دو متغیر توسط رابطه زیر به دست می آوریم.

$$\begin{aligned}
 P(U_2 \leq u_2 | U_1 = u_1) &= \lim_{\delta \rightarrow 0} \frac{P(U_2 \leq u_2, U_1 \in [u_1 - \delta, u_1 + \delta])}{P(U_1 \in [u_1 - \delta, u_1 + \delta])} \\
 &= \lim_{\delta \rightarrow 0} \frac{C(u_1 + \delta, u_2) - C(u_1 - \delta, u_2)}{2\delta} \\
 &= \frac{\partial}{\partial u_1} C(u_1, u_2)
 \end{aligned}$$

با استفاده از این مفصل شرطی، می توان یک فاصله اطمینان برای U_2 به دست آورد. با توجه به توزیع U_2 و این فاصله اطمینان نرخ مربوط به این متغیر را پیش بینی می کنیم. حال با فرض اینکه مفصل کلیتون برازش مناسبی برای داده‌ها می باشد می خواهیم نرخ سهام نفت بهران را در روز ۱۹ دی ۱۳۹۰ پیش بینی کنیم. نرخ سهام نفت پارس در

آن روز ۳۹۶۶ ریال می باشد. فاصله اطمینان برای U_2 به طریق زیر به دست می آید.

$$\left(\left(\frac{71}{182}\right)^{-5/26} + u_2^{-5/26} - 1\right)^{\frac{-1}{5/26}} - 1 \left(\frac{71}{182}\right)^{-6/26} = 0.05$$

$$\left(\left(\frac{71}{182}\right)^{-5/26} + u_2^{-5/26} - 1\right)^{\frac{-1}{5/26}} - 1 \left(\frac{71}{182}\right)^{-6/26} = 0.95$$

در نتیجه $u_2 \in (0.25, 0.69)$ و با توجه به اینکه $X_2 \sim \Gamma(94/8, 0.005)$ می باشد بنابراین با محاسبه چندک ها داریم $x_2 \in (17587, 19854)$ در نتیجه نرخ سهام نفت بهران عددی بین ۱۷۵۸۷ و ۱۹۸۵۴ ریال می باشد.

بحث و نتیجه گیری

تعیین وابستگی در مدل های چند متغیره یک مسئله مهم در آمار است. مفاصل یک روش مفید برای مدل بندی وابستگی در مدل های چند متغیره می باشند. در این مقاله در ابتدا به معرفی مفصل و ویژگی های آن پرداختیم. به برخی از روش ها برای برآورد مفصل و پارامترهای آن اشاره کردیم. برای داده های نرخ سهام نفت بهران و نفت پارس با استفاده از $K - Plot$ نوع وابستگی آن ها را مورد بررسی و سپس مناسب ترین مفصل را به داده ها برازش دادیم. با توجه به مفصل برازش داده شده قیمت سهام متغیر مجهول را پیش بینی کردیم. بنابراین می توان از مفاصل در بورس برای پیش بینی نرخ سهام استفاده کرد. مفاصل امروزه در علوم مختلفی از جمله اقتصاد، تحلیل بقا، امور مالی و بیمه کاربرد دارند.

مراجع

- [۱] راسخی نژاد، س. (۱۳۹۰)، استنباط آماری برای مفاصل، پایان نامه کارشناسی ارشد، دانشگاه بیرجند.
- [2] Cherubini, U., Luciano, E. and Vecchiato, W. (2004), Copula Methods in Finance, John Wiley, Sons Ltd, England.
- [3] Clayton, D.G. (1978), A Model for Association in Bivariate Life Table and its Application in Epidemiological Studies of Familial Tendency in Chronic Disease Incidence, *Biometrika*, **65**, 141–151.
- [4] Cook, R.L. and Johnson, M.L. (1981), A Family of Distributions for Modeling Nonelliptically Symmetric Multivariate Data. *Journal of the Royal Statistical Society*, **43**, 210–218.

- [5] Deheuvels, P. (1979), La fonction de dépendance empirique et ses propriétés .Un test non paramétrique d'indépendance, *Acad. Roy. Belg. Bull. Cl. Sci*, **65**, 274–292.
- [6] Frank, M.J. (1979), On Simultaneous Associativity of $F(x,y)$ and $x+y -F(x,y)$, *Equationes Mathematicae*, **19**, 194–226.
- [7] Genest, C. (1987), Frank's family of bivariate distributions, *Biometrika*, **74**, 549–555.
- [8] Genest, C. and Rivest, L.P. (1993), Statistical Inference Procedures for Bivariate Archimedean Copula, *Journal of the American Statistical Association*, **88**, 1034–1043.
- [9] Genest, C. and Favre, A.C. (2007), Everything you always wanted to know about copula modeling but were afraid to ask, *Journal of Hydrologic Engineering*, **12**, 347–368.
- [10] Hougaard, P. (1986), Survival Models for Heterogeneous Populations Derived From Stable Distributions, *Biometrika*, **73**, 387–396.
- [11] Joe, H. (1997), Multivariate Models and Dependence Concepts, Chapman Hall, London.
- [12] Mari, D.D. and Kotz, S. (2001), Correlation and Dependence, Imperial College Press, London.
- [13] Nelsen, R.B. (1986), Properties of a one-parameter family of bivariate distributions with specified marginals, *Comm Statist Theory Methods*, **15**, 3277–3285.
- [14] Nelsen, R.B. (2006), An introduction to copulas, Lecture Notes in Statistics. Springer-Verlag, New York.
- [15] Oakes, D. (1982), A Model for Association in Bivariate Survival Data, *Journal of the Royal Statistical Society*, **44**, 414–422.
- [16] Omelka, M., Gijbels, I. and Veraverbeke, N. (2009), Semiparametric Estimation in Copula Model-Improved Kernel Estimation of Copula, *Institute of Mathematical Statistics*, **37**, 3023–3058.
- [17] Tsukahara, H. (2005), Semiparametric Estimation in Copula Models, *The Canadian Journal of Statistics*, **33**, 357–375.

- [18] Wu, F., Valdez, E.A. and Sherris, M. (2006), Simulating Exchangeable, Multivariate Archimedean Copulas and its Applications, *Communications in Statistics-Simulation and Computation*, **36**, 1019–1034.

آیا اطلاعات نامتقارن در بازار محصولات بیمه ایران وجود دارد؟

داریوش همدانی ح. و صفری کته سری، ۱.۵

گروه آمار دانشگاه شهید بهشتی

چکیده

در این مقاله با مدل‌بندی فراوانی ادعاهای بیمه‌ای و نوع پوشش بیمه‌ای انتخابی افراد بیمه‌گذار در مورد وجود اطلاعات نامتقارن بحث می‌کنیم. با استفاده از توابع مفصل همبستگی متغیرهای ذکر شده را نشان می‌دهیم و از نتایج این مدل‌ها برای رد یا قبول وجود اطلاعات نامتقارن در بازار بیمه استفاده می‌کنیم. وجود اطلاعات نامتقارن در بازار بیمه هر کشوری می‌تواند اثرات جبران‌ناپذیری روی بنگاه‌های اقتصادی و انجام تعهدات شرکت‌های بیمه داشته باشد. با بررسی وجود اطلاعات نامتقارن در این بازارها می‌توان راهکارهای مناسب از جمله طبقه‌بندی ریسک افراد را برای جلوگیری از رفتارهای کژگزینی و کژمنشی بیمه‌گذاران به دست آورد. **واژه‌های کلیدی:** اطلاعات نامتقارن، بیمه اتومبیل، کژگزینی، کژمنشی، مدل‌های مفصل.

۱ مقدمه

پس از آکرلوف (۱۹۷۰) اطلاعات نامتقارن^۲ به یکی از مهمترین مفاهیم نظری تأثیرگذار در توصیف کنش و واکنش مجموعه نیروهای بازار تبدیل شده است. به‌طور شفاف اطلاعات نامتقارن مسأله‌ای پیچیده است، چون برای بنگاه‌های اقتصادی فرصت تقلب و دروغگویی در معاملات را فراهم می‌کند که می‌تواند سبب نا کارآمدی بازار یا حتی سقوط آن شود. در بازارهای بیمه اطلاعات نامتقارن زمانی که یک شکاف اطلاعاتی بین بیمه‌گذار و بیمه‌گر وجود داشته باشد رخ می‌دهد به‌طوری که بیمه‌گر اطلاع کمتری در باره عوامل ریسک بیمه‌گذار داشته باشد. در نتیجه عدم تقارن در دستیابی به ریسک واقعی فرصت‌هایی را برای رفتارهای استراتژیک مانند کژگزینی^۳ و کژمنشی^۴ فراهم می‌کند. تلاش‌های زیاد و هزینه‌برداری به‌صورت تجربی در تایید وجود اطلاعات نامتقارن در بازارهای بیمه مختلف انجام شده است. اگرچه مطالعات انجام شده نشان دهنده وجود اطلاعات نامتقارن در تعداد زیادی از بازارهای بیمه مانند بیمه خدمات درمانی و بیمه عمر است، اما با شواهد موجود نمی‌توان وجود اطلاعات

^۱email: ssafari.hadi@gmail.com

^۲Information Asymmetric

^۳Selection Adverse

^۴Hazard Moral

نامتقارن را در بازار بیمه اتومبیل نتیجه‌گیری کرد. در گذشته طیف وسیعی از تحقیقات شواهدی مبنی بر وجود اطلاعات نامتقارن در بازار بیمه اتومبیل از جمله در فرانسه، کانادا و ژاپن را نشان ندادند. این مطالعات باور عدم وجود اطلاعات نامتقارن در بازار بیمه اتومبیل را تقویت می‌کردند. هرچند اخیراً این باور به چالش کشیده شده است. به عنوان مثال کوهن (۲۰۰۵) شواهدی در وجود اطلاعات نامتقارن در بازار بیمه اتومبیل اسرائیل یافت. در حقیقت این یافته را نمی‌توان به صورت عام پذیرفت زیرا بازار اسرائیل منحصر به فرد است، از این حیث که این بازار فاقد سیستم اشتراک اطلاعات است تا شرکت‌های بیمه اطلاعات خود درباره عوامل تأثیرگذار در ریسک بیمه‌نامه و سایر پیشنهادات فنی برای مشاهده کثرتاری با کاربردهای وسیع‌تر را به اشتراک بگذارند. دینانی، میکود و ده‌چور (۲۰۱۰) داده‌های واقعی ۳-ساله فرانسه را تحلیل کردند و دلایل محکمی مبنی بر کثرتاری رانندگان بخصوص رانندگان با ۱۵-۵ سال سابقه رانندگی را یافتند. با یک روش مشابه لی، لیو و یه (۲۰۰۷) و همچنین وانگ، چانگ و ژنگ (۲۰۰۸) کثرتاری رانندگان را هنگام وضع قانون افزایش مالیات در تایوان گزارش کردند (کیم و همکاران، ۲۰۰۹).

اطلاعات نامتقارن در همه بازارها وجود دارد. به هر ترتیب وقتی محصولی فروخته یا خدماتی ارائه شد، اغلب فروشنده آگاهی در مورد ترجیحات خریدار ندارد، از جمله حداکثر مبلغی که خریدار حاضر است بپردازد تا نیاز خود را برطرف سازد. به طور مشابه خریدار نیز عموماً آگاهی درباره فن‌آوری تولید یا هزینه نهایی فروشنده ندارد. با این وجود اغلب اوقات این عدم تقارن بی‌تأثیر است. به‌طور کلی در محیط رقابتی^۵، فروشنده منفعت خاصی از یک آگاهی جزئی در مورد مبلغی که خریدار حاضر است بپردازد نمی‌برد، زیرا وی باید قیمت رقابتی را لحاظ کند. خریدار نیز نیازی به داشتن آگاهی درباره فن‌آوری به کار رفته در تولید محصول ندارد، زیرا همه اطلاعاتی که وی نیاز دارد در قیمت محصول نهفته است. از این رو عموماً اطلاعات نامتقارن هم پراهمیت و هم کم اهمیت است. یک ویژگی مهمی که در مثال‌های بالا مشاهده می‌شود این است که اطلاعات پنهان از دسترس طرفین، خصوصی است. فاگارت (۱۹۹۶) تحت پیش‌فرض‌های ضعیف ثابت کرد که رقابت در قالب اطلاعات خصوصی (که ما از آن با نام "اطلاعات نامتقارن نامربوط" یاد می‌کنیم) ما را به سمت تعادلی هدایت می‌کند که اغلب کاراست. به عنوان مثال در بازار ماشین‌های اسقاطی پرداختی خریدار به کیفیت ماشین بستگی دارد که فقط فروشنده از آن آگاه است. به‌طور مشابه سود یک بیمه‌گر به ریسک بیمه‌گذاری که بیمه‌نامه را از وی می‌خرد بستگی دارد (چیاپوری و همکاران، ۲۰۰۶).

۲ سیری در پیشینه نظری و تجربی

یک راه‌کار معمول برای مطالعه اطلاعات نامتقارن آزمون کردن همبستگی بین انتخاب یک قرارداد و رخداد یا شدت یک تصادف از روی مشاهدات است. تحت ریسک‌های حاصل از کژمنشی، افراد "پر ریسک" به کسانی

^۵Environmentn Competitive

اطلاق می‌شود که با احتمال زیادی هم قراردادی با پوشش بیمه‌ای بالا انتخاب می‌کنند و هم با احتمال زیادی یک تصادف خواهند داشت. داستان پیدایش کژمنشی به کژگزینی بسیار نزدیک است و رابطه بین آنها رابطه معکوس و علت و معلولی است. در مفهوم کژمنشی، افراد ابتدا قراردادهای مختلفی را انتخاب می‌کنند. آنگاه یک فرد وقتی با پوشش بیمه‌ای بهتری مواجه می‌شود تشویق می‌شود که احتیاط کمتری را پیشه کار خود کند و تصادف بیشتری خواهد داشت. در هر دو مورد تنها یک الگو پدیدار می‌شود و آن اینست که با کنترل بر روی مشاهدات می‌بینیم که پوشش بیمه‌ای جامع‌تر باید با ریسک تحقق یافته بالاتر مرتبط باشد و این یک خاصیتی است که می‌تواند با استفاده از روش‌های مناسب پارامتری یا ناپارامتری آزمون شود.

روش همبستگی حاشیه‌ای چند مزیت دارد. روشی ساده و خیلی قوی است که دیانی و همکاران (۲۰۰۰) روی آن کار کرده‌اند. علاوه بر این، این روش می‌تواند روی داده‌های نمونه‌ای مورد استفاده در تحقیق که معرف جامعه‌اند مفید واقع شود به طوری که نسبتاً این داده‌ها آسان به دست می‌آید، هر چند که رسیدن به چنین کیفیتی تا حدودی هزینه‌بر است (ابرینگ و همکاران، ۲۰۰۳ ب). کژگزینی زمانی در یک بازار بیمه وجود دارد که خریداران بیمه اطلاعاتی در رابطه با ریسکشان دارند که بیمه‌گرانی که در برابر بیمه‌نامه‌ها متعهد می‌شوند، این اطلاعات را کم داشته و در نتیجه تعهد بیمه‌نامه‌های ناقص را به عهده می‌گیرند. بیمه‌گذار نیز در خرید بیمه‌نامه از این نقصان استفاده می‌کند. بیمه‌گذار ممکن است در باره احتمال یک زیان و یا توزیع اندازه یک زیان در هنگام وقوع حادثه و یا هر دو با اطلاع باشد. اگرچه کارهای اساسی در بیرون از بازار بیمه به روی کژگزینی انجام شده است، ما به چند دلیل تمرکز خود را بر روی مفهوم بیمه‌ای آن قرار خواهیم داد.

نخست، واژه ”کژگزینی” ابتدا در مفهوم بیمه‌ای مشاهده شده است و بازار بیمه، محلی برای برخی از نخستین اقتصاددانان بوده است که تحقیقات نظری روی آن انجام داده‌اند (آرو، ۱۹۶۳، پائولی، ۱۹۷۴ و روتشیلد و استیجلیتس، ۱۹۷۶). در مراسم اهدای جایزه نوبل به جوزف استیجلیتس، پروفیسور یورگن دابلو. ویبل مفهوم کژگزینی را معرفی کرد و چنین تفسیری داشت ”یک مثال عمده که در بیمه می‌توان یافت این است که شرکت‌های بیمه معمولاً قراردادهای جایگزینی را ارائه می‌کنند که حداکثر تخفیف آنها با سبک و سنگین کردن کمترین حق بیمه بدست می‌آید. بدین گونه مشتری‌های آنان، با انتخاب قرارداد، بصورت مؤثر در سطوح ریسک متمایزی تقسیم‌بندی می‌شوند.”

دوم اینکه بازارهای بیمه یک زمینه نسبتاً خوبی برای آزمون عملی نظریه کژگزینی فراهم می‌کنند و در نتیجه محققان به مطالعه کیفیت داده‌ها در چنین بازارهایی می‌پردازند. به عبارت دیگر در صورتی که کژگزینی رخ دهد، کیفیت یک طرف قرارداد و اطلاعات در دسترس درباره چنین کیفیتی معمولاً تدریجی و غیرقابل تحقیق و رسیدگی است (کوهن و سیجلمن، ۲۰۰۹). رابطه پیش‌بینی شده بین انتخاب نوع پوشش و تعداد تصادفات در نظریه اطلاعات نامتقارن را می‌توان به دو حالت بخش‌بندی کرد:

۱. نظریه‌های کژگزینی و کژمنشی، که یک رابطه مثبت بین انتخاب نوع پوشش و تعداد تصادفات را پیش‌بینی می‌کنند.

۲. نظریه‌هایی که رابطه صفر یا منفی را بین انتخاب نوع پوشش و تعداد تصادفات فرد پیش‌بینی می‌کنند. برای حالت اول می‌توان به می‌توان به مقالات روتشیلد و استیجلیتس (۱۹۷۶)، آرنوت و استیجلیتس (۱۹۸۸) و چیاپوری و همکاران (۲۰۰۶) اشاره کرد که تمام آنها با ارایه فرض‌های مختلف و در چارچوب‌های نظری متفاوت، وجود رابطه مثبت بین انتخاب نوع پوشش و تعداد تصادفات را تصریح می‌کنند. همچنین از کارهای انجام شده در این زمینه می‌توان به مقالات پونلز و اسنو (۱۹۹۴) و کوهن (۲۰۰۱) اشاره کرد. برای حالت دوم که همبستگی صفر یا منفی را بین ریسک و پوشش بیمه پیش‌بینی می‌کنند، می‌توان از مقاله‌های دی‌متزا و وب (۲۰۰۱)، ژولین و همکاران (۲۰۰۱) و کوفوپولوس (۲۰۰۴) نام برد و از کارهای تجربی انجام شده در این قالب نیز، می‌توان به سایتو (۲۰۰۶) اشاره کرد.

۳ خصوصیات مدل مفصل

در آزمون وجود اطلاعات نامتقارن در بازار بیمه خودرو از استدلالی مشابه آنچه چیاپوری و سالانی (۲۰۰۰) انجام دادند به‌وسیله بررسی مدل توزیع توام اندازه‌های چندجمله‌ای نوع پوشش و تعداد تصادفات پیروی می‌کنیم. چنین چیزی ما را به سمت بهره‌برداری از انعطاف‌پذیری ذاتی مدل‌های مفصل^۶ در این کار هدایت می‌کند. روش مفصلی که در اینجا استفاده شده برگرفته از کار زیمر و ترای‌ودی (۲۰۰۶) است، البته با مفهوم متفاوت که آن‌ها از یک مفصل سه متغیره برای مدل توام تقاضا برای بیمه خدمات درمانی خانواده و تصمیمات مدیریتی یا خرید بیمه خدمات درمانی بکار گرفته بودند. با وجود این که مفصل‌ها به‌طور وسیعی مدل‌های وابستگی را برای داده‌های پیوسته کاوش می‌کنند، کاربردهای آن‌ها برای داده‌های گسسته نیز مشاهده شده است.

چون توزیع‌های حاشیه‌ای گسسته هستند، بر اساس قضیه اسکالر مدل مفصل یکتا نیست. بر اساس داده‌های بازار بیمه خودرو فرانسه در کار چیاپوری و سالانی (۲۰۰۰) انتخاب نوع پوشش بیمه‌نامه معمولاً به‌وسیله یک اندازه دو بخشی توصیف شده است. همین‌طور بر اساس اظهارات کیم و همکاران (۲۰۰۹) یک اندازه چند بخشی از انتخاب نوع پوشش اطلاعات مفیدتری برای یک خریدار بیمه‌نامه فراهم می‌کند. بر همین اساس در این جا یک اندازه سه بخشی بررسی شده و مورد آزمون قرار گرفته شده است. بر اساس جامع بودن در ارائه خدمات از کمترین به بیشترین بصورت زیر طبقه‌بندی شده است:

۱. فقط بیمه شخص ثالث

۲. بیمه شخص ثالث، آتش‌سوزی و سرقت

۳. بیمه جامع

رده اول یا همان "فقط بیمه شخص ثالث" آن پوشش بیمه‌ای را که در پیشامد یک تصادف بیمه‌گذار مقصر شناخته می‌شود ارایه می‌کند و این کمترین تعهد پوشش مورد نیاز به‌وسیله قانون در تعداد زیادی از حوزه‌های قضایی

^۶ Copula

برای بیمه اتومبیل می‌باشد. رده دوم یا همان "بیمه شخص ثالث آتش سوزی و سرقت" علاوه بر پوشش فوق، پوشش بیمه‌ای در مقابل آتش سوزی و سرقت را نیز فراهم می‌کند. بالاخره، رده سوم یعنی "پوشش بیمه‌ای جامع" پوشش اضافه‌تری از پوشش‌های فوق را فراهم می‌کند به طوری که هر گونه آسیب در یک تصادف را مورد حمایت و پشتیبانی قرار می‌دهد.

لازم به ذکر است که پوشش بیمه‌ای جامع می‌تواند بر اساس سطوح کسریذیری‌های متفاوت در رده‌های مختلفی قرار گیرد. همچنین بر اساس کار چیاپوری و سالانی (۲۰۰۰) تقسیم کردن پوشش‌ها بر اساس کسریذیری مدل را بسیار پیچیده و مشکل خواهد کرد.

بر اساس مجموعه‌ای از مشاهده‌های چندبخشی با فرض‌های زیر شروع می‌کنیم: که y_{i1} و y_{i2} نشان‌دهنده‌ی انتخاب نوع پوشش و تعداد تصادفات برای بیمه‌نامه‌ی i ام می‌باشد. در اینجا، y_{i1} مقادیر ممکن ۱، ۲ یا ۳ که نشان‌دهنده‌ی به‌ترتیب انتخاب بیمه شخص ثالث، بیمه شخص ثالث آتش سوزی و دزدی، و پوشش جامع را به خود می‌گیرد. توجه داشته باشید که y_{i1} و y_{i2} متغیرهای مشاهده شده‌ای هستند که توسط متغیرهای پنهان که به‌ترتیب با y_{i1}^* و y_{i2}^* آنها را نشان می‌دهیم مقادیرشان ارزیابی می‌شود.

می‌توان y_{i1}^* را به عنوان پوشش بیمه‌ای که بیمه‌گذار آن را ترجیح می‌دهد و y_{i2}^* را به عنوان سطح ریسک ذاتی بیمه‌نامه در نظر گرفت. همان‌طور که در بخش ۱ شرح داده شد یک انتخاب پوشش بیمه‌نامه بهینه و سطح ریسک ذاتی تواما با هم تعیین می‌گردند. بنابراین بررسی این دو مورد با هم اطلاعات بهتری نسبت به بررسی آنها به صورت جداگانه در اختیار ما قرار خواهد داد.

یک رابطه خطی برای هر دو متغیر پنهان می‌تواند بر اساس مشخصه‌های زیر مفروض شود:

$$y_{i1}^* = x_i' \beta + \nu_i + \epsilon_{i1} \quad (1)$$

و

$$y_{i2}^* = z_i' \gamma + \nu_i + \epsilon_{i2} \quad (2)$$

در این جا x_i و z_i را به‌ترتیب بردارهای متغیرهای کمکی که انتخاب نوع پوشش را شرح می‌دهد و ریسک ضمنی بیمه‌نامه i ام تعریف می‌کنیم. بردارهای ضرایب β و γ همان پارامترهایی هستند که باید برآورد شوند. جمله‌ی عمومی ν_i کاراکترهای مشاهده نشده‌ی بیمه‌نامه i ام است که می‌تواند تواما بر روی انتخاب نوع پوشش و درجه‌ی ریسک تاثیرگذار باشد و بنابراین به‌عنوان پیوند بین این دو فرایند معرفی می‌شود. جمله‌های خطای ϵ_{i1} و ϵ_{i2} مستقل از ν_i فرض می‌شوند. مدل فوق می‌توانست توسط مشخص کردن توزیع‌های ν_i و خطا یا به‌وسیله‌ی مشخص کردن ساختار کوواریانس متغیرهای پنهان مدل‌بندی شود. به‌عنوان یک راه‌حل جایگزین، مدلی را برای متغیرهای قابل مشاهده y_{i1} و y_{i2} با یک مفصل پارامتری انتخاب و به‌صورت $C(\dots)$ تعریف می‌کنیم. آنگاه تابع جرم احتمال توام y_{i1} و y_{i2} را می‌توان به‌صورت زیر بیان کرد:

$$f_i(y_{i1}, y_{i2}) = P(Y_{i1} = y_{i1}, Y_{i2} = y_{i2})$$

$$\begin{aligned}
&= P(Y_{i1} \leq y_{i1}, Y_{i2} \leq y_{i2}) - P(Y_{i1} \leq y_{i1} - 1, Y_{i2} \leq y_{i2}) \\
&\quad - P(Y_{i1} \leq y_{i1}, Y_{i2} \leq y_{i2} - 1) - P(Y_{i1} \leq y_{i1} - 1, Y_{i2} \leq y_{i2} - 1) \\
&= F_{Y_{i1}, Y_{i2}}(y_{i1}, y_{i2}) - F_{Y_{i1}, Y_{i2}}(y_{i1} - 1, y_{i2}) - F_{Y_{i1}, Y_{i2}}(y_{i1}, y_{i2} - 1) + F_{Y_{i1}, Y_{i2}}(y_{i1} - 1, y_{i2} - 1) \\
&= C(F_{i1}(y_{i1}), F_{i2}(y_{i2})) - C(F_{i1}(y_{i1} - 1), F_{i2}(y_{i2})) \\
&\quad - C(F_{i1}(y_{i1}), F_{i2}(y_{i2} - 1)) + C(F_{i1}(y_{i1} - 1), F_{i2}(y_{i2} - 1)) \tag{۳}
\end{aligned}$$

که F_{i2} و F_{i1} توابع توزیع تجمعی y_{i2} و y_{i1} هستند. به دلیل جنبه پارامتری مدل مفصل برای شناسایی مدل نیاز به مشخص کردن توابع توزیع F_{i2} و F_{i1} داریم. انتخاب نوع پوشش با یک مقیاس ترتیبی اندازه گیری می‌شوند. بنابراین مدل چند جمله‌ای مرتب برای توصیف رابطه‌ی بین متغیر پاسخ y_{i1} و متغیر پنهان y_i^* استفاده می‌شود:

$$y_{i1} = \begin{cases} 1 & \text{if } y_i^* \leq \alpha_1 \\ 2 & \text{if } \alpha_1 < y_i^* \leq \alpha_2 \\ 3 & \text{if } y_i^* > \alpha_2 \end{cases} \tag{۴}$$

که α_1 و α_2 حدود نامعلومی هستند که به صورت اختیاری تخمین زده می‌شوند. بنابراین تابع توزیع y_{i1} را می‌توان به صورت زیر بیان نمود:

$$F_{i1}(y_{i1}) = P(Y_{i1} \leq y_{i1}) = \begin{cases} \pi(\alpha_1 - x_i' \beta) & y_{i1} = 1 \\ \pi(\alpha_2 - x_i' \beta) & y_{i1} = 2 \\ 1 & y_{i1} = 3 \end{cases} \tag{۵}$$

معمولاً برای یک پاسخ عددی روش‌های رگرسیون مفید واقع می‌شوند که شامل مدل پروبیت مرتب با $\pi(a) = \Phi(a)$ ، که $\Phi(\cdot)$ تابع توزیع تجمعی یک متغیر تصادفی نرمال استاندارد است و مدل لوجیت مرتب با $\pi(a) = \frac{1}{1+e^{-a}}$ می‌باشند. معمولاً این دو روش نتایج سازگاری را فراهم کرده و انتخاب بین این دو روش بر اساس ترجیحات استفاده‌کنندگان می‌باشد. برای اهدافمان، مدل رگرسیونی لوجیت مرتب را در برآوردمان مورد بررسی قرار می‌دهیم. بنابراین، مدل زیر را استفاده می‌کنیم:

$$F_{i1}(y_{i1}) = P(Y_{i1} \leq y_{i1}) = \begin{cases} \frac{1}{1+\exp(-(\alpha_1 - x_i' \beta))}, & y_{i1} = 1 \\ \frac{1}{1+\exp(-(\alpha_2 - x_i' \beta))}, & y_{i1} = 2 \\ 1, & y_{i1} = 3 \end{cases} \tag{۶}$$

تعداد تصادفات y_{i2} با استفاده از یک مدل رگرسیون دو جمله‌ای منفی مشخص شده است. به‌طور خاص تر تابع جرم احتمال آن به صورت زیر مشخص شده است:

$$\begin{aligned}
f_{i2}(y_{i2}) &= P(Y_{i2} = y_{i2}) = \left(\frac{\psi + y_{i2} - 1}{y_{i2}} \right) \left(\frac{\psi}{\lambda_i + \psi} \right)^\psi \left(1 - \frac{\psi}{\lambda_i + \psi} \right)^{y_{i2}} \\
&= \Gamma(y_{i2} + \psi) \Gamma(\psi) \Gamma(y_{i2} + 1) (\psi \lambda_i + \psi)^\psi (\lambda_i \lambda_i + \psi)^{y_{i2}} \tag{۷}
\end{aligned}$$

یعنی:

$$Y_{i2} \sim NB\left(\psi, \frac{\psi}{\lambda_i + \psi}\right)$$

که در آن

$$\begin{aligned} E(Y_{i2}) &= \lambda_i \\ \text{Var}(Y_{i2}) &= \lambda_i + \frac{1}{\psi} \lambda_i^2 \\ M_{Y_{i2}}(t) &= \left(\frac{\psi}{\psi + \lambda_i(1 - e^t)} \right)^\psi \end{aligned}$$

که یک تابع ربط لگ برای میانگین شرطی آن استفاده شده که به صورت زیر داده شده است:

$$E(y_{i2}|z_i) = \lambda_i = \omega_i \exp(\gamma z_i')$$

با ω_i که پارامتر وزن برای بیمه‌نامه i ام می‌باشد. این پارامتر وزنی بر مبنای طول زمانی که بیمه‌گذار به وسیله بیمه‌نامه برای سال مورد بررسی تحت پوشش قرار گرفته است، تفسیر می‌شود. پارامتر پراکنش ψ در واریانس شرطی که به صورت زیر ظاهر می‌شود:

$$\text{Var}(y_{i2}|z_i) = \lambda_i + \lambda_i^2/\psi$$

قابلیت انعطاف‌پذیری بیشتری برای برازش داده‌های ادعایمان فراهم می‌کند، به‌ویژه اغلب در مواردی که با داده‌های شمارشی ادعاها روبرو می‌شویم ما را قادر به اصلاح بیش‌پراکنشی یا کم‌پراکنشی می‌سازد. به‌عنوان مثالی در این زمینه به کامرون و ترای‌وی‌دی (۱۹۸۶) مراجعه کنید. مدل مشخص شده در این بخش به دلیل ماهیت ذاتی اش تمام پارامتری است (در واقع ناپارامتری نیست) و بنابراین به آسانی می‌تواند با استفاده از روشهای بر پایه - درستنمایی برآورد شود. لگاریتم تابع درستنمایی می‌تواند با استفاده از جمع بستن لگاریتم تابع احتمال توام بیان شده در (۱) برای تمام بیمه‌نامه‌های مجموعه داده قابل مشاهده داده شده به صورت

$$(y_{i1}, y_{i2}, x_i, z_i)$$

برای هر بیمه‌نامه i به دست آید. تابع درستنمایی $f(y_{i1}, y_{i2})$ به صورت زیر است:

$$\prod_{i=1}^n f(y_{i1}, y_{i2})$$

حال از تابع درستنمایی $f(y_{i1}, y_{i2})$ لگاریتم می‌گیریم:

$$\begin{aligned} \log \prod_{i=1}^n f(y_{i1}, y_{i2}) &= \sum_{i=1}^n \log f(y_{i1}, y_{i2}) \\ &= \sum_{i=1}^n [\log C(F_{i1}(y_{i1}), F_{i2}(y_{i2})) - \log C(F_{i1}(y_{i1} - 1), F_{i2}(y_{i2})) \\ &\quad - \log C(F_{i1}(y_{i1}), F_{i2}(y_{i2} - 1)) + \log C(F_{i1}(y_{i1} - 1), F_{i2}(y_{i2} - 1))] \end{aligned} \quad (۸)$$

در لگاریتم تابع درستنمایی، $F_{i1}(y_{i1})$ با عبارت (۴) جایگذاری می‌شود و $F_{i2}(y_{i2}) = \sum f_{i2}(y_{i2})$ با عبارت (۵) مشخص می‌گردد. برای تطبیق این واقعیت که انتخاب نوع پوشش و تعداد فراوانی تصادفات به‌طور مثبت یا منفی

با هم در ارتباط هستند، مفصل فرانک را که چنین انعطاف پذیری را به ما می دهد مورد استفاده قرار می دهیم:

$$C(u_1, u_2; \theta) = -\frac{1}{\theta} \log \left[1 + \frac{(e^{-\theta u_1} - 1)(e^{-\theta u_2} - 1)}{e^{-\theta} - 1} \right] \quad (9)$$

که θ همان پارامتر وابستگی بوده که بیانگر ارتباط بین دو پاسخ می باشد. انعطاف پذیری برای ارتباط مستقیم بیان شده در بالا از نخستین دلایل محبوبیت مفصل فرانک در کاربردهای بیمه، مالی و آمار پزشکی می باشد. به سادگی می توان نشان داد وقتی که $\theta \rightarrow 0$ مفصل فرانک به مفصل حاصل ضرب تبدیل می شود، که حالت خاص استقلال را نشان می دهد. با این وجود در حالتی که $\theta > 0$ یک ارتباط مثبت بین دو پاسخ را نشان می دهد؛ عکس این حالت زمانی رخ می دهد که $\theta < 0$. این ارتباطات به ما اجازه می دهد که نوع ارتباطی را که ما از داده هایمان (به طور تجربی) به دست می آوریم را بفهمیم. خاصیت های آماری بیشتری از خانواده مفصل فرانک در (۶) را می توانید در گنست (۱۹۸۷) پی گیری کنید. خواننده را برای درک بیشتر درباره این خانواده از مفصل ها به این مقاله خیلی جذاب ارجاع می دهیم.

۴ مقایسه بیمه های عمر و غیر عمر ایران با کشورهای جهان

در کشورهایی که آمار تصادفات به دلایل مختلفی چون کمبود جاده های استاندارد، نقص فنی ماشین ها، خطای بالای انسانی و عواملی از این قبیل بالاست، بخش قابل توجهی از دارشتمان شرکت های بیمه به رشته های عمومی که فرصت قلب و دروغگویی زیادی را برای بیمه گذار فراهم می کند، اختصاص می یابد. رشته های بیمه ای چون بیمه حوادث راننده، بیمه بدنه اتومبیل، بیمه شخص ثالث و بیمه نامه هایی از این قبیل در گروه بیمه های عمومی (غیرزندگی) قرار می گیرند. چنین اتفاقی بیشتر در کشورهای در حال توسعه که از لحاظ ساختار سیستم بیمه ای دچار مشکل هستند، می افتد. با بررسی دارشتمان محصولات بیمه ای شرکت های بیمه در سرتاسر جهان و تقاضا برای رشته های مختلف بیمه ای به نتایج جالب توجهی می رسیم (شکل ۴). براساس آمار بیمه مرکزی ایران در سال ۱۳۸۷ (پیوست ۱) سهم بیمه های عمومی مانند شخص ثالث و مازاد، حوادث راننده و بدنه اتومبیل از دارشتمان بیمه نامه های صادره شرکت های بیمه کشور حدود ۶۲.۶۷ درصد بوده است و این در حالی است که سهم همین بیمه نامه های عمومی از خسارت های واقع شده برای دارشتمان شرکت های بیمه کشور حدود ۷۰ درصد در این سال می باشد. همین طور در سال ۱۳۸۸ (پیوست ۲) سهم بیمه نامه های عمومی از دارشتمان بیمه نامه های صادره شرکت های بیمه کشور حدود ۶۰.۰۴ درصد و از خسارت های واقع شده حدود ۶۷ درصد در این سال بوده است. در سال ۱۳۸۹ (پیوست ۳) سهم بیمه های عمومی از دارشتمان بیمه نامه های صادره شرکت های بیمه کشور حدود ۵۲.۹۵ درصد بوده و این در حالی است که سهم همین بیمه نامه های عمومی از خسارت های واقع شده برای دارشتمان شرکت های بیمه کشور حدود ۶۰ درصد در این سال می باشد.

در صورتی که سهم بیمه های اجتماعی چون بیمه عمر که از رشته های سودآور برای شرکت های بیمه و یکی از عوامل



شکل ۱: سهم حق بیمه غیر زندگی از کل حق بیمه های تولیدی در ایران و جهان ۲۰۰۱ تا ۲۰۱۰

رفاه در هر کشور و نشانه پویایی شرکت های بیمه در فعالیت های اقتصادی است از داشتن شرکت های بیمه کشور در سال ۱۳۸۷ رقم ناچیز ۵.۲۲ درصد، در سال ۱۳۸۸ حدود ۶.۹ درصد و در سال ۱۳۸۹ حدود ۷.۹۲ درصد می باشد که این رقم در دنیا به طور متوسط حدود ۵۹ درصد و در کشورهای توسعه یافته مانند انگلیس حدود ۸۰ درصد می باشد.

با این تفاسیر در بازارهایی مانند بازار بیمه ایران و اسرائیل که پیش بینی وجود اطلاعات نامتقارن امری محتمل است، علاوه بر مفصل فرانک می توان از مفصل هایی مانند کلایتون، گامبل، جو و گالامبوس که پارامتر وابستگی آنها فقط می تواند توجیه کننده رابطه مثبت بین انتخاب نوع پوشش و تعداد تصادفات بیمه گذار باشد، نیز استفاده کرد. آزمون های نیکویی برازش هم برای توزیع های حاشیه ای و هم برای مدل های مفصل قابل اعمال است. همان طوری که در بخش قبل ذکر شد مدل مفصل فرانک ارتباط های مثبت و منفی را برای ما تبیین می کند. چون آزمون آماری پارامتر وابسته از مسائلی که در فرض روی مرز با آن روبرو می شویم دوری می کند، چنین مزیتی آزمون نیکویی برازش را به ما پیشنهاد می کند. علاوه بر این می توان یک آزمون نسبت درستی را برای آزمون کیفیت برازش مدل مفصل فرانک اعمال کرد. آزمون به وسیله مدل مفصل، یک آزمون نیرومند و قوی است. اطلاعات پیشین درباره رابطه بین نوع پوشش و تعداد رخدادها هم یک رابطه مثبت و هم یک رابطه منفی را بین انتخاب نوع پوشش و تعداد خسارت ها پیشنهاد می کند. به خاطر چنین اطلاعات پیشینی، برای برازش مدل ناخود آگاه از مفصل فرانک که اجازه چنین انعطاف پذیری را به ما می دهد، استفاده می کنیم. به دلیل بهینه یابی غیرخطی مورد نیاز در ارزیابی تابع درستی، علاقمند به برآورد نتایج آزمون نیرومند به وسیله آزمون کردن سایر مشخصه های مفصل هستیم. به این منظور مدل مفصل فرانک را تحت سه تا دیگر از مفصل های از نوع ارشمیدسی متداول، یعنی مفصل گامبل، مفصل کلایتون و مفصل جو با ویژگی های زیر دوباره مدل بندی می کنیم:

$$C(u_1, u_2, \theta) = \exp[-((- \log u_1)^\theta + (- \log u_2)^\theta)^{\frac{1}{\theta}}], \theta \geq 1 \quad (10)$$

$$C(u_1, u_2, \theta) = (u_1^{-\theta} + u_2^{-\theta} - 1)^{\frac{1}{\theta}}, \quad \theta > 0 \quad (11)$$

$$C(u_1, u_2) = 1 - [(1 - u_1)^\theta + (1 - u_2)^\theta - (1 - u_1)^\theta(1 - u_2)^\theta]^{\frac{1}{\theta}}, \quad \theta \geq 1 \quad (12)$$

که فرمول‌های (۱۰)، (۱۱) و (۱۲) به ترتیب تابع مفصل‌های گامبل، کلایتون و جو هستند که θ نمایانگر پارامتر وابسته در هر سه نوع مفصل می‌باشد. توجه داشته باشید که مفصل‌های گامبل، کلایتون و جو نمی‌توانند رابطه منفی را تطبیق دهند. در مواردی که اطلاعات نامتقارن به شدت در بازار وجود داشته باشد این سه نوع مفصل برای آزمون کردن آزمون‌های نیرومند ممکن واجد شرایط می‌شوند. اگر در مفصل گامبل و مفصل جو $\theta = 1$ و در مفصل کلایتون $\theta = 0$ قرار دهیم به آسانی استقلال را نتیجه خواهیم گرفت.

از انواع مفصل‌های غیر ارشمیدسی که می‌توانیم در کارمان برای مشاهده ارتباط بین انتخاب نوع پوشش و تعداد تصادفات استفاده کنیم، مفصل نرمال بوده که از انواع مفصل‌های بیضوی است که ساختار آن به صورت زیر می‌باشد:

$$C(u_1, u_2) = \int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} \frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}} \exp\left\{-\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)}\right\} dy dx \quad (13)$$

باید توجه داشت که اگر همه ضرایب ماتریس همبستگی P صفر باشد آنگاه شرط استقلال را خواهیم داشت. مفصل غیر ارشمیدسی دیگری که می‌تواند این ارتباط را نشان دهد مفصل گالامبوس بوده که ساختار تک پارامتری زیر را داراست:

$$C(u_1, u_2, \theta) = u_1 u_2 \exp\{[-(\ln u_1)^{-\theta} + (-\ln u_2)^{-\theta}]^{-\frac{1}{\theta}}\}, \quad \theta > 0 \quad (14)$$

۵ نتیجه‌گیری

این مقاله به طور مقدماتی سعی بر ارائه روشهای مفید نظری برای آزمون وجود اطلاعات نامتقارن در بازارهای مختلف و به ویژه بازار بیمه خودرو داشته است. ما یک روش مفصل رگرسیونی دومتغیره را برای آزمون کردن توام انتخاب نوع پوشش و سطح ریسک بیمه‌گذار ارائه کردیم. برای آزمون کردن وجود اطلاعات نامتقارن، انتخاب پوشش بیمه‌گذار توسط یک متغیر طبقه‌بندی شمارشی اندازه‌گیری شده و درجه ریسک فرد نیز تقریباً به وسیله اندازه ریسک قبلی فرد یعنی تعداد دفعاتی که فرد در طول یک سال ادعای خسارت کرده مورد ارزیابی قرار گرفته است. روش‌های موجود مشکلات یکسان زیادی برای مدل کردن دو متغیر گسسته دارند و مورد انتقاد زیادی به خاطر مدل‌های به کار برده شده برای آزمون عدم تقارن اطلاعات قرار گرفته‌اند. مدل‌های مفصل ارائه شده ما به خاطر انعطاف‌پذیری بالایش نقصان مدل‌های قبلی را پوشش داده و در عین حال هم رابطه خطی و هم غیرخطی بین انتخاب نوع پوشش و سطح ریسک بیمه‌گذار را نشان داده است که چنین چیزی در مقاله‌های تقارن اطلاعات از دست می‌رفت. برای آزمون کردن مدل مفصل می‌توان از یک مشاهده تجربی مقطعی یک‌ساله از یک شرکت بیمه بزرگ استفاده کرد. آنچه که در این مقاله مورد توجه قرار گرفته، نوع‌آوری ساختار مفصل برای آزمون اطلاعات نامتقارن بوده و همین‌طور تحلیل

بازار بیمه از لحاظ ساختار اطلاعاتی و تاثیرات احتمالی این عدم تقارن اطلاعات بر روی سود بیمه‌گر و بیمه‌گذار بوده است.

مراجع

- [1] Abbring, J.H., Chiappori, P.A. and Pinquet, J. (2003), Moral hazard and dynamic insurance data, *European Economic Association*, **1**, 767–820.
- [2] Abbring, J.H., Heckman, J.J., Chiappori, P.A. and Pinquet, J. (2003), Adverse selection and moral hazard in insurance: can dynamic data help to distinguish?, *European Economic Association*, **1**, 512–521.
- [3] Akerlof, G.A. (1970), The market for ‘lemons’: quality uncertainty and the market mechanism, *Quarterly Journal of Economics*, **84**, 488–500.
- [4] Chiappori, P.A., and Salanié, B. (2000), Testing for asymmetric information in insurance market, *Political Economy*, **108**, 56–78.
- [5] Chiappori, P.A., Jullien, B., Salanié, B. and Salanié, F. (2006), Asymmetric information in insurance: general testable implications, *RAND Journal of Economics*, **37**, 783–798.
- [6] Cohen, A. (2005), Asymmetric information and learning: evidence from the automobile insurance market, *Review of Economics and Statistics*, **87**, 197–207.
- [7] Cohen, A. and Siegelman, P. (2010), Testing for adverse selection in insurance markets, *Risk and Insurance*, **77**, 39–84.
- [8] De Meza, D. and Webb, D.C. (2001), Advantageous selection in insurance markets, *RAND Journal of Economics*, **32**, 249–262.
- [9] Dionne, G., Michaud, P.C. and Dahchour, D. (2010), Separating moral hazard from adverse selection and learning in automobile insurance: Longitudinal evidence from France, Working Paper.
- [10] Frees, E.W., Shi, P. and Valdez, E.A. (2009), Actuarial applications of a hierarchical claims model, *ASTIN Bulletin*, **39**, 165–197.

جدول ۱-الف / عملکرد بازار بیمه در سال ۱۳۸۷																	
ضریب خسارت			خسارت						حق بیمه						رشته		
تغییر (بند)	۱۳۸۷ (%)	۱۳۸۶ (%)	واقع شده	موقوف اول سال	موقوف پایان سال	پرداختی			عاید شده	ذخیره اول سال	ذخیره پایان سال	تولیدی					
						سهم از (بازار) (%)	درخ رشد (%)	۱۳۸۷				۱۳۸۶	درخ رشد (%)	۱۳۸۷		۱۳۸۶	
۱۸/۵۵	۴۱/۲۷	۲۲/۷۲	۱۰۲۰/۳	۲۹۶/۶	۳۷۴/۶	۳/۸۱	۱۱۰/۱۲	۹۴۲/۳	۴۴۸/۴	۲۴۴/۳	۱۰۵۴/۴	۱۱۷۴/۷	۶/۳۹	۴۰/۰۱	۲۵۹۲/۶	۲۶۴۵/۸	آتش سوزی
۳/۴۷	۲۱/۶۳	۱۷/۱۶	۲۲۸/۵	۳۵۳/۱	۳۷۹/۰	-/۸۲	۷۳/۷۲	۴۰۲/۷	۱۱۶/۰	۱۰۵۶/۳	۳۳۴/۶	۴۲۱/۲	۲/۵۷	-۴/۳۱	۱۰۴۴/۹	۱۰۸۹/۹	باربری
۲/۳۸	۳۹/۴۰	۳۷/۰۱	۲۷۲/۲	۵۲/۹	۸۵/۸	-/۹۷	۱۵/۹۶	۲۳۹/۲	۲۰۶/۳	۶۹۰/۹	۲۴۸/۰	۳۹۲/۷	۲/۰۶	۳۵/۹۷	۸۳۵/۶	۶۱۴/۵	حوادث
-۹/۹۱	۲۳/۵۵	۳۳/۴۵	۳۷۳/۰	۱۷۲/۷	۱۴۶/۲	۱/۶۱	۱۴/۱۹	۳۹۹/۵	۳۴۹/۹	۱۵۸۴/۲	۵۰۸/۹	۹۷۶/۹	۵/۰۶	۶۱/۳۳	۲۰۵۲/۳	۱۷۷۲/۱	حوادث سرنشین
-۱/۶۹	۶۲/۲۰	۶۳/۸۹	۲۶۰۱/۲	۲۱۳/۲	۱۴۵/۷	۱۰/۸۶	۶/۰۶	۲۶۸۸/۶	۲۵۳۵/۱	۴۱۸۱/۹	۱۶۶۴/۵	۲۴۶۰/۶	۱۴/۲۷	۱۹/۶۳	۳۹۷۷/۹	۴۱۶۱/۱	بدنه اتومبیل
-۹/۵۰	۹۳/۳۶	۱۰۳/۸۶	۱۴۸۳۰/۵	۳۱۳۲/۶	۲۹۸۲/۷	۶۰/۵۲	۱۰/۷۳	۱۳۹۸۰/۵	۱۳۵۲۷/۷	۱۵۷۱۶/۹	۶۰۱۱/۹	۸۶۸۳/۳	۳۵/۳۳	۲۲/۳۵	۱۸۳۸۹/۳	۱۵۰۳۰/۱	شخص ثالث
۲۱/۲۲	۹۵/۱۷	۷۳/۹۶	۳۰۱۳/۱	۱۰۷/۸	۲۱۹/۵	۱۱/۷۲	۵۱/۴۳	۲۹۰۱/۳	۱۹۱۶/۱	۳۱۶۵/۸	۱۳۳۳/۰	۱۶۱۴/۴	۸/۹۹	۲۸/۷۱	۳۶۴۷/۱	۲۸۳۳/۵	درمان
-۴۲/۴۵	۳۳/۱۶	۸۵/۶۱	۶۱/۴	۷۱/۸	۴۶/۴	-/۳۵	۳۱/۶۱	۸۶/۸	۶۶/۰	۱۴۲/۲	۵۷/۶	۵۵/۵	-/۳۵	-۴/۵۷	۱۴۰/۱	۱۴۶/۸	کشتی
۱۹/۱۴	۵۰/۰۵	۳۰/۹۱	۲۲۳/۱	۱۸۳/۲	۲۴۴/۶	-/۶۶	۲۳/۶۴	۱۶۲/۶	۱۳۱/۵	۴۴۵/۹	۱۲۸/۰	۱۲۸/۰	۱/۱۰	۳۹/۳۳	۴۴۵/۸	۳۲۰/۰	هواپیما
۹/۰۶	۳۲/۱۶	۲۳/۰۰	۲۴۷/۳	۲۳۸/۵	۳۲۱/۱	-/۶۷	۲۶/۹۵	۱۶۴/۸	۱۲۹/۸	۷۶/۹۱	۳۷۴/۹	۳۶۷/۳	۱/۸۸	-۱۸/۷۷	۷۶۱/۵	۹۳۷/۵	مهندسی
۶/۸۰	۱۳/۱۹	۶/۳۸	۴/۳	۱۶/۱	۱۵/۰	-/۰۲	۶۳/۱۲	۵/۴	۳/۳	۳/۶	۱۵/۶	۱۶/۴	-/۰۸	-۱۴/۲۶	۳۳/۵	۳۹/۰	پول
۱۴/۸۶	۵۸/۰۲	۴۳/۱۶	۱۴۳۹/۷	۵۸۸/۴	۸۷۴/۳	۳/۸۶	۵۷/۹۰	۹۵۵/۹	۶۰۵/۴	۲۱۳۶/۶	۷۳۳/۵	۹۷۰/۳	۵/۸۵	۲۹/۰۱	۲۳۷۳/۵	۱۸۳۹/۸	مسئولیت
۱۷/۱۹	۱۱۵/۸۹	۹۸/۷۰	۸۸/۹	۲۷۸/۱	۳۴۵/۲	-/۰۹	-۲۵/۵۷	۲۱/۷	۲۹/۲	۷۶/۷	۳۶/۳	۲۹/۷	-/۱۷	-۲۳/۵۶	۷۰/۱	۹۱/۷	اعتبار
-۱۴/۶۰	۱۱/۵۹	۲۶/۱۹	۱۰۱/۶	۸۲۳/۲	۸۸۰/۲	-/۱۸	۹۷/۶۲	۴۴/۷	۲۲/۶	۸۷۶/۶	۳۲۸/۹	۵۴۸/۶	۲/۶۵	۱۹/۰۴	۱۰۷۶/۳	۹۰۴/۱	نفت و انرژی
-۱۴/۲۸	۶۷/۲۸	۷۹/۵۵	۳/۳	۴/۱	۴/۴	-/۰۱	-۱۰/۳۶	۳/۰	۳/۳	۴/۹	۴/۳	۴/۲	-/۰۲	۱۸/۷۵	۶/۹	۵/۸	سایر انواع
-/۴۰	۷۴/۸۸	۷۳/۴۹	۲۴۳۰۸/۵	۶۵۳۳/۳	۷۰۴۴/۸	۹۶/۱۵	۱۸/۴۶	۲۳۷۹۹/۰	۴۰۰۹۰/۶	۳۳۳۵۳/۰	۱۴۷۳۴/۳	۱۷۸۲۴/۶	۹۴/۷۸	۴۰/۴۰	۴۸۴۴۵/۳	۳۱۹۳۱/۹	جمع غیر زندگی
۱۱۴/۳۶	۱۹۹/۸۳	۸۷/۴۶	۹۴۶/۶	۸۹/۷	۸۴/۶	۳/۸۵	۳۰/۱۲	۹۵۳/۶	۷۳۴/۹	۴۷۳/۷	۳۹۶۵/۶	۵۶۰/۷۶	۵/۲۲	۱۱/۸۱	۲۱۱۵/۸	۱۸۹۲/۲	زندگی (عمر)
-/۹۹	۷۴/۶۶	۷۳/۶۷	۲۵۲۵۵/۱	۶۶۲۳/۰	۷۱۲۵/۵	۱۰۰/۰۰	۱۸/۸۷	۲۴۷۵۲/۶	۲۰۸۲۳/۵	۳۳۸۴۶/۷	۱۶۶۹۷/۹	۲۳۳۳۲/۲	۱۰۰/۰۰	۱۹/۹۲	۴۰۵۶۱/۰	۳۳۸۲۴/۱	جمع کل

شکل ۲: پیوست ۱

- [11] Klugman, S.A., Panjer, H.H. and Willmot, G.E. (2008), Loss Models From Data To Decision, Willy series in probability and statistics, Willy.
- [12] Nelsen, R.B. (1998), An introduction to copula, Portland, Oregon.
- [13] Shi, P. and Valdez, E.A. (2011), A copula approach to test asymmetric information with applications to predictive modeling, *Insurance: Mathematics and Economics*.49, 226--239.

جدول ۱-ح/ عملکرد بازار بیمه در سال ۱۳۸۸															
رشته	حق بیمه						خسارت				ضریب خسارت		تغییر (واحد)		
	تولیدی		ذخیره پایان سال	ذخیره اول سال	عاید شده	پرداختی	موقوف پایان سال	موقوف اول سال	واقع شده	۱۳۸۸ (%)	۱۳۸۷ (%)	مجموع		۱۳۸۸ (%)	
	نرخ رشد (%)	سهم از کل (%)													نرخ رشد (%)
آتش سوزی	۲۵۹۱/۶	۲۹۱۵/۴	۱۲/۴۵	۶/۲۸	۱۳۰۰/۲	۱۱۷۳/۷	۲۷۹۰/۰	۹۴۲/۳	۷۶۰/۲	۱۹/۳۲	۲/۳۷	۳۷۴/۵	۳۰۲/۱	۳۴/۶۶	-۱۶/۶۱
باربری	۱۰۴۲/۶	۷۹۴/۸	-۳/۸۰	۳۴۲/۶	۳۲۱/۲	۸۷۳/۳	۴۰۲/۷	۲۱۳/۳	۵/۲۵	-۰/۶۹	۳۴۳/۱۰	۳۷۸/۶	۴۱/۶۲	۴۰/۳۳	-۱/۲۹
حوادث	۸۳۶/۳	۸۵۸/۰	۲/۵۹	۱/۸۵	۳۹۵/۱	۸۳۷/۸	۲۴۰/۲	۲۸۳/۰	۱۷/۸۳	-۲/۹۲	۶۷/۰۸	۸۵/۴	۲۶۴/۷	۳۹/۵۸	-۷/۹۹
حوادث سرقتین	۲۰۵۲/۳	۲۳۹۰/۵	۱۶/۴۸	۵/۱۵	۱۱۲۰/۸	۹۷۴/۵	۲۲۴۴/۲	۳۹۸/۶	۶/۵۰	۳/۳۸	۱۵۸/۱	۱۴۶/۲	۲۳۶/۴	۲۳/۴۵	-۳/۰۰
بدنه انومبیل	۴۹۷۷/۹	۵۴۵۱/۹	۹/۵۲	۹/۵۲	۴۷۱۵/۹	۲۴۶۱/۰	۵۱۹۷/۰	۲۶۸۸/۶	۱۴/۱۵	۹/۹۸	۲۲۴/۱	۱۴۵/۷	۳۱۶۷/۴	۶۰/۹۵	-۱/۲۶
شخص ثالث	۱۸۳۸۸/۶	۲۰۰۵۱/۹	۹/۰۵	۴۳/۱۶	۹۳۵۷/۳	۸۶۸۴/۹	۱۹۳۷۹/۵	۱۴۹۸۰/۵	۱۶/۱۲	۱۶/۳۸	۳۲۱۳/۱	۲۹۸۲/۷	۱۷۶۲۶/۳	۹۰/۹۵	-۳/۲۱
درمان	۳۶۴۷/۱	۵۳۹۳/۳	۴۷/۸۸	۱۱/۶۱	۲۷۷۵/۷	۱۶۱۴/۲	۴۲۳۱/۸	۲۹۰۱/۳	۷/۰۴	۳۹۶۲/۳	۳۹۶/۵	۲۱۹/۵	۵۱۳۹/۲	۱۴۱/۴۴	۲۶/۲۷
کشتی	۱۴۰/۱	۱۳۵/۱	-۳/۵۶	-۰/۶۹	۵۵/۴	۶۰/۱	۱۳۰/۵	۸۶/۸	-۳۴/۵۹	-۰/۱۸	۱۷۱/۱	۴۴/۰	۱۸۳/۹	۱۴۰/۹۵	۹۹/۴۶
هولیتما	۴۴۵/۸	۵۳۱/۶	۲۱/۴۹	۱/۱۷	۱۶۳/۶	۱۲۸/۰	۵۰/۶۰	۱۶۲/۶	۱۰/۶۰	۳۳۶/۰	۲۳۸/۳	۲۲۴/۶	۳۲۹/۷	۶۵/۱۱	۱۵/۱۱
مهندسی	۷۶۱/۵	۱۰۷۹/۳	۴۱/۷۳	۲/۳۲	۲۴۶/۷	۳۶۷/۳	۹۹۶/۹	۱۶۲/۸	۸۰/۶۳	۲۹۷/۶۳	۲۴۵/۰۱	۳۲۱/۰۷	۴۲/۱۶	۴۲/۱۶	۱۰/۰۰
پول	۳۳/۵	۵۸/۹	۷۵/۹۴	-۲/۱۳	۲۵/۳	۱۶/۴	۵۰/۱	۵/۴	۱۱۷/۶۳	-۲/۰۴	۱۸/۷	۱۵/۰	۱۵/۵	۳۰/۹۱	۱۷/۷۱
مسئولیت	۲۳۷۳/۵	۲۶۱۹/۴	۱۰/۳۶	۵/۶۴	۱۱۶۶/۴	۹۷۰/۴	۲۴۲۳/۵	۹۵۵/۹	۱۹/۰۱	۱۱۳/۶۶	۳/۷۰	۹۴۹/۹	۸۶۹/۲	۵۰/۲۷	-۷/۶۴
اعتبار	۷۰/۱	۳۳/۰	-۵۲/۸۶	۰/۰۷	۲۹/۴	۲۹/۸	۴۶/۷	۲۱/۷	۱۴۷/۳۶	-۰/۱۷	۳۰۱/۰	۳۳۵/۲	۹/۵	۲۰/۲۶	-۹۵/۷۲
نفت و انرژی	۱۰۷۶/۳	۹۱۴/۵	-۱۵/۴۱	۱/۹۶	۴۳۴/۷	۵۲۸/۶	۱۰۰۶/۴	۴۴/۷	۷۸۶/۷۰	۳۹۶/۱	۱/۲۹	۶۸۲/۹	۸۸۰/۲	۱۹/۷۶	۸/۱۷
سایر انواع	۶/۹	۱۷/۹	۱۶۱/۲۱	۰/۰۴	۹/۹	۴/۳	۱۲/۳	۳/۰	۱۱۸/۲۰	۶/۵	۱/۷	۲/۴	۳/۸	۳۰/۳۱	-۳۷/۳۱
جمع غیر زندگی	۳۸۴۴/۳	۴۳۵۳/۵	۱۲/۵۱	۹۳/۱۰	۲۰۳۵۰/۵	۱۷۸۲۵/۹	۴۰۳۵۰/۵	۳۳۷۹۰/۰	۲۳/۵۵	۲۶۳/۰۴	۷۵۱۲/۶	۷۰۳۶/۷	۲۹۸۸۰/۳	۷۳/۳۶	۰/۵۰
زندگی (عمر)	۲۱۱۵/۷	۳۲۰۶/۲	۵۱/۵۴	۶/۹۰	۷۷۳۴/۹	۵۶۰۷/۶	۱۰۷۹/۰	۹۵۳/۶	۴۱/۳۳	۱۳۴/۷۸	۷۷/۹	۸۲/۶۴	۱۳۴۳/۱	۱۲۴/۴۸	-۷۵/۳۶
جمع کل	۴۰۵۱۶/۰	۴۶۳۵۹/۷	۱۴/۵۳	۱۰۰/۰۰	۲۸۰۸۵/۴	۲۳۴۳۴/۵	۴۱۸۰۷/۸	۲۴۷۵۲/۶	۳۰۷۵۲/۳	۲۴/۲۴	۷۵۹۰/۶	۷۱۱۹/۴	۳۱۲۳۳/۵	۷۴/۶۸	۰/۰۴

شکل ۳: پیوست ۲

جدول ۱-۴ عملکرد بازار بیمه در سال ۱۳۸۹

رشته	حق بیمه												خسارت					ضرریب خسارت	
	تولیدی					ذخیره پایان سال	ذخیره اول سال	عاید شده	پرداختی				معوق پایان سال	معوق اول سال	واقع شده	۱۳۸۸ (%)	۱۳۸۹ (%)	تغییر (واحد)	
	۱۳۸۹	۱۳۸۸	نرخ رشد (%)	سهم از رشته (%)	۱۳۸۹				۱۳۸۸	نرخ رشد (%)	سهم از رشته (%)	۱۳۸۹							۱۳۸۸
						۱۳۸۹	۱۳۸۸	نرخ رشد (%)					سهم از رشته (%)						
آتش سوزی	۲۹۱۵/۴	۳۳۵۲/۳	۱۴/۹۸	۵/۶۷	۱۵۶۶/۸	۱۳۰۱/۰	۳۰۸۶/۵	۷۶۰/۲	۸۳۶/۲	۹/۷۳	۲/۱۳	۳۲۱/۸	۳۰۲/۰	۸۵۴/۰	۲۴/۶۶	۲۷/۶۷	۳/۰۱		
بادبری	۷۹۴/۸	۸۷۲/۴	۹/۷۷	۱/۴۷	۳۹۹/۳	۳۴۴/۴	۸۱۶/۵	۴۱۳/۴	۱۵۸/۳	-۲۵/۸۱	-۲/۴	۴۱۱/۴	۳۴۲/۴	۲۲۷/۳	۲۰/۴۸	۲۷/۸۴	۷/۵۶		
حوادث	۸۶۲/۸	۹۳۵/۶	۸/۴۵	۱/۵۸	۴۵۰/۲	۴۱۵/۲	۹۰۰/۶	۴۸۳/۰	۲۷۹/۳	-۱/۳۱	-۰/۷۱	۸۷/۶	۶۷/۱	۳۰۰/۲	۳۱/۴۱	۳۳/۳۳	۱/۹۲		
حوادث راننده	۲۳۸۵/۶	۱۵۸۸/۰	-۳۳/۴۴	۲/۶۸	۶۵۰/۰	۱۱۲۰/۸	۲۰۵۸/۸	۴۲۴/۵	۴۷۸/۶	۱۲/۸۲	۱/۲۲	۲۳۶/۹	۱۵۸/۱	۵۵۷/۷	۱۹/۴۹	۲۷/۰۹	۷/۶۰		
بدنه اتومبیل	۵۴۵۱/۹	۵۴۵۹/۵	-۰/۱۴	۹/۲۳	۴۵۱۴/۵	۲۷۱۵/۷	۵۶۶۰/۷	۳۰۶۹/۳	۳۵۱۹/۴	۹/۰۲	۱۵/۳۲	۲۸۰/۶	۲۲۴/۱	۳۵۹۶/۰	۶۰/۹۵	۶۳/۵۳	۲/۵۸		
شخص ثالث ومالک	۲۰۰۵۱/۹	۲۴۲۷۹/۰	۲۱/۰۸	۳۱/۰۴	۱۰۹۳۳/۷	۹۳۵۶/۶	۲۲۷۰/۱۸	۱۷۳۹۶/۶	۱۳۳۱/۶	۵۱/۱۲	۱۵/۲۵	۳۰۱۶/۷	۳۲۱۳/۱	۲۰۸۵۳/۰	۹۰/۹۵	۹۱/۸۶	-۰/۹۰		
درمان	۵۳۹۳/۳	۱۱۱۷۱/۷	۱۰۷/۱۴	۱۸/۸۸	۵۹۷۲/۳	۲۷۷۶/۱	۷۷۵/۵	۴۹۶۲/۳	۹۷۸/۲	۹۷/۱۳	۲۴/۹۴	۳۹۶/۶	۵۳۵/۳	۹۹۲۱/۰	۱۲/۴۶	۲۴۶/۳۹	۲/۹۴		
کشتی	۱۳۵/۱	۳۹۶/۷	۱۹۳/۵۹	-۰/۶۷	۱۹۳/۵۹	۳۹۶/۷	۵۰/۱	۲۲۷/۲	۵۶/۸	۷۸/۴۰	۱۰/۱۳	۳۰۹/۴	۱۷۱/۱	۲۶۹/۴	۱۳/۳۶	۱۱۷/۳۶	-۲۳/۶۱		
هواپیما	۵۴۱/۶	۵۴۳/۹	-۰/۴۲	-۰/۹۲	۱۷۰/۷	۱۶۳/۶	۵۳۶/۸	۳۳۶/۰	۱۴۸/۹	-۳۸	-۳۸	۳۰۹/۳	۲۳۸/۳	۴۱۹/۸	۶۵/۱۵	۴۰/۹۵	-۲۴/۴۰		
مهندسی	۱۰۵۷/۷	۱۴۵۷/۵	۱۸/۸۹	۲/۱۳	۵۶۱/۸	۴۴۶/۷	۱۱۴۲/۴	۲۹۷/۹	۳۴۰/۰	۱۴/۱۴	-۰/۸۷	۷۵۳/۷	۴۴۵/۰	۶۲۸/۷	۴۳/۱۲	۵۶/۷۸	۱۳/۶۶		
پول	۵۸/۹	۶۴/۴	۹/۱۳	-۰/۱۱	۲۹/۰	۲۵/۳	۶۰/۵	۱۱/۸	۱۳/۶	۱۵/۵۰	-۰/۳	۱۹/۰	۱۸/۷	۱۳/۹	۳۰/۰	۲۳/۰۱	-۷/۸۹		
مستوفیت	۲۶۱۹/۴	۳۴۱۱/۱	۲۲/۵۹	۵/۴۳	۱۳۸۲/۱	۱۱۶۶/۷	۲۹۹۵/۸	۱۱۳۷/۶	۱۳۶۶/۷	۲۰/۱۴	۳/۴۸	۱۱۰۶/۰	۹۴۹/۹	۱۵۲۲/۸	۵۰/۸۳	۵۰/۸۳	-۰/۵۶		
اعتیاد	۳۲/۳	۳۸۱/۱	۱۰۷۹/۳۳	-۰/۶۴	۲۴۳/۲	۱۶/۶	۱۵۴/۴	۵۳/۷	۳/۵	-۳/۴۸	-۰/۸	۴۰۳/۹	۳۲۲/۶	۱۱۳/۸	۷۳/۶۹	۷۳/۶۹	۵/۴۶		
نفت و انرژی	۹۱۲/۶	۹۳۱/۸	۲/۱۰	۱/۵۷	۴۷۳/۱	۳۵۵/۱	۹۰۲/۷	۳۹۶/۸	۳۸۹/۲	-۱/۶۸	-۰/۹۹	۵۳۵/۶	۶۸۲/۹	۲۵۱/۹	۱۹/۹۴	۲۷/۹۰	۷/۹۶		
سایر انواع	۱۸/۷	۳۰/۶	۶۳/۸۷	-۰/۰۵	۱۳/۱	۹/۸	۲۷/۳	۱۱/۰	۶/۵	۶۸/۵۴	-۰/۳	۲/۶	۱/۷	۱۱/۸	۴۸/۹۴	۴۳/۳۹	۱۴/۳۹		
جمع غیر زندگی	۴۲۸۶۰/۰	۵۳۴۷۵/۳	۲۷/۰۹	۹۲/۰۸	۲۵۵۸۸/۰	۲۰۳۶۴/۶	۴۲۴۶/۹	۲۶۰۵/۶	۳۷۵۴/۶	۲۷/۶۱	۹۵/۶۷	۹۳۹/۸	۷۵۳۳/۵	۳۶۶۱/۲	۷۳/۶۴	۷۶/۹۲	۶/۴۳		
زندگی (عمر) *	۳۴۰۶/۲	۴۶۸۶/۰	۳۶/۱۷	۷/۹۲	۹۳۰۶/۹	۷۷۳۴/۹	۱۸۱۱/۵	۱۳۴۷/۸	۱۶۹۸/۲	۲۶/۰۰	۱۶۹/۰	۱۳۹/۵	۷۷/۹	۱۱۵۲/۹	۶۷/۴۸	۶۳/۶۴	-۳/۸۳		
جمع کل	۴۶۴۳۸/۲	۵۹۱۶۶/۰	۲۷/۴۰	۱۰۰/۰۰	۳۳۸۹۰/۹	۲۸۰۹۷/۵	۵۲۳۶۸/۶	۳۰۷۵۳/۴	۳۹۲۳۳/۱	۲۷/۵۴	۱۰۰/۰۰	۹۵۰۹/۳	۷۶۱۱/۴	۴۱۱۲۱/۱	۷۴/۸۰	۷۸/۵۲	۳/۸۳		

شکل ۴: پیوست ۳