

تحلیل داده‌های بورس بر اساس تابع مفصل

سیده آزاده فلاح مرتضی نژاد^۱، غلامرضا محتشمی برزادران^۲، بهرام صادق‌پور گیلده^۳، و محمد امینی^۴

تاریخ دریافت: ۱۳۹۹/۱۰/۰۸

تاریخ پذیرش: ۱۴۰۰/۰۸/۲۵

چکیده:

تابع مفصل یک ابزار مفید در شناسایی ساختار وابستگی داده‌های وابسته و در نتیجه برازش یک توزیع توأم مناسب به داده‌ها است. در این مقاله با استفاده از تابع مفصل داده‌های بورس شامل سه متغیر ضعف مالی، سود انباشته، و دارایی ملموس مربوط به ۱۱۰ شرکت تجاری ایران از سال ۱۳۸۵ تا ۱۳۸۹ تحلیل می‌شود و به‌خصوص یک توزیع سه‌بعدی به این داده‌ها برازش خواهد شد. برای بررسی نوع وابستگی بین داده‌ها از ابزارهای مختلفی از جمله نمودار پراکنش، نمودار کای، و نمودار کندال استفاده می‌کنیم. همچنین اندازه وابستگی جهت‌دار و دمی داده‌ها را مورد بررسی قرار می‌دهیم و ضرایب وابستگی کندال و اسپیرمن را نیز محاسبه خواهیم کرد. در انتها آزمون نیکویی برازش برای چند مفصل معروف را انجام می‌دهیم تا بتوان مفصل مناسب داده‌های بورس این مقاله را دست آوریم.

واژه‌های کلیدی: تابع مفصل، وابستگی جهتی و دمی، نمودار کای و کندال.

۱ مقدمه

قراردادند. آن‌ها تأثیر ساختار وابستگی بر تغییرات مرز کارا، را با به‌کارگیری توابع مفصل تحلیل کردند. شاکری و همکاران (۱۳۹۰) با استفاده از تابع مفصل مشاهدات سانسور وابسته از بیماران سرطانی بررسی کردند. در مدل رایج یکی از فرضیات اساسی استقلال میان زمان سانسور و زمان پیشامد می‌باشد، ولی همواره نگرانی در خصوص اعتبار مبتنی بر فرضیه استقلال سانسورها وجود داشت. آنان این مشکل را با توابع مفصل حل کردند. گنجعلیخانی و همکاران (۱۳۹۵) روشی برای درون‌یابی پهنه‌بندی کیفی آب‌های زیرزمینی دشت‌های کرمان با توابع مفصل ارائه کردند. مرتضی‌نژاد و همکاران (۱۳۹۸) بر ساخت پیوندی بین آنتروپی و توابع مفصل کار نمودند.

در این مقاله تحلیل داده‌های بورس ۱۱۰ شرکت تجاری با استفاده از توابع مفصل انجام می‌شود. متغیرهای مورد بررسی ضعف مالی، سود انباشته و دارایی ملموس می‌باشد. هدف اولیه یافتن تابع مفصل این داده‌ها است. برای این منظور ابتدا با استفاده از چند نمودار گرافیکی با توجه به فیشر و اسویتزر (۲۰۰۱) و جنست و بويس (۲۰۰۳) یک نمای کلی از ساختار وابستگی داده‌ها نمایش می‌دهیم. در این مرحله مشخص می‌شود که یک نوع وابستگی مثبت میان داده‌ها وجود دارد. اما تنها از روی مشاهده این نمودارها نمی‌توان تصمیم‌گیری کرد. بنابراین نیازمند یک سری آزمون‌های اولیه نسبت

یکی از روش‌های مهم برای تشخیص توزیع توأم متغیرهای وابسته با استفاده از توزیع‌های حاشیه‌ای، توابع مفصل هستند. مدلی که بر اساس توابع مفصل برای متغیرهای وابسته تولید می‌شود، دارای تمام ویژگی‌های وابستگی میان متغیرها است. تابع مفصل برای اولین بار توسط اسکالار (۱۹۵۹) بیان شد. پس از آن محققانی مانند راشندورف (۱۹۹۶) و بنز و استین (۱۹۹۷) مطالعات اساسی در مورد توابع مفصل و کاربردهای آن انجام دادند. فیشر و اسویتزر (۲۰۰۱) و جنست و بويس (۲۰۰۳) از نمودارهای کای و کندال برای نمایش تصویری ساختار وابستگی داده‌ها استفاده کردند. نلسن (۲۰۰۶) خانواده‌های جدیدی از تابع مفصل ساخت. گاندورف (۲۰۱۰) چند ویژگی توابع مفصل را بررسی کرد. عباسیان و همکاران (۱۳۹۳) برای پیش‌بینی سیلاب‌ها با استفاده از توابع مفصل، مطالعاتی را انجام دادند. تحقیقات آن‌ها بر اساس تحلیل سه متغیره بوده است. محمودیان و همکاران (۱۳۹۳) مدل فضایی برای تحلیل مقادیر کرانگین با توزیع حاشیه‌ای تعمیم‌یافته را معرفی کردند و در آن وابستگی‌های کوچک مقیاس با استفاده از تابع مفصل تی به دست آوردند. بدافلو و همکاران (۲۰۱۵) تحلیل بیزی عوامل مؤثر بر دفع پیوند قرنيه با استفاده از تابع مفصل کلايتون کارکردند. صالحی و زمانی (۱۳۹۵) داده‌های بورس اوراق بهادار تهران را بر اساس مفصل مورد تحلیل

^۱ گروه آمار - دانشگاه فردوسی azadehfallah@mail.um.ac.ir

^۲ گروه آمار - دانشگاه فردوسی (نویسنده مسئول): grmohtashami@um.ac.ir

^۳ گروه آمار - دانشگاه فردوسی sadehgpour@um.ac.ir

^۴ گروه آمار - دانشگاه فردوسی m-amini@um.ac.ir

$$\widehat{V}_i = \frac{1}{n} \sum_{j=1}^n I(X_{j\tau} \leq X_{i\tau}),$$

$$\widehat{W}_i = \frac{1}{n} \sum_{j=1}^n I(X_{j\tau} \leq X_{i\tau}).$$

۲ مروری بر وابستگی بر اساس مفصل

در این بخش اندازه وابستگی جهتی بردارهای تصادفی و اندازه وابستگی دمی چند متغیره را برای داده‌های وابسته بیان می‌کنیم. این اندازه‌های وابستگی کمک به تشخیص ساختار وابستگی داده‌ها می‌کند. برای مطالعه بیشتر بر این اندازه‌های وابستگی به کوشادا و یوبدا (۲۰۱۲) و لوکا و رویجو (۲۰۱۲) رجوع شود.

۱.۲ وابستگی جهتی بردارهای تصادفی

فرض کنید (Ω, F, P) یک فضای احتمال باشند که Ω یک مجموعه‌ی غیر تهی و F یک زیر σ میدان از زیرمجموعه‌های Ω و P یک اندازه احتمال بر F باشند. همچنین $X = (X_1, X_2, \dots, X_n)$ یک بردار تصادفی از Ω به $\mathbb{R}^n = [-\infty, \infty]^n$ با تابع توزیع توأم H است.

تعریف ۱.۲. بردار تصادفی پیوسته $X = (X_1, X_2, \dots, X_n)$ با تابع توزیع توأم H و $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$ به طوری که برای تمام $i = 1, \dots, n$ $|\alpha_i| = 1$ در نظر بگیرید، آنگاه

- بردار تصادفی X را وابسته مثبت در جهت α یا $(PD(\alpha))$ گوئیم اگر برای هر $x \in \mathbb{R}^n$ نامساوی زیر برقرار باشد:

$$P(\cap_{i=1}^n (\alpha_i X_i > x_i)) \geq \prod_{i=1}^n P(\alpha_i X_i > x_i).$$

- X را وابسته منفی در جهت α یا $(ND(\alpha))$ گوئیم اگر عبارت زیر برای هر $x \in \mathbb{R}^n$ درست باشد:

$$P(\cap_{i=1}^n (\alpha_i X_i > x_i)) \leq \prod_{i=1}^n P(\alpha_i X_i > x_i)$$

به‌عنوان یک نتیجه از تعریف (۱.۲)، بردار تصادفی X را وابسته مثبت در جهت α گوئیم، اگر و تنها اگر برای هر $x \in \mathbb{R}^n$ ، $I = \{1, \dots, n\}$ و $J = \{i \in I | \alpha_i > 0\}$ نامساوی زیر برقرار باشد:

$$P[\cap_{i \in J} (X_i > x_i), \cap_{i \in I \setminus J} (X_i < x_i)] \geq \prod_{i \in J} P[X_i > x_i] \cdot \prod_{i \in I \setminus J} P[X_i < x_i]$$

با شرایط مشابه، برای وابستگی منفی با تغییر جهت نامساوی‌ها در روابط فوق، نتیجه حاصل می‌شود.

به داده‌ها هستیم، تا از عدم استقلال و وجود وابستگی میان داده‌ها اطمینان حاصل کنیم. سپس برای تشخیص تابع مفصل مناسب از آزمون نیکویی برازش استفاده کردیم و مقادیر آکائیکه و تابع درست نمایی مفصل‌های مختلف را مقایسه کردیم و در نتیجه بهترین تابع مفصل قابل برازش به داده‌ها تعیین می‌شود. علاوه بر این اندازه وابستگی جهتی و دمی از نکات مهم دیگر مقاله است. برای شروع یک سری تعریف و مفاهیم اولیه از مفصل‌ها را به‌اختصار بیان می‌کنیم، نلسن (۲۰۰۶).

تعریف ۱.۱. یک مفصل n - متغیره (یا به‌طور خلاصه n - مفصل) یک تابع C از I^n به I با خاصیت‌های زیر است.

۱. اگر حداقل برای یک j از $(j = 1, \dots, n)$ ، $u_j = 0$ ، آنگاه $C(\vec{u}) = 0$.

۲. اگر برای $i \neq j$ ($j = 1, \dots, n$) و $u_j = 1$ آنگاه $C(\vec{u}) = u_i$.

۳. برای هر $\vec{a}, \vec{b} \in I^n$ به طوری که $\vec{a} \leq \vec{b}$ ؛ $V_C([\vec{a}, \vec{b}]) \geq 0$ ، که در آن

$$V_C([\vec{a}, \vec{b}]) = \Delta_{\vec{a}}^{\vec{b}} C(\vec{u}) = \Delta_{a_n}^{b_n} \Delta_{a_{n-1}}^{b_{n-1}} \dots \Delta_{a_1}^{b_1} C(\vec{u}).$$

قضیه ۲.۱. (قضیه اسکالر در حالت n - متغیره) فرض کنید H یک تابع توزیع n - متغیره با حاشیه‌های F_n, \dots, F_1 باشد. در این صورت یک n - مفصل C در I^n وجود دارد به طوری که:

$$H(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n)); \quad (1)$$

$$\forall x_i \in R_i, (i = 1, \dots, n).$$

اگر F_n, \dots, F_1 همگی پیوسته باشند، آنگاه C یکتاست. در غیر این صورت C به‌طور یکتا بر $\text{Ran} F_1 \times \dots \times \text{Ran} F_n$ تعریف می‌شود. برعکس اگر C یک n - مفصل و F_n, \dots, F_1 توابع توزیع باشند، آنگاه تابع H به‌صورت (۱) یک تابع توزیع n - بعدی با حاشیه‌های F_n, \dots, F_1 است.

در این مقاله برای محاسبه ضرایب وابستگی جهتی که در ادامه به آن‌ها اشاره خواهد شد، به برآوردگر مفصل تجربی نیازمندیم. از این رو با توجه به چن و هانگ (۲۰۰۷) برآوردگر مفصل تجربی به‌صورت زیر تعریف می‌شود:

$$\widehat{C}(u, v, w) = \frac{1}{n} \sum_{i=1}^n I(\widehat{U}_i \leq u, \widehat{V}_i \leq v, \widehat{W}_i \leq w)$$

که در آن $\widehat{U}_i, \widehat{V}_i$ و \widehat{W}_i از روابط زیر تعیین می‌شوند:

$$\widehat{U}_i = \frac{1}{n} \sum_{j=1}^n I(X_{j1} \leq X_{i1}),$$

۲.۲ اندازه وابستگی دمی چند متغیره

وابستگی دمی یک اندازه هم‌راستا بین متغیرها است. این اندازه هم‌راستا تمایل به تجمع در دم‌های پایینی و بالایی توزیع توأم دارد. مقدار این اندازه بین ۰ و ۱ می‌باشد. هر چه اندازه وابستگی دمی به ۱ نزدیک‌تر باشد، وابستگی میان داده‌ها، هم‌راستای مثبت خواهد بود. در لوکا و رویچو (۲۰۱۲) حالت چند متغیره اندازه وابستگی دمی را بیان کرده است. در این مقاله فقط سه متغیر موجود است، بنابراین فقط بر حالت سه متغیره تأکید می‌کنیم. حال برای حالت سه متغیره، فرض کنید $n = 3$ متغیر داریم و $h = 1$ متغیر مستقل در احتمال شرطی است، ضرایب وابستگی دمی بالا به صورت زیر به دست می‌آیند:

$$\lambda_U^{1|23} = \lim_{u \rightarrow 1^-} \frac{\widehat{C}(1-u, 1-u, 1-u)}{\widehat{C}_{12}(1-u, 1-u)}$$

$$= \frac{1-3u + C_{12}(u, u) + C_{13}(u, u) + C_{23}(u, u) - C(u, u, u)}{1-2u + C_{12}(u, u)}$$

$$\lambda_U^{2|13} = \lim_{u \rightarrow 1^-} \frac{\widehat{C}(1-u, 1-u, 1-u)}{\widehat{C}_{13}(1-u, 1-u)}$$

$$= \frac{1-3u + C_{12}(u, u) + C_{13}(u, u) + C_{23}(u, u) - C(u, u, u)}{1-2u + C_{13}(u, u)}$$

$$\lambda_U^{3|12} = \lim_{u \rightarrow 1^-} \frac{\widehat{C}(1-u, 1-u, 1-u)}{\widehat{C}_{23}(1-u, 1-u)}$$

$$= \frac{1-3u + C_{12}(u, u) + C_{13}(u, u) + C_{23}(u, u) - C(u, u, u)}{1-2u + C_{23}(u, u)}$$

و به‌طور مشابه ضرایب وابستگی دمی پایین برای همان مقادیر n و h :

$$\lambda_L^{1|23} = \lim_{u \rightarrow 0^+} \frac{C(u, u, u)}{C_{12}(u, u)}, \quad \lambda_L^{2|13} = \lim_{u \rightarrow 0^+} \frac{C(u, u, u)}{C_{13}(u, u)},$$

$$\lambda_L^{3|12} = \lim_{u \rightarrow 0^+} \frac{C(u, u, u)}{C_{23}(u, u)},$$

و برای $n = 3$ و $h = 2$ ، ضرایب وابستگی دمی بالا و پایین مانند زیر تعیین می‌شوند:

$$\lambda_U^{1|23} = \frac{1-3u + C_{12}(u, u) + C_{13}(u, u) + C_{23}(u, u) - C(u, u, u)}{1-u},$$

$$\lambda_L^{1|23} = \frac{C(u, u, u)}{u}.$$

واضح است که در حالت فوق سه ضریب وابستگی دمی بالا $\lambda_U^{1|23}$ ، $\lambda_U^{2|13}$ و $\lambda_U^{3|12}$ با هم برابر بوده، پس کافی است یکی را محاسبه کنیم. این مطلب برای ضرایب وابستگی دمی پایین نیز برقرار است.

قضیه ۲.۲. مفصل سه‌بعدی C وابسته مثبت در جهت هر α است، اگر و تنها اگر $C = \Pi^3$.

بر اساس مطالب بالا می‌توان نوع وابستگی جهتی میان داده‌ها را شناخت و طبق آن اندازه وابستگی جدید مانند کندال و اسپیرمن را معرفی کرد. فرض کنید (X, Y, Z) یک بردار تصادفی باشد، در این صورت اندازه‌های وابستگی جهتی کندال و اسپیرمن برای داده‌های سه‌بعدی به صورت زیر تعریف می‌شوند:

$$\tau_{\alpha}^+(C) = \frac{1}{3}(\lambda P(\alpha X > \alpha Y) - 1)$$

$$\rho_{\alpha}^+(C) = \lambda \int_{[0,1]^3} Q_{\alpha}(u, v, w) du dv dw$$

که در آن:

$$Q_{\alpha}(u, v, w) = P[\alpha_1 X > \alpha_1 u, \alpha_2 Y > \alpha_2 v, \alpha_3 Z > \alpha_3 w] - P[\alpha_1 X > \alpha_1 u]P[\alpha_2 Y > \alpha_2 v]P[\alpha_3 Z > \alpha_3 w].$$

برای توضیح اولیه قضیه (۳.۲) با توجه به نلسن و یوبدا (۲۰۱۲)، فرض کنید (X, Y, Z) یک بردار تصادفی یکنواخت بر I^3 باشد که تابع توزیع آن مفصل سه‌بعدی C است، آنگاه $\rho_{\alpha}^+(C)$ و $\rho_{\alpha}^-(C)$ برای هر $u, v, w \in I^3 = [0, 1]^3$ با روابط زیر تعیین می‌شود:

$$\rho_{\alpha}^+(C) = \lambda \int_{I^3} [P(X > u, Y > v, Z > w) - P(X > u)P(Y > v)P(Z > w)] du dv dz$$

$$\rho_{\alpha}^-(C) = \lambda \int_{I^3} [P(X \leq u, Y \leq v, Z \leq w) - P(X \leq u)P(Y \leq v)P(Z \leq w)] du dv dz$$

قضیه ۳.۲. فرض کنید بردار تصادفی (X, Y, Z) دارای مفصل سه‌بعدی C باشد، در این صورت برای هر جهت $(\alpha_1, \alpha_2, \alpha_3)$:

$$\rho_{\alpha}^{(\alpha_1, \alpha_2, \alpha_3)}(C) = \frac{\alpha_1 \alpha_2 \rho_{XY} + \alpha_2 \alpha_3 \rho_{YZ} + \alpha_3 \alpha_1 \rho_{ZX}}{3} + \alpha_1 \alpha_2 \alpha_3 \frac{\rho_{\alpha}^+ - \rho_{\alpha}^-}{2}$$

و

$$\tau_{\alpha}^{(\alpha_1, \alpha_2, \alpha_3)} = \frac{\alpha_1 \alpha_2 \tau_{XY} + \alpha_2 \alpha_3 \tau_{YZ} + \alpha_3 \alpha_1 \tau_{ZX}}{3}$$

برای به دست آوردن اندازه‌های وابستگی جهتی کندال و اسپیرمن ابتدا باید درستی قضیه ۲ در مقاله کوئسادا و یوبدا (۲۰۱۲) را برای داده‌ها در جمات هشت‌گانه بررسی کنیم. سپس برای جهاتی که داده‌های مقاله در آن صدق کرده‌اند، ضرایب وابستگی جهتی کندال و اسپیرمن متناظر را محاسبه کنیم. به‌عنوان مثال اگر داده‌ها در جهت $(-1, 1, 1)$ باشند، این بدان معناست که مقادیر کوچک متغیر اول هم‌زمان تمایل به مقادیر بزرگ متغیرهای دوم و سوم دارد. سایر جهات نیز به‌طور مشابه این جهت تفسیر می‌شوند.

۳ نمودارهایی برای تشخیص ساختار وابستگی

$$F_i = \frac{1}{n-1} \#\{j \neq i : X_j \leq X_i\}$$

$$G_i = \frac{1}{n-1} \#\{j \neq i : Y_j \leq Y_i\}$$

بنا بر فرضیه استقلال $H_i = F_i \times G_i$ ، فیشر (۱۹۸۵) و اسویتزر (۲۰۰۱) زوج (λ_i, χ_i) را رسم کردند و آن را نمودار کای نامیدند:

$$\chi_i = \frac{H_i - F_i G_i}{\sqrt{F_i(1-F_i)G_i(1-G_i)}} \quad (۳)$$

$$\lambda_i = \text{sing}(\tilde{F} \tilde{G}) \max(\tilde{F}_i^\uparrow, \tilde{G}_i^\uparrow)$$

$$\tilde{F}_i = F_i - \frac{1}{4}, \quad \tilde{G}_i = G_i - \frac{1}{4} \quad 1 \leq i \leq n$$

برای $\lambda_i \in [-1, 1]$ فاصله زوج (χ_i, λ_i) از مرکز داده‌ها است. برای اجتناب از داده‌های نادرست پیشنهاد می‌شود، فقط نقاطی که در نامساوی زیر صدق می‌کنند، را رسم کنیم.

$$|\lambda_i| < 4 \left\{ \frac{1}{n-1} - \frac{1}{4} \right\}^2$$

قابل توجه است که برای محاسبه فرمول (۳) از زوج (X_{ij}, Y_{ij}) برای تمام $i \neq j$ به شکل زیر استفاده می‌کنیم:

$$X_{ij} = \begin{cases} 1 & X_j \leq X_i \\ 0 & \text{جاها سایر} \end{cases}$$

$$Y_{ij} = \begin{cases} 1 & Y_j \leq Y_i \\ 0 & \text{جاها سایر} \end{cases}$$

برای هر $1 \leq i \leq n$ ، $-1 \leq \chi_i \leq 1$ است. به علاوه $\sqrt{n}\chi_i$ آماره کای اسکور است که معمولاً برای آزمون استقلال 2×2 به وسیله نقاط برش (X_i, Y_i) مورد استفاده قرار می‌گیرد. برای کمک به تشخیص تغییر تصادفی در مقادیر مشاهدات χ خطوط کنترل افقی را به نمودار اضافه می‌کنیم. فیشر و اسویتزر (۱۹۸۵) و اسویتزر (۲۰۰۱) این پیشنهاد را دادند. این حدود کنترل به صورت $\chi = \pm c_p / \sqrt{n}$ هستند که در آن c_p به وسیله شبیه‌سازی مونت کارلو به دست می‌آید، که به طور تقریبی $100 \times p$ درصد از زوج‌های (λ_i, χ_i) بین خطوط قرار می‌گیرند، فیشر و اسویتزر (۲۰۰۱) و جنست و بويس (۲۰۰۳). برای درک بهتر از نمودار کای به شکل ۱ توجه کنید. در آن با رسم نقاط شبیه‌سازی شده از دو مفصل مختلف FGM و فرانک، رفتار این نمودار را در ساختار وابستگی مثبت و منفی به نمایش کشیدیم. مفصل FGM به صورت (۴) تعریف می‌شود. داده‌های شکل به ازای $\alpha = 0.99$ شبیه‌سازی شده‌اند. همان‌طور که در شکل ۱ سمت چپ مشاهده می‌شود، به دلیل وجود ساختار وابستگی مثبت نقاط بالای خط افقی $\chi = 0$ قرار دارند. هر چه ساختار وابستگی مثبت بین نقاط مشاهده شده قوی‌تر باشد، تجمع نقاط از این خط افقی دورتر و به سمت بالا خواهد بود.

$$C_\alpha(u, v) = uv[1 + \alpha(1-u)(1-v)] \quad , \alpha \in \mathcal{R} \quad (۴)$$

یک نمونه تصادفی $(X_1, Y_1), \dots, (X_n, Y_n)$ به حجم $n \geq 2$ از یک تابع توزیع تجمعی دومتغیره پیوسته H با حاشیه‌های F و G در اختیار داریم. ساده‌ترین نمودار برای تشخیص مستقل بودن متغیرهای X و Y همان نمودار پراکنش مشاهدات است. اما برای رسم سایر نمودارها نیازمند توزیع‌های حاشیه‌ای داریم که معمولاً نامعلوم هستند. یک روش ساده برای کنترل این تأثیر رسم جفت‌های $(\hat{F}_n(X_i), \hat{G}_n(Y_i))$ برای $1 \leq i \leq n$ است، که در آن \hat{F}_n و \hat{G}_n به ترتیب توابع توزیع تجربی X_i و Y_i به فرم زیر هستند:

$$\hat{F}_n(t) = \frac{1}{n} \#\{i : X_i \leq t\} = \frac{1}{n} \sum_{i=1}^n I\{X_i \leq t\}$$

$$\hat{G}_n(t) = \frac{1}{n} \#\{i : Y_i \leq t\}$$

واضح است که جفت $(\bar{F}_n(t), \bar{G}_n(t))$ همان $(R_i/n, S_i/n)$ هستند که در آن $R_i = r(X_i)$ و $S_i = r(Y_i)$ برای $i = 1, \dots, n$ است. رسم زوج رتبه‌ها از رسم زوج مشاهدات مناسب‌تر است، زیرا مفصل‌ها تحت تبدیلات صعودی پایا هستند، بنابراین برای رسم نمودار پراکنش می‌توان از زوج‌های $(R_i/n, S_i/n)$ برای $i = 1, \dots, n$ استفاده کرد. از این نمودار برای تشخیص درستی فرضیه استقلال استفاده می‌شود، اما تشخیص تصادفی بودن با مشاهده این نمودار دشوار است. برای تشخیص دقیق وابستگی، نمودارهای دیگری مانند کای و کندال وجود دارند، که نسبت به این نمودار مناسب‌تر هستند.

۱۰۳ نمودار کای

یکی از نمودارهایی که برای مشاهده ساختار وابستگی داده‌ها مورد استفاده قرار می‌گیرد، نمودار کای معرفی شده در فیشر و اسویتزر (۲۰۰۱) و جنست و بويس (۲۰۰۳) است. با به‌کارگیری این نمودار می‌توان شدت وابستگی مثبت یا منفی را تشخیص داد. ابتدا با استفاده از داده‌ها، زوج اعداد (χ_i, λ_i) را محاسبه کرده و آن‌ها را بر نمودار رسم می‌کنیم. هر چه این اعداد به منحنی بالایی نزدیک‌تر باشند، داده‌ها دارای ساختار وابستگی مثبت و هر چه از آن منحنی دور شده و در پایین نمودار قرار گیرند، ساختار وابستگی داده‌ها منفی خواهد بود. برای محاسبه زوج (χ_i, λ_i) به روش زیر عمل می‌کنیم. برای هر زوج (χ_i, λ_i) توزیع‌های تجربی توأم و حاشیه‌ای با استفاده از عبارت زیر محاسبه می‌شوند:

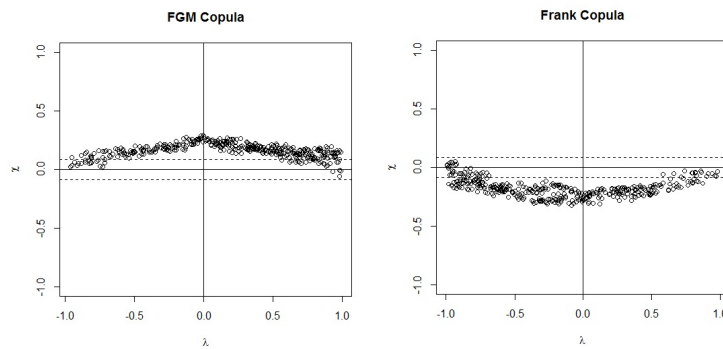
$$H_i = \frac{1}{n-1} \#\{j \neq i : X_j \leq X_i, Y_j \leq Y_i\} \quad (۲)$$

توجه به نمودار مذکور، هنگامی که وابستگی میان مشاهدات وجود نداشته باشد، نقاط در پیرامون همان خط افقی تجمع خواهند کرد و هیچ تمایلی به سمت بالا یا پایین نخواهند داشت.

(۵)

$$C_{\theta}(u, v) = -\frac{1}{\theta} \log \left[1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{\theta} - 1} \right], \theta \in \mathcal{R} \setminus \{0\}$$

مفصل فرانک نیز در (۵) نشان داده شده است. مشاهدات این مفصل به ازای $\theta < 0$ ، شبیه سازی شده اند. با توجه به شکل ۱ سمت راست، به دلیل منفی بودن ساختار وابستگی، نقاط در پایین خط $\chi = 0$ اجتماع کرده اند. این موضوع نشان دهنده وجود ساختار وابستگی منفی بین داده ها است. شکل ۲ برای داده های شبیه سازی شده از مفصل استقلال $C(u, v) = uv$ رسم شده است. با



شکل ۱: به ترتیب از چپ به راست نمودار کای مربوط به مفصل FGM و فرانک

برای مشاهدات دومتغیره کاربرد ندارد. الگوریتم زیر زمانی استفاده می شود، که مشاهدات دومتغیره به صورت $(X_1, Y_1), \dots, (X_n, Y_n)$ باشند. با مشاهده این نمودار می توان تشخیصی در مورد ساختار وابستگی نمونه تصادفی مذکور داد. الگوریتم یاد شده از قرار زیر است:

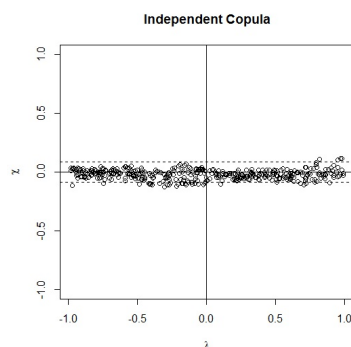
۱. ابتدا H_i ها را در عبارت (۲) محاسبه می کنیم.
۲. آماره های ترتیبی $H_{(1)} \leq \dots \leq H_{(n)}$ به دست می آوریم.
۳. زوج $(W_{n:i}, H_{(i)})$ برای $1 \leq i \leq n$ رسم می کنیم. در آن $W_{i:n}$ بیانگر امید i امین آماره ترتیبی در یک نمونه تصادفی به حجم n از یک توزیع K از H_i تحت فرضیه صفر که همان استقلال می باشد، است. به نمودار به دست آمده از این

۲.۳ نمودار کندال یا K

بر اساس جنس و بویس (۲۰۰۳) هنگامی که با یک نمونه تصادفی تک متغیره X_1, \dots, X_n مواجه هستیم، یک روش رایج برای شناسایی مشخصه گاوسی یعنی مقایسه آن با توزیع نرمال استاندارد، به روش گرافیکی، رسم نمودار چندک برای نقاط $(Z_{i:n}, X_{(i)})$ است که در آن $X_{(1)} \leq \dots \leq X_{(n)}$ آماره های ترتیبی نمونه هستند و $Z_{i:n}$ رتبه نرمال i امین مشاهده از یک نمونه نرمال استاندارد به حجم n است.

$$Z_{i:n} = E(Z_{(i)}); \quad 1 \leq n.$$

در آن $Z_{(1)} \leq \dots \leq Z_{(n)}$ آماره های ترتیبی از یک نمونه تصادفی Z_1, \dots, Z_n از یک توزیع نرمال استاندارد می باشد. اما این راه کار



شکل ۲: نمودار کای مفصل استقلال

نقاط را نمودار کندال یا K می‌گویند.

برای توضیح کامل فرآیند کار تنها کافی است شکل K را تحت فرضیه صفر مشخص کنیم. تابع چگالی یک آماره ترتیبی برای $1 \leq i \leq n$ به صورت زیر تعریف می‌شود:

$$W_{i:n} = n \binom{n-1}{i-1} \int_0^1 w \{K_0(w)\}^{i-1} \times \{1-K_0(w)\}^{n-i} dK_0(w)$$

هنگامی که داده‌ها را به بازه $[0, 1]$ منتقل کنیم، به آن‌ها شبه مشاهدات می‌گوییم. بر اساس جنست و ریوست (۱۹۹۳) تابع توزیع تجربی K_n با استفاده از شبه مشاهدات، H_1, \dots, H_n به صورت مجانبی توزیع کندال خواهد بود. \hat{H}_n تابع توزیع تجربی بر اساس (X_i, Y_i) برای $i \neq j$ تعریف می‌شود. آنگاه زمانی که n به بی‌نهایت میل کند، H_i به H همگراست. اثبات این نتیجه کمی دشوار و دقیق است زیرا H_i ها نسبت به هم، به طور تصادفی مستقل‌اند، اما می‌توان به وسیله توزیع کندال به راحتی اثبات کرد و برای محاسبه $W_{i:n}$ از این K استفاده می‌کنیم.

در اینجا برخی خواص نمودار K را بیان خواهیم کرد. وقتی n به ∞ میل کند، $K_n(w)$ در احتمال همگرا به $K(w)$ برای تمام $0 \leq w \leq 1$ است و با توجه به داشتن فرضیه کران‌داری، $K_n^{-1}(p)$ همگرا در احتمال به $K^{-1}(p)$ برای تمام $0 \leq p \leq 1$ خواهد بود.

• اگر برای یک عدد صحیح دلخواه $n \geq 1$ و $0 \leq p \leq 1$ کوچک‌ترین عدد صحیح بزرگ‌تر یا مساوی np را با $[np]$ نشان دهیم، آنگاه تحت فرضیه صفر داریم:

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

و همچنین

$$\lim_{n \rightarrow \infty} E(H_{([np])}) = \lim_{n \rightarrow \infty} W_{[np]:n} = K_0^{-1}(p)$$

• برای یک نمونه به اندازه کافی بزرگ n زوج‌های $(W_{i:n}, H_{(i)})$ به منحنی زیر متمایل می‌شوند.

$$p \rightarrow (K_0^{-1}(p), K^{-1}(p))$$

به عبارت دیگر نقاط نمودار K شبیه به نمودار $w \rightarrow K^{-1}\{K_0(w)\}$ است.

• هنگامی که تحت فرضیه استقلال، $K = K_0$ را داشته باشیم، نمودار به صورت خطی خواهد بود.

• وقتی X و Y هم‌یکنوا با $\tau(X, Y) = -1$ باشند، تمام نقاط روی خط افقی ($p \equiv 0$) قرار خواهد گرفت، زیرا برای تمام مقادیر ممکن $0 \leq p \leq 1$ ، $K^{-1}(p) = 0$ را داریم.

• وقتی که X و Y هم‌یکنوا با $\tau(X, Y) = 1$ باشند، تمام نقاط بر منحنی $K_0(p)$ قرار می‌گیرد، زیرا بر بازه $[0, 1]$ ، $K^{-1}(p) \equiv p$ است.

برخلاف نمودار کای که نمی‌توان آن را برای حالت چند متغیره تعمیم داد، نمودار K به حالت چند متغیره تعمیم داده شده است. با استفاده از یک نمونه تصادفی p متغیره $(X_{n1}, \dots, X_{np}), \dots, (X_{11}, \dots, X_{1p})$ فرآیند رسم نمودار K چند متغیره به صورت زیر است:

۱. برای هر $1 \leq i \leq n$ ، H_i را به صورت زیر محاسبه می‌کنیم:

$$H_i = \frac{1}{n-1} \#\{j \neq i : (X_{j1}, \dots, X_{jp}) \leq (X_{i1}, \dots, X_{ip})\}.$$

۲. این H_i های به دست آمده را به صورت $H_{(1)} \leq \dots \leq H_{(n)}$ باید مرتب شوند، که این مرحله با حالت دوم متغیره یکسان است.

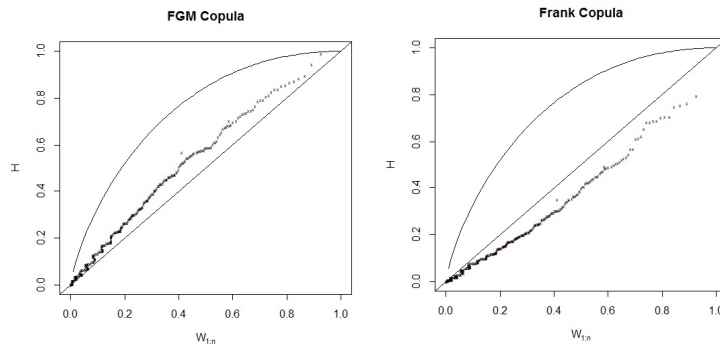
۳. زوج‌های $(W_{i:n}, H_{(i)})$ را رسم می‌کنیم. $W_{i:n}$ رتبه i ام از یک نمونه تصادفی به حجم n از توزیع مجانبی H_i تحت فرضیه استقلال دوبه‌دو در بین p مؤلفه است. $W_{i:n}$ مانند (۶) محاسبه می‌شود، با این تفاوت که باید از K_0 زیر استفاده کرد:

$$K_0(w) = w + w \sum_{k=1}^{p-1} \frac{1}{k!} \log^k(1/w), \quad 0 \leq w \leq 1.$$

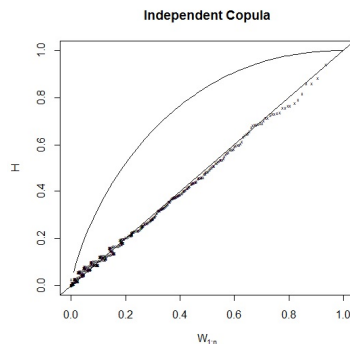
برای تفسیر نمودار K بر روی شکل، به مفصل‌های FGM و فرانک توجه کنید. این بار نمودار K را برای این دو مفصل با همان مقدار پارامتر بیان شده در زیر بخش قبل، در شکل ۳ رسم کردیم. با توجه به شکل سمت چپ که مربوط به مفصل FGM است، برداشت می‌شود که هر چه میزان وابستگی مثبت میان داده‌ها قوی‌تر باشد، نقاط بالای خط نیمساز ربع اول قرار خواهند گرفت. همانند نمودار سمت راست در شکل ۳ از مفصل فرانک، هر چه ساختار وابستگی منفی میان داده‌ها بیشتر باشد، نقاط در پایین این خط نیمساز واقع خواهند شد. همان‌گونه که در شکل ۴ مشخص است، اگر مشاهدات فاقد هرگونه وابستگی باشند، آنگاه داده‌ها نه به سمت بالا و نه به سمت پایین متمایل خواهند بود. در مواقعی که داده‌ها روی خط واقع شوند، میان آن‌ها استقلال وجود دارند.

۴ مثال‌ها

در این بخش برای تشریح بهتر از ابزارهای نموداری در تشخیص ساختار وابستگی داده‌های وابسته با استفاده از چند مثال بررسی



شکل ۳: به ترتیب از چپ به راست نمودار K مربوط به مفصل FGM و فرانک



شکل ۴: نمودار K مفصل استقلال

خواهد شد.

نمودار K برای شبه مشاهدات نمونه تصادفی به حجم $n = ۱۰۰$ از مفصل C_α با $\alpha = ۳/۴, ۱/۲, ۱/۴$ رسم کردیم.

مثال ۳.۴. برای توضیح حالت سه متغیره نمودار K از شبه مشاهدات یک نمونه تصادفی به حجم $n = ۱۰۰$ از یک بردار (X_1, X_2, X_3) که هر مؤلفه‌اش دارای توزیع یکنواخت بر بازه $[0, 1]$ است، استفاده می‌کنیم. در شکل ۷ سمت چپ شدیدترین نوع وابستگی را برای $X_1 = X_2 = X_3$ داریم. مفصل متناظر با آن‌ها کران بالایی فرشه $M(x_1, x_2, x_3) = \min(x_1, x_2, x_3)$ است. نمودار سمت راست، با استفاده از مشاهدات $X_1 = X_2 = 1 - X_3$ رسم شده است. این نمودار نشان‌دهنده وجود شدیدترین نوع وابستگی منفی میان داده‌ها است. اما باید توجه داشت، برای هر مفصل $p \geq ۳$ متغیره کران پایین فرشه W تابع توزیع نیست.

$$C(u_1, \dots, u_p) \geq W(u_1, \dots, u_p) \equiv \max(0, u_1 + \dots + u_p - 1)$$

مثال ۴.۴. فرض کنید X_2, X_3, X_4 و z متغیرهای تصادفی دوبه‌دو مستقل از توزیع نرمال استاندارد و $X_1 = |z| \text{sing}(X_2 X_3)$ باشند. در این صورت ساختار وابستگی بردار (X_1, X_2, X_3) بر اساس $n = ۱۰۰$ مشاهده از مفصل متناظر به دست می‌آید. فرض کنید مؤلفه اول بردار (X_1, X_2, X_3) به صورت $X_1 = |z| \text{sing}(X_2 X_3)$ باشد. متغیرهای X_2, X_3 و z دوبه‌دو مستقل از توزیع نرمال

مثال ۱.۰۴. به عنوان اولین مثال فرض کنید که یک زوج (X, Y) از متغیرهای تصادفی هم‌یکنوا داریم. اگر $Y = G^{-1}\{F(X)\}$ باشد، مفصل متناظر (X, Y) کران بالایی فرشه $M(u, v) = \min(u, v)$ (نمودار سمت راست) خواهد بود و اگر $Y = G^{-1}\{1 - F(X)\}$ باشد، مفصل (X, Y) ، کران پایینی فرشه $W(u, v) = \max(0, u + v - 1)$ (نمودار سمت چپ) است. در شکل ۵ نمودار K برای نمونه به حجم $n = ۱۰۰$ از شبه مشاهداتی تهیه شده است، که حاشیه‌های F و G آن از توزیع یکنواخت بر بازه $[0, 1]$ است و توزیع توأم آن‌ها مفصل ذکر شده می‌باشد.

مثال ۲.۰۴. خانواده مفصل کلاسیون را در نظر بگیرید.

$$C_\alpha(u, v) = (u^{-\alpha} + v^{-\alpha} - 1)^{-1/\alpha}, \quad 0 \leq u, v \leq 1, \quad \alpha > 0$$

ضریب کندال τ برای این خانواده برحسب تابعی از α به شکل $\tau(X, Y) = \frac{\alpha}{\alpha+2}$ است، که می‌توان آن را با استفاده از α مرتب کرد:

$$\alpha \leq \alpha^* \Rightarrow C_\alpha \prec_k C_{\alpha^*}.$$

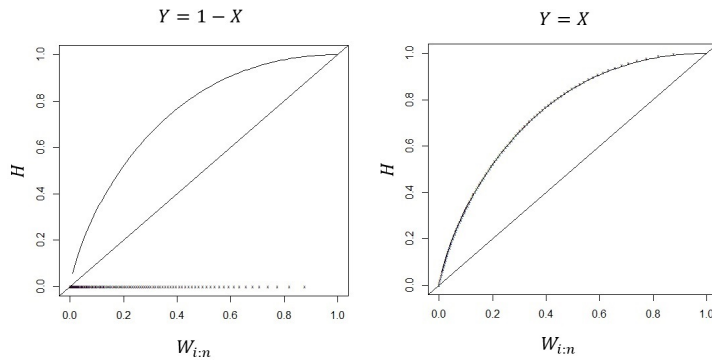
بر اساس تابع مفصل ارشمیدسی، توزیع کندال متناظر برای $0 \leq w \leq 1$ به شکل $K_\alpha(w) = w + \frac{w}{\alpha}(1 - w^\alpha)$ تعیین می‌شود. در شکل ۶

مثال ۵.۴. در شکل ۸ سمت چپ نمودار پراکنش یک نمونه شیب‌سازی شده از توزیع نرمال دومتغیره با ضریب همبستگی ۰ مطابق (۷) را رسم کردیم. در این نمودار مشاهده می‌شود، نمونه در شعاع ۱/۵ میانگین پراکنده شده است. همچنین در سمت راست نمودار کای داده‌ها، حول ۰ مرکزی هستند. کاملاً مشخص است که داده‌ها دارای ساختار وابستگی مثبت ولی ضعیف هستند.

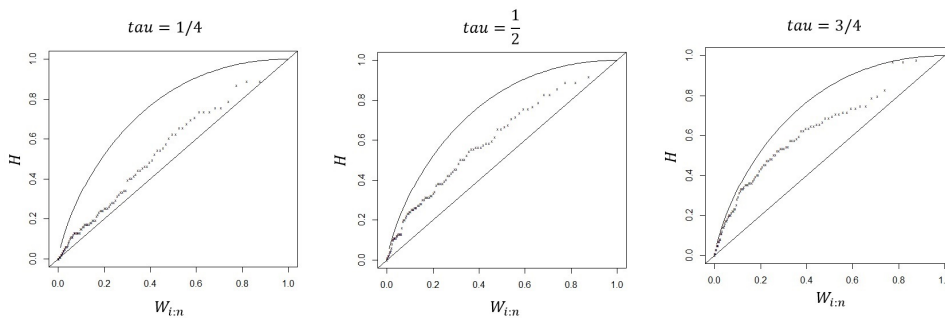
استاندارد شیب‌سازی شده‌اند. X_i ها دوبه‌دو مستقل‌اند، ولی توأمًا مستقل نمی‌باشند. نمودار K بر اساس شبه مشاهدات از نمونه تصادفی به حجم $n = 100$ از این توزیع در شکل ۹ رسم شد. در نمودار سه متغیره به دلیل وجود استقلال نقاط حول نیمساز پراکنده شده‌اند. اما در نمودارهای دومتغیره، داده‌ها کمی به بالا و به پایین متمایل هستند. لازم به ذکر است که تنها با مشاهده از روی نمودارها نمی‌توان در مورد وابستگی به‌طور کامل استدلال کرد ولی یک ابزار مفید برای آشنایی اولیه با داده‌ها است.

$$f(x, y) = \frac{1}{\sqrt{2\pi}\sigma_x\sigma_y} \exp\left\{-\frac{1}{2}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 + \left(\frac{y-\mu_y}{\sigma_y}\right)^2\right]\right\},$$

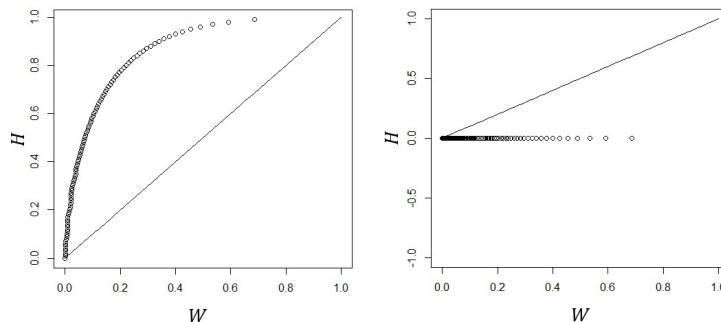
$$x, y \in \mathcal{R}$$



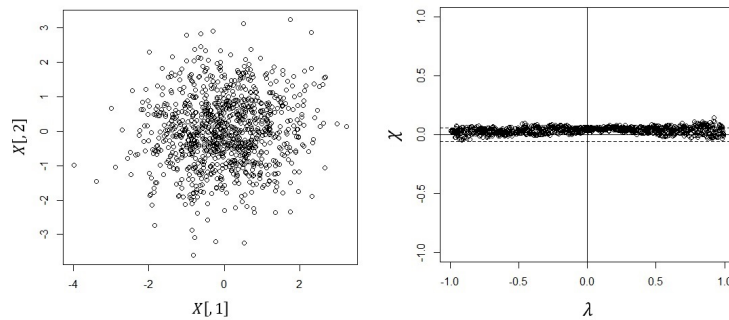
شکل ۵: به ترتیب از چپ به راست نمودارهای مربوط به کران پایین و کران بالای فرشه



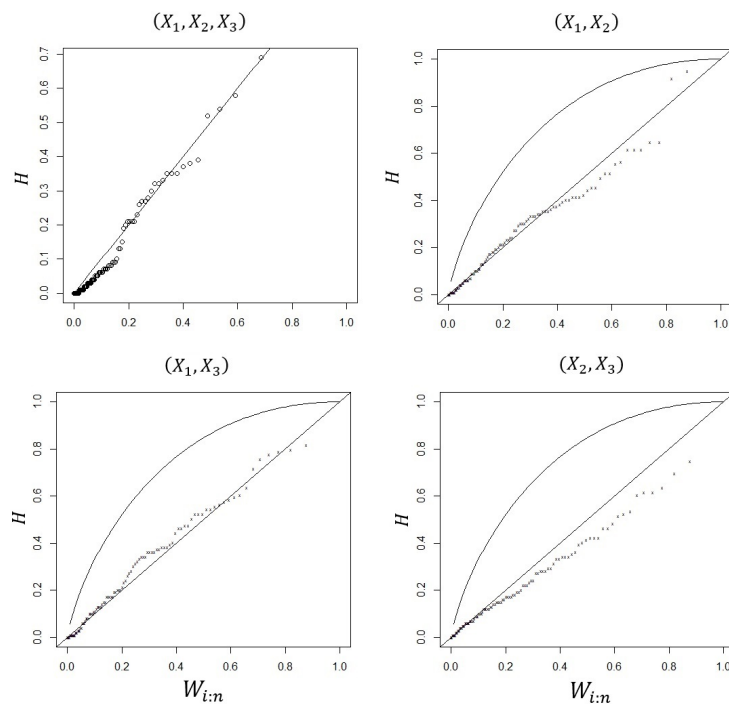
شکل ۶: نمودار K



شکل ۷: به ترتیب از چپ به راست نمودار مربوط به شدیدترین نوع وابستگی مثبت و منفی



شکل ۸: به ترتیب از چپ به راست نمودارهای پراکنش و کای از توزیع نرمال دومتغیره



شکل ۹: نمودارهای K برای حالت سه متغیره و دومتغیره از توزیع نرمال

نمودارهای پراکنش و کای را در اشکال ۱۰ و ۱۲ برای زوج متغیره‌ها نشان دادیم، تا رفتار دوبه‌دوی متغیره‌ها را نسبت به هم مشاهده کنیم. همچنین نمودار K را در حالت سه متغیره در شکل ۱۱ رسم کرده‌ایم. با یک نگاه کلی به نمودارهای ذکر شده، این برداشت حاصل می‌شود که یک نوع وابستگی مثبت میان داده‌ها وجود دارد. تنها با مشاهده این سه نمودار نمی‌توان در مورد میزان و نوع وابستگی میان داده‌های بورس قضاوت کرد. بنابراین احتیاج به ابزارهای دیگری از جمله آزمون‌ها و اندازه‌های مختلف داریم. حال برای اینکه از وجود وابستگی بین متغیره‌ها اطمینان حاصل کنیم، آزمون کرامرون مایس، برای فرضیه استقلال میان متغیره‌ها را انجام می‌دهیم. با توجه به جدول ۱، فرضیه استقلال میان متغیره‌ها رد می‌شود و هیچ استقلالی میان این سه متغیره بورس وجود ندارد و داده‌ها شدیداً به هم وابسته هستند. ضرایب τ کندال و ρ اسپیرمن جفتی را در جدول ۲ برای

۵ تحلیل داده‌های بورس بر اساس ویژگی‌هایی از مفصل

در این بخش با بررسی رابطه بین سه متغیره ضعف مالی، سود انباشته، و دارایی ملموس ۱۱۰ شرکت تجاری می‌خواهیم تابع مفصل سه‌بعدی داده‌های بورس را برآورد کنیم. ابتدا آزمون کرامرون مایس برای فرضیه استقلال میان داده‌ها را انجام می‌دهیم، تا از عدم وجود استقلال میان داده‌ها اطمینان حاصل کنیم. با رسم نمودار پراکنش، K ، و کای و سپس با استفاده از آزمون‌ها و محاسبه معیار آکائیکه و مقدار تابع درست‌نمایی مفصل سه‌بعدی را برآورد می‌کنیم. سپس مقادیر τ کندال و ρ اسپیرمن جفتی و جهت‌ها را برای داده‌ها محاسبه و بررسی کردیم. اندازه‌های وابستگی دمی بالا و پایین را نیز با استفاده از مفصل تجربی به دست می‌آوریم.

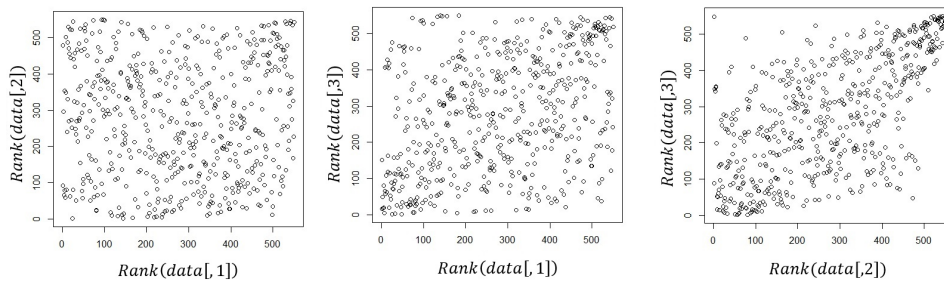
داده‌ها ساختار وابستگی مثبت وجود دارد. بنابراین از نمودارها، آزمون‌ها، و اندازه‌های وابستگی مختلف به این نتیجه رسیدیم که میان داده‌ها یک نوع وابستگی مثبت وجود دارد و مفصل مناسب داده‌ها، مفصل تی است.

۶ بحث و نتیجه‌گیری

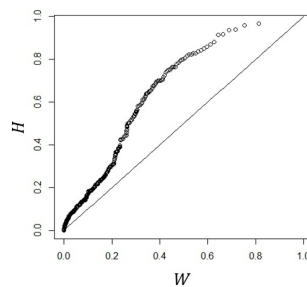
هدف این مقاله تحلیل سه متغیره از داده‌های بورس ۱۱۰ شرکت تجاری بوده است. برای این منظور نیازمند ابزارهای مختلفی هستیم تا کار تحلیل آسان شود. از آنجایی که داده‌ها به هم وابسته بودن، مفصل را برای این کار مناسب دیدیم. ولی سؤال اینجاست که چه مفصلی و با چه نوع ساختار وابستگی باید برای این داده‌ها انتخاب کنیم. از این رو در ابتدای مطالعه باید از چند نمودار برای مشاهده ساختار وابستگی داده‌ها استفاده کرد. بنابراین یک بخش را به معرفی و بیان سازوکار چند نمودار مناسب پرداختیم. ولی تنها با مشاهده آن‌ها نمی‌توان به نتیجه‌گیری نهایی و انتخاب مفصل مناسب رسید. بنابراین نیاز به ابزارهای دیگری است. در اینجا از آزمون نیکویی برازش کرامر ون مایس برای چند مفصل استفاده کردیم. مفصل‌های مورد آزمون گامبل، کلایتون، فرانک، جو، تی، و نرمال بوده است. مشاهده می‌شود که تنها مفصل تی به داده‌های بورس برازش شد. فقط استفاده از این آزمون برای استقلال کافی نیست، پس مقادیر آکائیکه و تابع درستنمایی را به کار گرفتیم و آن‌ها را باهم مقایسه کردیم. در این مرحله نیز مفصل تی مناسب داده‌ها بود. بعد از شناسایی مفصلی که به داده‌های سه متغیره بورس برازش شد، چند اندازه وابستگی مختلف اعم از اندازه وابستگی جهتی و دمی برای تحلیل دقیق‌تر بر داده‌ها محاسبه کردیم. با به‌کارگیری اندازه وابستگی جهتی می‌توان رفتار سه‌تایی متغیرها را نسبت به هم یافت. یعنی با کاهش و افزایش هر کدام از متغیرها می‌توان پیش‌بینی کرد که متغیر دیگر چه رفتاری خواهد داشت. اندازه وابستگی دمی، در مورد رفتار داده‌ها در دم اطلاعات می‌دهد. بنابراین در این مقاله یک روش بررسی برای داده‌های وابسته سه متغیره را به تفصیل بیان کردیم و توانستیم یک مفصل به آن‌ها برازش کنیم. با استفاده از این مفصل و قضیه اسکالر می‌توان تابع توزیع این متغیرها را محاسبه کرد و با آن فاصله اطمینان و آزمون فرض‌های مختلف و غیره را برای داده‌های وابسته چند متغیره به راحتی انجام داد.

تمام زوج متغیرها محاسبه کرده‌ایم. این مقادیر نشان‌دهنده‌ی یک نوع وابستگی مثبت میان داده‌های بورس هستند. این زمان که از وابستگی میان داده‌ها اطمینان حاصل کردیم، باید بررسی کنیم که کدام مفصل مناسب آن‌ها می‌باشد. از این رو آزمون نیکویی برازش کرامر ون مایس را برای مفصل‌های گامبل، کلایتون، فرانک، جو، تی، و نرمال انجام دادیم. با توجه به p - مقدار به دست آمده، فقط مفصل تی برای داده‌ها مناسب است. نتایج در جدول ۳ نشان داده شده است. حال مقادیر آکائیکه و مقدار تابع درستنمایی را برای داده‌ها محاسبه کردیم، تا از نتیجه‌گیری ناشی از آزمون استقلال کرامر ون مایس اطمینان حاصل کنیم. آکائیکه از فرمول $AIC = 2k - 2 \ln l$ به دست آمده است. در آن k تعداد متغیرها و l مقدار تابع درستنمایی می‌باشد. هر چه مقدار آکائیکه کمتر و مقدار تابع درستنمایی بیشتر باشد، مدل مناسب‌تر است. این بار نیز مفصل تی مناسب بود و تأییدی بر درستی بر درستی نتایج آزمون است. نتایج در جدول ۴ است.

ضرایب τ کندال و ρ اسپیرمن جهتی را در جدول ۵ نمایش دادیم. همان‌طور که قبلاً گفته شد، برای محاسبه اندازه وابستگی جهتی، ابتدا درستی قضیه ۲ از مقاله کوئسادا و یویدا (۲۰۱۲) را باید بررسی کنیم. آن‌ها در آن قضیه با استفاده از مفصل سه‌بعدی و حاشیه‌های دوطرفی در مورد جهت وابستگی مثبت نتیجه‌گیری می‌کردند. داده‌های این مقاله فقط در پنج جهت از قضیه مذکور صدق کردند، بنابراین اندازه وابستگی را فقط برای این پنج جهت محاسبه می‌کنیم. مثبت بودن اندازه وابستگی جهتی کندال و اسپیرمن برای جهت $(-1, -1, -1)$ بدان معناست که سه متغیر مورد بررسی هم‌راستا باهم کاهش می‌یابد. همچنین این اندازه برای جهت $(1, 1, 1)$ نیز مثبت می‌باشد، یعنی اگر ضعف مالی را کاهش دهیم، سود انباشته و دارایی ملموس بیشتری خواهیم یافت. این دو اندازه وابستگی برای سه جهت دیگر، منفی شده است. برای جهت $(-1, 1, -1)$ و $(1, -1, 1)$ به این صورت تفسیر می‌شود، که اگر ضعف مالی و دارایی ملموس را کاهش دهیم، سود انباشته نیز کاهش می‌یابد. در جهت $(1, 1, -1)$ با کاهش دارایی ملموس به ضعف مالی کمتر و سود انباشته کمتری می‌رسیم و این فقط به دلیل منفی بودن این اندازه وابستگی جهتی می‌باشد. نتایج حاصل از محاسبه اندازه وابستگی دمی در جدول ۶ آمده است. همان‌طور که می‌دانیم این اندازه عددی بین ۰ و ۱ است. اندازه ضریب وابستگی دمی پایین و بالا به عدد ۱ نزدیک‌تر است. یعنی میان



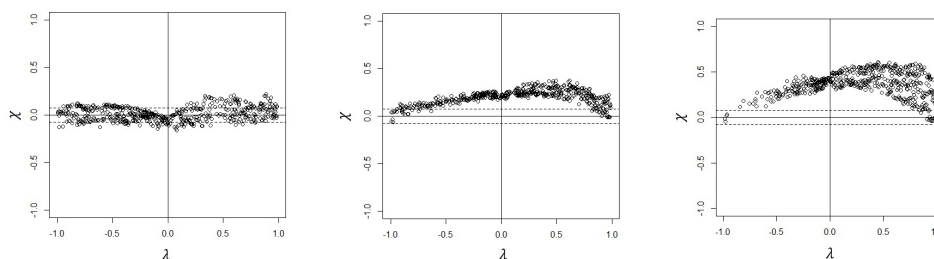
شکل ۱۰: نمودار پراکنش داده‌های بورس به ترتیب از چپ به راست برای داده ضعف مالی و سود انباشته، ضعف مالی و دارایی ملموس، و سود انباشته و دارایی ملموس می‌باشد.



شکل ۱۱: نمودار K داده‌های بورس

مراجع

- [۱] بدافلور، ر.، خیری، س.، سدهی، م.، و آخوند، م. (۲۰۱۵). بررسی عوامل مؤثر بر دفع پیوند قرنیه در بیماران قوز قرنیه با به‌کارگیری مدل‌بندی بیزی خطر و ایبول دومتغیره بر اساس تابع مفصل کلاسیون. تحقیقات نظام سلامت، ۱۰(۴): ۷۹۵-۸۰۵.
- [۲] صالحی، م.، و زمانی مقدم، س. (۱۳۹۵). تأثیر وابستگی ساختاری بر تغییرات مرز کارا پرتفوی و مقایسه آن با روش سنتی در بورس اوراق بهادار تهران: توابع مفصل. دانش مالی تحلیل اوراق بهادار (مطالعات مالی)، ۲۹(۹): ۲۲۵۱-۶۸۵۹.
- [۳] شاکری، م.، قوامی قنبرآبادی، و.، اسماعیلی، ح.، جباری نوقابی، ه.، همایی شاندیز، ف.، و باغی‌شنی، ح. (۱۳۹۰). تحلیل داده‌های بقای بیماران سرطان پستان، با مدل خطرهای متناسب کاکس برای سانسور وابسته بر اساس تابع مفصل. مجله دانشگاه علوم پزشکی خراسان شمالی (ویژه‌نامه آمار زیستی و اپیدمیولوژی)، ۳(۳): ۲۰۰۸-۸۷۰۱.
- [۴] گنجعلی‌خانی، م.، ذونعمتی کرمانی، م.، رضاپور، م.، و رهنما، م. (۱۳۹۵). ارزیابی عملکرد تابع مفصل در پهنه‌بندی کیفی آب‌های زیرزمینی، مطالعه موردی دشت‌های کرمان و راور. تحقیقات آب‌و‌خاک ایران (علوم کشاورزی ایران)، ۴۷(۳): ۴۷۹-۴۷۹.



شکل ۱۲: نمودار کای داده‌های بورس به ترتیب از چپ به راست برای داده ضعف مالی و سود انباشته، ضعف مالی و دارایی ملموس، و سود انباشته و دارایی ملموس می‌باشد.

جدول ۱: آزمون کرامر ون مایس برای فرضیه استقلال میان متغیرها

| Tipper's p-value | Fisher's p-value | p-value | آماره کرامر ون مایس |
|------------------|------------------|-------------|---------------------|
| ۰/۰۰۰۰۴۹۹۵۵ | ۰/۰۰۰۰۴۹۹۵۵ | ۰/۰۰۰۰۴۹۹۵۵ | ۱,۴۳۵۴۱۷ |

جدول ۲: ضرایب وابستگی برای زوج متغیرها

| زوج متغیرها | τ کندال | ρ اسپیرمن |
|--------------|--------------|----------------|
| (X_1, X_2) | ۰/۰۰۵۹۵ | ۰/۰۱ |
| (X_1, X_3) | ۰/۲۶ | ۰/۳۷ |
| (X_2, X_3) | ۰/۴۵۸ | ۰/۶۳ |

جدول ۳: آزمون نیکویی برازش برای چند مفصل

| مفصل | p-value | مقدار پارامتر | آماره آزمون |
|---------|------------|---------------|-------------|
| گامبل | ۰/۰۰۰۰۴۹۹۵ | ۱,۴۰۷۳ | ۰/۰۷۳۴۳ |
| کلایتون | ۰/۰۰۰۰۴۹۹۵ | ۰/۸۱۴۶۲ | ۰/۲۰۱۵۲ |
| فرانک | ۰/۰۰۰۰۴۹۹۵ | ۲,۵۴۷۵ | ۰/۱۱۱۶۳ |
| جو | ۰/۰۰۰۰۴۹۹۵ | ۱,۷۴۵۲ | ۰/۰۹۱۵۵۵ |
| تی | ۰/۲۱۶۳ | ۰/۳۴۹۵۲ | ۰/۵۸۵۲۳ |
| نرمال | ۰/۰۰۰۰۴۹۹۵ | ۰/۳۵۷۸۵ | ۰/۱۰۰۴۹ |

[۵] عباسیان، م.، جلالی، س.، و موسوی ندوشنی، س.س. (۱۳۹۳). تحلیل فراوانی چندمتغیره سیلاب با استفاده از تابع مفصل و توزیع‌های حاشیه‌ای پارامتری و ناپارامتری. مهندسی عمران مدرس، ۱۴(۴): ۸۱-۹۲.

[۶] محمودیان، ب.، محمدزاده درودی، م.، و شهبازی، ل. (۱۳۹۳). مدل بندی مقادیر کرانگین فضایی با میدان تصادفی و تابع مفصل. نشریه علوم دانشگاه خوارزمی، ۱۴(۲): ۱۲۷-۱۴۰.

[7] Benes, V., and Stephan, J. (1997). *Distributions with Given Marginals and Moment Problems*. Kluwer, Dordrecht, Netherlands.

[8] Chen, S. X., and Huang, T. M. (2007). Nonparametric estimation of copula functions for dependence modelling. *Canadian Journal of Statistics*, **35(2)**, 265-282.

[9] De Luca, G., and Riveccio, G. (2012). Multivariate tail dependence coefficients for archimedean copulae. *In Advanced Statistical Methods for the Analysis of Large Data-Sets* (Springer Berlin Heidelberg), 287-296.

[10] Fisher, N. I., and Switzer, P. (1985). Chi-plots for assessing dependence. *Biometrika*, **72(2)**, 253-265.

[11] Fisher, N. I., and Switzer, P. (2001), *Statistical computing and graphics*, **55**, 233-238.

[12] Genest, C., and Boies, J. C. (2003). Detecting dependence with Kendall plots. *The American Statistician*, **57(4)**, 275-284.

[13] Genest, C., and Rivest, L. P. (1993). Statistical inference procedures for bivariate Archimedean copulas. *Journal of the American Statistical Association*, **88(423)**, 1034-1043.

جدول ۴: مقادیر آکائیکه و تابع درستنمایی

| مقدار تابع درستنمایی | AIC | مفصل |
|----------------------|-----------|---------|
| ۹۷,۴۶۵۶ | -۲۰۱,۰۲۸۲ | گامبل |
| ۱۵,۷۳۴۵۲ | ۴۹,۲۰۲۸ | کلایتون |
| ۸۵,۷۷۹۱ | -۱۸۸,۳۱۴۷ | فرانک |
| ۶۹,۰۵۳۵ | -۱۶۴,۲۴۴۸ | جو |
| ۱۱۵,۹ | -۲۲۰,۷۸۸۷ | تی |
| ۹۳,۵۲۶۵۸ | -۱۸۱,۶۴۹۶ | نرمال |

جدول ۵: ضرایب وابستگی جهتی کندال و اسپیرمن

| جهت | τ کندال | ρ اسپیرمن |
|--------------|--------------|----------------|
| (-۱, -۱, -۱) | ۰,۲۴۱۱۶۴۲ | ۰,۳۲۸۲۶۰۷ |
| (-۱, ۱, -۱) | -۰,۰۶۸۰۳۹۳ | -۰,۰۸۱۵۹۴ |
| (-۱, ۱, ۱) | ۰,۶۴۰۷۳۸۲ | ۰,۰۷۴۹۲۷۳۳ |
| (۱, -۱, ۱) | -۰,۰۶۸۰۳۹۳ | -۰,۰۹۸۴۰۶ |
| (۱, ۱, -۱) | -۰,۲۳۷۱۹۸۷ | -۰,۳۳۸۴۰۶ |

جدول ۶: ضرایب وابستگی دمی

| λ | ضریب وابستگی دمی بالا | ضریب وابستگی دمی پایین |
|-----------|-----------------------|------------------------|
| ۱/۲۳ | ۰,۵۹۸۹۵۸۳ | ۰,۶۰۱۰۳۶۳ |
| ۲/۱۳ | ۰,۶۹۲۷۷۱۱ | ۰,۶۹۰۴۷۶۲ |
| ۳/۱۲ | ۰,۸۹۱۴۷۲۹ | ۰,۸۸۵۴۹۶۲ |
| ۱۲/۳ | ۰,۴۱۸۱۸۱۸ | ۰,۰۵۴۷۹۱۶۴ |

- [14] Genest, C., and Rivest, L. P. (2001). On the multivariate probability integral transformation. *Statistics and Probability Letters*, **53**(4), 391-399.
- [15] Gudendorf, G., and Segers, J. (2010). Extreme-value copulas. In *Copula Theory and Its Applications (Springer Berlin Heidelberg)*, 127-145.
- [16] Mortezaejad, S. A. F., Borzadaran, G. M., and Sadeghpour Gildeh, B. (2019). Joint dependence distribution of data set using optimizing Tsallis copula entropy. *Physica A: Statistical Mechanics and its Applications*, **533**, 121897-121911.
- [17] Nelsen, R. B. (2006). *An Introduction to Copulas*, Ser. Lecture Notes in Statistics. New York: Springer.
- [18] Nelsen, R. B., and Úbeda-Flores, M. (2012). Directional dependence in multivariate distributions. *Annals of the Institute of Statistical Mathematics*, **64**(3), 677-685.
- [19] Quesada, M. J. J., and Úbeda, F. M. (2012). Directional dependence of random vectors. *Information Sciences*. **215**, 67-74.
- [20] Rüschendorf, L. (1996). On c-optimal random variables. *Statistics and Probability Letters*. **27**, 267-270.
- [21] Schweizer, B., and Taylor, M. D. (1996). *Distributions with Fixed Marginals and Related Topics*. IMS.
- [22] Schweizer, M. (2001). A guided tour through quadratic hedging approaches. In *Option Pricing, Interest Rates and Risk Management*, ed. by Jouini, E; Cvitanic, J; Musiela, M. (Cambridge University Press. Cambridge. UK.) 538-574.
- [23] Sklar, M. (1959). Fonctions de Répartition à n Dimensions et Leurs Marges. Université Paris 8.

Analysis of stock data based on copula function

Seyedeh Azadeh Fallah Mortezaejad ¹, Gholamreza Mohtashami Borzadaran²,
Bahram Sadeghpour Gildeh³, and Mohammad Amini ⁴

Department of Statistics and Order and Spatial Data Center for Excellent, School of Mathematical Sciences,
Ferdowsi University of Mashhad, Iran.

Abstract:

A copula function is a useful tool in identifying the dependency structure of dependent data and thus fitting an appropriate joint distribution to the data. In this paper, stock exchange data, including three variables of financial weakness, accumulated profit, and tangible assets related to 110 Iranian companies from 2006 to 2010, are analyzed relating to the copula functions, and especially a three-dimensional distribution is fitted into this data. We use various tools to examine the type of dependencies between data, including scatter plots, chi-square charts, and Kendall charts. We also consider the directional and tail dependencies of the data and calculate the dependence coefficients of Kendall and Spearman. Finally, we perform a goodness of fit test for some well-known copulas to get the appropriate copula function of the available stock data of this article.

Keywords: Copula function, Tail and directional dependency, Chi- and Kendall plot.

¹azadeh.fallah@mail.um.ac.ir

²Corresponding author: grmohtashami@um.ac.ir

³sadeghpour@um.ac.ir

⁴m-amini@um.ac.ir