

مقایسه آزمون‌های نیکویی برازش برای تابع مفصل

زینب بهباش^۱، غلامعلی پرهام^۲

تاریخ دریافت: ۱۳۹۴/۱۱/۲۴

تاریخ پذیرش: ۱۳۹۵/۱۲/۲۵

چکیده:

توابع مفصل می‌توانند ساختار وابستگی بین متغیرها را به صورت مدل نشان دهند. تابع مفصل مناسب برای یک کاربرد خاص، تابعی است که به بهترین نحو ممکن وابستگی بین داده‌ها را نشان دهد. آزمون‌های نیکویی برازش از دیدگاه نظری بهترین روش انتخاب تابع مفصل می‌باشند. روش‌های آزمون نیکویی برازش متعددی برای توابع مفصل پیشنهاد شده است. در این مقاله در پی انتخاب روش‌های آزمون نیکویی برازش مناسب، سه روش متفاوت آزمون نیکویی برازش برای تابع مفصل را مورد بررسی و مقایسه عددی قرار داده، به بررسی نقاط قوت و ضعف آن‌ها نسبت به یکدیگر می‌پردازیم. در نهایت روش‌های آزمون مطرح‌شده را با استفاده از داده‌های شاخص بورس اوراق بهادار تهران تحلیل قرار می‌کنیم.

واژه‌های کلیدی: تابع مفصل، آزمون نیکویی برازش.

۲ تابع مفصل

۱ مقدمه

تابع مفصل، تابعی چندمتغیره است که توابع توزیع چندمتغیره را به توزیع‌های حاشیه‌ای آن‌ها پیوند می‌دهد و دارای حاشیه‌های یکنواخت بر بازه $[0, 1]$ می‌باشد.

تعریف ۱.۲. یک تابع مفصل، تابعی مانند $I^2 \rightarrow I : C$ با خصوصیات زیر است:

آ (زمین گیر^۴ است؛ یعنی، برای هر $u, v \in I$ داریم $C(u, 0) = C(0, v) = 0$ ؛

ب) برای هر u و v در I ، داریم $C(u, 1) = u$ و $C(1, v) = v$ ؛

پ) C تابعی دوسعودی است؛ یعنی، برای هر u_1, u_2, v_1, v_2 در I ، اگر $u_1 \leq u_2$ و $v_1 \leq v_2$ باشد، داریم

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

قضیه ۲.۲ (اسکلار [۱۵]). فرض کنید $H(x, y)$ تابع توزیع توأم با حاشیه‌های F و G باشد. در این صورت تابع مفصلی مانند C وجود دارد، به طوری که برای هر $x, y \in \bar{R}$ داریم:

$$H(x, y) = C(F(x), G(y)),$$

توابع مفصل^۳ توابعی هستند که توابع توزیع چندمتغیره را به توزیع‌های حاشیه‌ای پیوند می‌دهند. به عنوان یکی از راه‌های اصلی مدل‌بندی وابستگی، می‌توان توابع مفصل را ابزاری توانا در تحلیل سری‌های چندمتغیره برشمرد. بنا بر این، این مسئله که آیا ساختار وابستگی داده‌ها را می‌توان به وسیله انتخاب یک مفصل به طور تقریبی مدل‌بندی کرد، موضوعی بااهمیت است. چنانچه تابع مفصل به درستی انتخاب نشود، می‌تواند اثر جدی روی تحلیل آماری بگذارد و باید به نحوی مطلوب، تابع مفصلی تعیین گردد که برازنده داده‌های نمونه حاضر باشد. به طور کلی هدف آزمون نیکویی برازش، این است که نشان دهد آیا نمونه از یک توزیع خاص پیروی می‌کند. روش نیکویی برازش برای تابع مفصل یک حالت خاص از مسئله کلی‌تر آزمون مدل‌های چگالی چندمتغیره است؛ اما به دلیل توزیع‌های حاشیه‌ای نامعین پیچیده است و p -مقدارها معمولاً به وسیله شبیه‌سازی محاسبه می‌شوند.

^۱ دانش‌آموخته کارشناس ارشد آمار دانشگاه شهید چمران اهواز، ایران

^۲ گروه آمار دانشگاه شهید چمران اهواز، ایران

^۳ copula

^۴ grounded

آزمون می‌شود، ایجاد می‌شوند. آزمون‌های نیکویی برازش برای توابع مفصل، همتای اغلب آزمون‌های نیکویی برازش توزیع چندمتغیره هستند که در آن‌ها F_n و F به ترتیب با C_n و C جایگذاری شده‌اند.

یا به‌طور معکوس اگر C یک تابع مفصل و F و G توابع توزیع باشند، H یک تابع توزیع توأم با حاشیه‌های F و G است. اگر F و G پیوسته باشند، C یکتا است.

فرض کنید H یک تابع توزیع دومتغیره با حاشیه‌های پیوسته F و G باشد. اگر $F^{(-1)}$ و $G^{(-1)}$ به ترتیب شبه‌معکوس دو تابع F و G باشند، طبق قضیه اسکالر، رابطه زیر بین تابع توزیع توأم و تابع مفصل برقرار است:

$$C(U, V) = H(F^{-1}(U), G^{-1}(V)),$$

که $U = F(X)$ و $V = G(Y)$ متغیرهای تصادفی با توزیع یکنواخت استاندارد هستند. به این طریق می‌توان برای هر توزیع توأم، یک تابع مفصل تعیین کرد که این روش، روش وارون نامیده می‌شود. برای هر $u \in [0, 1]^d$ توابع مفصل تجربی به‌صورت

$$C_n(u) = F_n(F_{n,1}^{(-1)}(u_1), \dots, F_{n,d}^{(-1)}(u_d))$$

تعریف می‌شود. بنا بر این تابع مفصل تجربی، یک تابع از تابع توزیع تجربی است. توزیع‌های حاشیه‌ای تجربی برای هر $i = 1, \dots, n, X_i \in R^d$ به‌صورت

$$F_{n,k}(x_k) = n^{-1} \sum_{i=1}^n I(X_{i,k} \leq x_k), k = 1, \dots, d$$

به‌دست می‌آیند و تابع توزیع تجربی توأم به‌صورت زیر است:

$$F_n(x) = n^{-1} \sum_{i=1}^n I(X_i \leq x).$$

اغلب آزمون‌های نیکویی برازش با در نظر گرفتن انواعی از فاصله‌ها بین تابع توزیع تجربی F_n و F که منجر به یک آماره

۳ توابع مفصل و اندازه‌های وابستگی

وابستگی بین متغیرهای تصادفی، یکی از گسترده‌ترین موضوعات مورد مطالعه در احتمال و آمار است، تا آن‌جا که بدون در نظر گرفتن نوع وابستگی متغیرها مدل‌بندی مقدور نیست. توابع مفصل، ابزاری مفید در تحلیل ساختار وابستگی هستند. نلسن [۱۱] با استفاده از توابع مفصل به تحلیل همبستگی و ضرایب مربوط به آن‌ها پرداخت. معمولاً برای محاسبه میزان وابستگی بین متغیرها از سه معیار ρ پیرسون^۵ و τ کندال^۶ و ρ_s اسپیرمن^۷ استفاده می‌شود. در حالی که ضریب همبستگی پیرسون دارای پیش‌فرض نرمال بودن است، ضرایب همبستگی ρ_s اسپیرمن و τ کندال معیارهایی براساس رتبه می‌باشند و بنا بر این در شرایطی که متغیرها غیر نرمال هستند مناسب‌ترند. علاوه بر پیش‌نیاز بودن τ کندال در تعیین پارامتر وابستگی در یک تابع مفصل، وابسته نبودن τ کندال به توزیع‌های حاشیه‌ای، انگیزه ما را در ادامه کار برای انتخاب این ضریب همبستگی بیشتر می‌کند. ضرایب همبستگی ساختار وابستگی کلی را نشان می‌دهند و براساس تابع توزیع توأم قابل محاسبه و بیان می‌باشند. در حالی که گاهی ما علاقه‌مند به بررسی وابستگی موضعی هستیم که در این حالت، از نوعی وابستگی که به وابستگی دمی یا دنباله‌ای معروف است استفاده می‌کنیم.

جدول ۱. اندازه‌های وابستگی دمی برای برخی از توابع مفصل

تابع مفصل	تابع چگالی	فضای پارامتر	وابستگی دمی بالایی	وابستگی دمی پایینی
گامبل	$\exp\left(-\left(-\log u\right)^\theta + (\log v)^\theta\right)^{\frac{1}{\theta}}$	$[1, \infty)$	$2 - 2^{\frac{1}{\theta}}$.
جو	$1 - \left[\left(1 - u\right)^\theta + \left(1 - v\right)^\theta - \left(1 - u\right)\left(1 - v\right)\right]^{\frac{1}{\theta}}$	$[1, \infty)$	$2 - 2^{\frac{1}{\theta}}$.
فرانک	$\frac{-1}{\theta} \ln \left[1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1}\right]$	$R - \{0\}$.	.
نرمال	$\int_{-\infty}^{\phi^{-1}(u)} \int_{-\infty}^{\phi^{-1}(v)} \frac{1}{\sqrt{2\pi} \sqrt{1-\rho^2}} \exp\left\{\frac{\rho uv - u^2 - v^2}{1-\rho^2}\right\} dudv$	$(-1, 1)$.	.

^۵ pearson

^۶ kendall

^۷ spearman

جدول ۲. تعداد رد فرضیه صفر تابع مفصل نرمال در ۱۰۰ بار تکرار (میانگین P -مقدارها) به وسیله روش های A_1 ([۱۴]) و A_2 ([۷]) و A_3

$$\alpha = 0.05 \text{ در سطح } ([۱۳])$$

A_3	A_2	A_1	مفصل واقعی	τ	n
۳۵(۰/۳۶۱)	۰(۰/۹۱۱)	۰(۰/۷۶۲)	نرمال	۰/۳	۳۰۰
۶۲(۰/۱۶۵)	۵(۰/۳۲۸)	۱۰(۰/۳۸۹)	فرانک		
۵۹(۰/۱۸۰)	۲(۰/۳۴۴)	۳(۰/۳۸۰)	گامبل		
۸۶(۰/۰۲۵)	۱۲(۰/۱۵۷)	۵۴(۰/۱۳۰)	جو		
۴۶(۰/۲۷۶)	۲(۰/۸۳۹)	۰(۰/۸۶۸)	نرمال	۰/۷	۳۰۰
۶۶(۰/۱۴۲)	۲۰(۰/۱۳۰)	۳۰(۰/۲۲۸)	فرانک		
۷۲(۰/۱۳۱)	۲(۰/۲۱۵)	۸(۰/۳۶۲)	گامبل		
۱۰۰(۰/۰۰۰)	۴۴(۰/۰۶۹)	۱۰۰(۰/۰۱۱)	جو		
۵۵(۰/۲۱۲)	۲(۰/۷۴۶)	۱(۰/۶۹۳)	نرمال	۰/۳	۵۰۰
۸۳(۰/۰۲۹)	۳۹(۰/۱۰۱)	۶(۰/۴۰۷)	فرانک		
۹۴(۰/۰۱۴)	۱۳(۰/۱۸۲)	۳۶(۰/۱۶۹)	گامبل		
۱۰۰(۰/۰۰۰)	۵۶(۰/۰۳۳)	۸۱(۰/۰۲۴)	جو		
۷۵(۰/۱۲۵)	۳(۰/۷۸۷)	۱(۰/۵۸۹)	نرمال	۰/۷	۵۰۰
۹۱(۰/۰۱۷)	۸۲(۰/۰۷۰)	۲۶(۰/۳۱۱)	فرانک		
۱۰۰(۰/۰۰۰)	۴۳(۰/۰۶۹)	۶۰(۰/۱۱۰)	گامبل		
۱۰۰(۰/۰۰۰)	۷۰(۰/۰۳۶)	۱۰۰(۰/۰۰۹)	جو		

تعریف ۱.۳ (برآورد τ کندال). فرض کنید وابستگی دمی، رابطه وابستگی بین مقادیر بزرگ (کوچک) از یک متغیر را با مقادیر بزرگ (کوچک) از متغیر دیگر توصیف می کند که تحت عنوان وابستگی دمی بالایی (پایینی) تعریف می شوند. در این بخش به معرفی معیار τ کندال و ارتباط آن با تابع مفصل می پردازیم.

۱.۳ کندال τ

یکی دیگر از اندازه های وابستگی، ضریب همبستگی رتبه ای کندال می باشد که به τ کندال معروف است و توسط کندال [۹] معرفی شد. اگر (x_1, y_1) و (x_2, y_2) مقادیر مشاهده شده مربوط به بردار تصادفی (X, Y) باشند، τ کندال برابر با تفاضل احتمال های هماهنگ بودن و ناهماهنگ بودن دو متغیر است و به صورت

$$\begin{aligned} \tau &= P((X_1 - X_2)(Y_1 - Y_2) > 0) \\ &\quad - P((X_1 - X_2)(Y_1 - Y_2) < 0) \in [-1, 1] \\ &= E[\text{sgn}(X_1 - X_2) \text{sgn}(Y_1 - Y_2)] \end{aligned}$$

تعریف می شود که در آن sgn تابع علامت است.

تعریف می شود. τ کندال را می توان براساس تابع مفصل به صورت

$$\tau = 4 \int \int C(u_1, u_2) dC(u_1, u_2) - 1$$

تعریف کرد. به علت پیچیدگی های تابع مفصل در اکثر موارد، محاسبه رابطه بالا همیشه آسان نیست. به همین دلیل، روش های آسان تری برای به دست آوردن τ کندال برای خانواده توابع معروف بیضوی و توابع مفصل ارشمیدسی در این بخش ارائه می شود.

۲.۳ محاسبه τ کندال برای توابع مفصل بیضوی و

ارشمیدسی

τ کندال برای خانواده مفصل بیضوی از فرمول زیر به دست می‌آید:

$$\rho = \sin\left(\frac{\pi}{4}\tau\right) \leftrightarrow \tau = \frac{2}{\pi} \sin^{-1}(\rho).$$

τ کندال برای مفصل ارشمیدسی به وسیله قضیه زیر به دست می‌آید.

قضیه ۲.۳. اگر φ یک تابع مولد از مفصل ارشمیدسی باشد، τ کندال به صورت

$$\tau = 1 + 4 \int_0^1 \frac{\varphi(t)}{\varphi'(t)} dt$$

محاسبه می‌شود که φ یک تابع نزولی نامنفی با $\varphi(0) = 1$ ، مشتق دوم نامنفی و $\varphi'(0)$ مشتق اول φ است [۱۲].

هر کدام از توابع مفصل، ساختار وابستگی متفاوتی را بر روی داده‌ها اعمال می‌کند. توابع مفصل بیضوی به دلیل طبیعت متقارن آن‌ها برای بسیاری از کارها مناسب نیستند. تابع مفصل نرمال (گوسی)^۸ متعلق به خانواده مفصل بیضوی است که عموماً در کاربرد برای این تابع مفصل، وابستگی دمی وجود ندارد. اکثر توابع مفصل به خانواده توابع مفصل ارشمیدسی تعلق دارند. توابع مفصل مربوط به این خانواده، وابستگی دمی را به خوبی نشان می‌دهند. تعدادی از توابع مفصل ارشمیدسی و وابستگی دمی مربوط به آن‌ها در جدول ۱ آمده است. تابع مفصل فرانک^۹ به دلیل متقارن بودن در دو دنباله و در برگرفتن همه حدود همبستگی در تحلیل پدیده‌هایی که دارای وابستگی مثبت و منفی هستند، استفاده می‌شود. تابع مفصل گامبل^{۱۰} وابستگی دمی بالایی را به خوبی نشان می‌دهد و تابع مفصل جو^{۱۱} دارای وابستگی دمی بالایی است.

جدول ۳. تعداد رد فرضیه صفر تابع مفصل فرانک در ۱۰۰ بار تکرار (میانگین P -مقدارها) به وسیله روش‌های A_1 ([۱۴]) و A_2 ([۷]) و A_3

([۱۳]) در سطح $\alpha = 0.05$

A_3	A_2	A_1	مفصل واقعی	τ	n
۴۵(۰/۱۲۹)	۷(۰/۲۶۳)	۳(۰/۳۹۶)	نرمال	۰/۳	۳۰۰
۳۲(۰/۲۱۴)	۲(۰/۵۰۶)	۰(۰/۷۹۹)	فرانک		
۷(۰/۱۴۰)	۵(۰/۳۴۲)	۹(۰/۳۳۲)	گامبل		
۶۸(۰/۱۶۳)	۴(۰/۳۸۹)	۴۷(۰/۲۱۵)	جو		
۵۱(۰/۰۸۵)	۷(۰/۲۱۰)	۱۶(۰/۳۶۴)	نرمال	۰/۷	۵۰۰
۴۴(۰/۱۴۰)	۳(۰/۴۱۳)	۱(۰/۷۱۴)	فرانک		
۸۹(۰/۰۴۰)	۶(۰/۲۸۹)	۶(۰/۳۳۹)	گامبل		
۷۵(۰/۰۵۲)	۲۰(۰/۲۱۱)	۱۰۰(۰/۰۱۰)	جو		
۶۸(۰/۰۷۲)	۱۳(۰/۲۰۲)	۲۳(۰/۲۱۵)	نرمال	۰/۳	۵۰۰
۴۸(۰/۱۳۶)	۲(۰/۴۸۹)	۲(۰/۴۲۵)	فرانک		
۹۲(۰/۰۳۱)	۲۴(۰/۲۱۶)	۳۵(۰/۲۲۶)	گامبل		
۹۸(۰/۰۱۱)	۱۰(۰/۲۸۱)	۹۸(۰/۰۲۱)	جو		
۸۵(۰/۰۴۳)	۸(۰/۲۵۳)	۴۰(۰/۱۷۰)	نرمال	۰/۷	۵۰۰
۶۰(۰/۱۱۵)	۵(۰/۳۲۲)	۳(۰/۳۷۳)	فرانک		
۱۰۰(۰/۰۰۰)	۵۸(۰/۱۹۷)	۸۰(۰/۰۴۹)	گامبل		
۱۰۰(۰/۰۰۰)	۹۱(۰/۱۱۵)	۱۰۰(۰/۰۱۳)	جو		

^۸ Gaussian

^۹ Frank

^{۱۰} Gumbel

^{۱۱} Joe

۴ آزمون نیکویی برازش به عنوان معیار ۱.۴ مقایسه بین آزمون‌های نیکویی برازش برای انتخاب تابع مفصل

به دلیل اهمیت انتخاب تابع مفصل مناسب، در پی انتخاب روش‌های مناسب هستیم. اخیراً مقایسه‌هایی بین برخی از روش‌های آزمون نیکویی برازش انجام شده است، از جمله برگ و باکن [۱] به بررسی ۹ آماره آزمون نیکویی برازش پرداخته‌اند که در برخی از این آزمون‌ها از روش کاهش بعد استفاده شده است. یک اشکال این روش، این است که کاهش ابعاد باعث کاهش اطلاعات شده و کاهش اطلاعات در دسترس، باعث کاهش توان آزمون خواهد شد. به هر حال، هنوز این روش‌ها وابسته به روش‌های مونتہ کارلو برای ارزیابی p -مقدارهای مربوط هستند. جنست و همکاران [۵] نیز مقایسه‌ای بین برخی آزمون‌های پوششی^{۱۲} انجام داده‌اند. براساس [۵] آزمون‌های نیکویی برازش برای تابع مفصل را می‌توان به سه گروه تقسیم کرد.

با استفاده از آزمون نیکویی برازش، مقادیر سطح معناداری توابع مفصل برازنده شده به داده‌ها تعیین می‌شود. این آزمون‌ها برای چند مفصل به کار می‌روند و هر کدام از آن‌ها که p -مقدار بزرگ‌تری داشته باشد، انتخاب می‌شود. البته باید از سطح α از قبل تعیین شده بزرگ‌تر باشد. چنانچه ماکسیمم p -مقدار کوچک‌تر از α باشد، همه مفصل‌ها رد خواهند شد و هیچ تابع مفصلی در این حالت انتخاب نمی‌شود. همچنین ممکن است در بین بهترین مفصل‌هایی که p -مقدار بزرگ‌تر از α داشته‌اند، مفصل‌هایی یافت شوند که p -مقدار نزدیک به هم داشته باشند. بنا بر این بهتر است فقط به این روش‌ها اکتفا نکنیم و روش‌های دیگری را نیز برای انتخاب مفصل مناسب در نظر بگیریم.

جدول ۴. تعداد رد فرضیه صفر تابع مفصل گامبل در ۱۰۰ بار تکرار (میانگین P -مقدارها) به وسیله روش‌های A_1 ([۱۴]) و A_2 ([۷]) و A_3

([۱۳]) در سطح $\alpha = 0.05$

A_3	A_2	A_1	مفصل واقعی	τ	n
۷۶(۰/۰۳۴)	۱۰(۰/۵۲۵)	۲(۰/۶۸۳)	نرمال	۰/۳	۳۰۰
۶۳(۰/۱۸۳)	۱۰(۰/۳۲۹)	۴۹(۰/۲۴۰)	فرانک		
۲۷(۰/۳۶۴)	۰(۰/۷۰۱)	۰(۰/۷۰۳)	گامبل		
۶۸(۰/۰۹۳)	۱(۰/۳۹۰)	۱۸(۰/۲۶۸)	جو		
۸۸(۰/۰۱۶)	۲۰(۰/۵۱۴)	۳(۰/۵۶۱)	نرمال	۰/۷	۵۰۰
۷۹(۰/۱۴۵)	۷۴(۰/۰۵۵)	۷۰(۰/۱۰۵)	فرانک		
۳۲(۰/۳۳۲)	۰(۰/۶۸۶)	۱(۰/۵۶۲)	گامبل		
۸۳۵(۰/۰۳۰)	۳۰(۰/۲۸۵)	۴۰(۰/۱۴۲)	جو		
۹۷(۰/۰۰۷)	۳۶(۰/۴۰۷)	۱۵(۰/۳۹۰)	نرمال	۰/۳	۵۰۰
۱۰۰(۰/۰۰۰)	۸۷(۰/۰۲۹)	۷۶(۰/۰۶۹)	فرانک		
۶۵(۰/۲۵۰)	۲(۰/۵۵۱)	۴(۰/۴۸۷)	گامبل		
۹۶(۰/۰۰۳)	۱۹(۰/۱۸۹)	۵۵(۰/۱۱۳)	جو		
۱۰۰(۰/۰۰۰)	۵۲(۰/۳۲۴)	۳۳(۰/۲۵۷)	نرمال	۰/۷	۵۰۰
۱۰۰(۰/۰۰۰)	۹۲(۰/۰۱۷)	۶۰(۰/۱۰۶)	فرانک		
۸۰(۰/۱۴۹)	۶(۰/۴۷۶)	۱۰(۰/۳۰۱)	گامبل		
۱۰۰(۰/۰۰۰)	۳۵(۰/۱۱۱)	۲۶(۰/۱۴۳)	جو		

^{۱۲} blanket tests

گرفته است که نسبت به آن‌ها دارای توانی کمتر می‌باشد. این آزمون در مقایسه با آزمون‌های پوششی دارای این مزیت است که دارای محاسباتی ساده‌تر است. هوآنگ و پروخوروف [۷] یک آزمون نیکویی برازش رتبه‌ای جدید برای توابع مفصل ارائه کردند که از تعیین پارامتری توزیع‌های حاشیه‌ای اجتناب می‌کند و شامل انتخاب استراتژی مانند برازش براساس تابع هسته یا انتخاب پهنای باند نیست. همچنین یک آمارهٔ آزمون با یک توزیع استاندارد است و نیازی به فرایند بوت‌استرپ برای محاسبه p -مقدار ندارد. در این مقاله آزمون ذکرشده را با دو آزمون دیگر ارائه‌شده در [۱۴] که جزو دسته آزمون‌هایی است که به استراتژی خاصی مانند انتخاب پهنای باند و برآورد چگالی هسته نیاز دارد و آزمون پیشنهاد شده در [۱۳] که براساس بوت‌استرپ پارامتری است، مورد مطالعه قرار می‌دهیم و به بررسی و مقایسهٔ این آزمون‌ها از نظر سهولت محاسبهٔ آماره‌ها و به‌دست آوردن p -مقادیر خواهیم پرداخت. مبنای ساخت این سه روش به‌اختصار به‌شرح زیر است:

گروه اول آزمون‌هایی برای آزمون خانواده‌های تابع مفصل پارامتری خاص هستند؛ مانند خانواده‌های توابع مفصل نرمال. گروه دوم آماره‌هایی هستند که می‌توانند برای آزمون کردن هر خانواده از توابع مفصل به کار روند. اما استفاده از آن‌ها نیاز به مواردی مانند تعیین پارامتر اختیاری، توابع هسته، توابع وزنی و پارامترهای هموارساز می‌یابد. در این زمینه می‌توان به [۳] اشاره کرد و گروهی دیگر از آزمون‌ها که به آزمون‌های پوششی معروف‌اند، برای همهٔ خانواده‌های توابع مفصل از هر ساختاری کاربرد دارند و نیاز به هیچ‌گونه انتخاب استراتژی به کار رفته در گروه ۲ نیست. از جمله می‌توان [۵] را نام برد. از آن‌جا که علاقه‌مند به یک فرایند کلی برای آزمون بدون هیچ محدودیتی در استفاده از آن هستیم، بنا بر این تمرکزمان را معطوف به آزمون‌های پوششی می‌کنیم و یک روش آزمون پوششی جدید را که توسط هوآنگ و پروخوروف [۷] ارائه‌شده است مورد بررسی قرار می‌دهیم. آزمون ارائه‌شده در [۷] مورد مقایسه با آزمون‌های پوششی ارائه‌شده در [۵] که شامل آزمون‌هایی براساس فرایند کندال، فرایند مفصل تجربی و تبدیل انتگرال احتمال هستند قرار

جدول ۵. تعداد رد فرضیهٔ صفر تابع مفصل جو در ۱۰۰ بار تکرار (میانگین P -مقادیرها) به‌وسیلهٔ روش‌های A_1 ([۱۴]) و A_2 ([۷]) و A_3 ([۱۳])

در سطح $\alpha = 0.05$

A_3	A_2	A_1	مفصل واقعی	τ	n
۴۸(۰/۰۳۰۲)	۱۸(۰/۵۱۹)	۵۰(۰/۱۲۱)	نرمال	۰/۳	۳۰۰
۷۶(۰/۱۸۱)	۴(۰/۵۹۰)	۲۴(۰/۳۰۹)	فرانک		
۹۳(۰/۰۰۳)	۱(۰/۷۱۲)	۱۰(۰/۴۰۲)	گامبل		
۴۳(۰/۲۳۵)	۰(۰/۷۳۱)	۰(۰/۸۲۴)	جو		
۵۲(۰/۲۱۱)	۳۰(۰/۱۳۳)	۷۰(۰/۰۹۸)	نرمال	۰/۷	
۸۳(۰/۰۹۴)	۴۶(۰/۱۴۵)	۵۸(۰/۲۶۸)	فرانک		
۶۹(۰/۱۲۷)	۲۱(۰/۵۴۱)	۱۳(۰/۳۴۶)	گامبل		
۵۶(۰/۱۱۲)	۲(۰/۶۴۶)	۰(۰/۴۶۳)	جو		
۱۰۰(۰/۰۰۰)	۲۶(۰/۴۰۷)	۹۱(۰/۰۱۳)	نرمال	۰/۳	۵۰۰
۱۰۰(۰/۰۰۰)	۸۲(۰/۰۲۸)	۶۶(۰/۱۵۳)	فرانک		
۸۵(۰/۰۷۲)	۷۶(۰/۲۷۸)	۳۶(۰/۲۲۱)	گامبل		
۶۲(۰/۱۰۸)	۴(۰/۵۱۳)	۱(۰/۴۳۹)	جو		
۹۸(۰/۰۰۱)	۴۵(۰/۳۸۴)	۱۰۰(۰/۰۱۳)	نرمال	۰/۷	
۱۰۰(۰/۰۰۰)	۱۰۰(۰/۰۰۱)	۱۰۰(۰/۰۰۹)	فرانک		
۹۳(۰/۰۱۴)	۱۰۰(۰/۰۰۰)	۹۱(۰/۰۱۹)	گامبل		
۷۰(۰/۰۹۷)	۸(۰/۴۸۵)	۱۰۰(۰/۰۱۳)	جو		

این سری داده‌ها را با (S_2) نشان می‌دهیم و به تعداد B بار از این داده‌ها نمونه‌گیری بوت‌استرپ انجام می‌دهیم (سری‌های S_{2j}^* ، $1 \leq j \leq B$). تحت فرض صفر سری تجربی تبدیل شده (S_1) و سری نمونه‌گیری شده (S_2) از تابع مفصل یکسان پیروی می‌کنند. از فاصله مربع \hat{Q} می‌توان استفاده نمود تا دو سری را با هم مقایسه کرد؛ زیرا تحت فرض صفر، این اندازه باید به مقداری کوچک بگراید. بر همین اساس، یک فاصله برآورد شده بین داده‌های تجربی تبدیل شده (S_1) و سری نمونه‌گیری شده اولیه (S_2) به دست می‌آوریم که با \hat{Q} نشان داده می‌شود. به همین ترتیب نیز یک فاصله برآورد شده بین S_{2j}^* و S_{1j}^* به دست می‌آوریم که با \hat{Q}_{2j}^* نشان می‌دهیم. به دلیل این که S_{2j}^* نمونه شبیه‌سازی شده از سری S_2 است فاصله \hat{Q}_{2j}^* مقداری کوچک می‌باشد و چنانچه فاصله \hat{Q} مقداری کوچک‌تر از \hat{Q}_{2j}^* داشته باشد می‌توان نتیجه گرفت که فرض صفر برقرار خواهد بود و دو سری S_1 و S_2 از تابع مفصل یکسان پیروی می‌کنند. پانچنکو [۱۴] بر مبنای فرایند مقایسه بین مقادیر \hat{Q} و \hat{Q}_{2j}^* -p مقدار را به صورت زیر به دست آورد:

$$\hat{P} = \frac{\sum_{i=1}^B I(\hat{Q} \geq \hat{Q}_i^*) + 1}{B + 1},$$

که در آن I تابع نشانگر است. فرض صفر رد می‌شود هرگاه $\hat{P} \leq \alpha$ باشد، که α یک اندازه آزمون است.

A₂: آزمون نیکویی برازش برای تابع مفصل براساس ماتریس اطلاع

هوآنگ و پروخوروف [۷] یک آزمون نیکویی برازش رتبه‌ای جدید برای توابع مفصل ارائه کردند. فرض کنیم θ بردار پارامتر وابستگی یک تابع مفصل باشد. یک تابع مفصل به درستی تعیین شده است اگر مقدار θ . وجود داشته باشد که به ازای آن $H(x_1, \dots, x_d) = C_\theta(F_1(x_1), \dots, F_d(x_d))$ برقرار باشد. در واقع، بنای ساخت آزمون نیکویی برازش برای تابع مفصل براساس این تعریف قرار داده شده است. فرض کنید $\mathbb{H}(\theta)$ امید ماتریس هسین از $\ln c_\theta$ و $\mathbb{C}(\theta)$ امید ضرب خارجی بردار امتیاز باشند که به صورت زیر تعریف می‌شوند:

$$\mathbb{H}(\theta) = E \nabla_\theta^\top \ln c_\theta(F_1(x_1), \dots, F_d(x_d))$$

$$\mathbb{C}(\theta) = E \nabla_\theta \ln c_\theta(F_1(x_1), \dots, F_d(x_d)) \times \nabla_\theta' \ln c_\theta(F_1(x_1), \dots, F_d(x_d))$$

A₁: روش ارائه شده براساس ضرب داخلی بین دو بردار

به عنوان اندازه‌ای از فاصله بین آن‌ها

پانچنکو [۱۴] برای ساخت آماره آزمون نیکویی برازش تابع مفصل با استفاده از ویژگی‌های صورت دوخطی که اولین بار در [۲] معرفی شده، یک فاصله مربع Q بین تابع‌های انتگرال پذیر f_1 و f_2 به صورت زیر در نظر گرفت:

$$Q = \langle (f_1 - f_2) | K_d | (f_1 - f_2) \rangle$$

که در آن d بیان‌کننده تعداد متغیرها است، $s_1, s_2 \in R^d$ ، $K_d(\cdot, \cdot)$ یک هسته متقارن معین مثبت است مانند تابع هسته گوسی

$$K_d(s_1, s_2) = e^{-\|s_1 - s_2\|^2 / (2dh^2)},$$

که در آن $\|\cdot\|$ نرم اقلیدسی در R^d تعریف می‌شود و $h > 0$ یک پارامتر هموارکننده^{۱۳} (پهنای باند) است. Q صفر می‌شود تنها وقتی که $f_1(\cdot)$ و $f_2(\cdot)$ مساوی باشند. با توجه به ویژگی‌های ضرب داخلی، Q می‌تواند به صورت زیر تجزیه شود:

$$Q = Q_{11} - 2Q_{12} + Q_{22},$$

که در آن $Q_{ij} = \langle f_i | K_d | f_j \rangle$ است. هر جمله در تجزیه بالا به طور سازگار با استفاده از آماره V به صورت زیر برآورد می‌شود [۱۴].

$$\hat{Q}_{ij} = \frac{1}{n^2} \sum_{t_1=1}^n \sum_{t_2=1}^n K_d(S_{i,t_1}^t, S_{j,t_2}^t),$$

که در آن S^t مقدار مشاهده بردار تصادفی S در نقطه t است. در نهایت، برآورد Q به صورت زیر به دست می‌آید:

$$\hat{Q} = \hat{Q}_{11} - 2\hat{Q}_{12} + \hat{Q}_{22}.$$

رابطه به دست آمده در فرایند آزمون مورد استفاده قرار می‌گیرد. با تبدیل داده‌های واقعی به وسیله رابطه

$$F_{n,k}(x_k) = n^{-1} \sum_{i=1}^n I(X_{i,k} \leq x_k), k = 1, \dots, d$$

توزیع‌های حاشیه‌ای تجربی را به دست می‌آوریم و توأم شده این داده‌ها می‌تواند به عنوان تابع مفصل واقعی تلقی شود. این سری داده‌ها را با (S_1) نشان می‌دهیم. برآورد پارامترهای تابع مفصل را به دست می‌آوریم و با استفاده از برآوردهای به دست آمده داده‌هایی از تابع مفصل تحت فرض صفر شبیه‌سازی می‌شود.

^{۱۳} smoothing parameter

که در آن ∇_{θ} مشتق نسبت به θ تعریف می‌شود و امیدها نسبت به توزیع دقیق \mathbb{H} است و c_{θ} تابع چگالی مفصل است. θ بردار پارامتر وابستگی تابع مفصل و دارای بعد p است. ایده اصلی این آزمون در ارتباط با آزمون ارائه شده در [۱۶] است که در اصل بیان می‌کند به شرط تعیین درست تابع مفصل داریم:

$$-\mathbb{H}(\theta) = \mathbb{C}(\theta)$$

وقتی این تساوی برقرار نشود، نتیجه می‌شود که مدل به درستی تعیین نشده است. صورت مسئله این آزمون به صورت

$$\begin{cases} H_0 : \mathbb{H}(\theta) + \mathbb{C}(\theta) = 0 \\ H_1 : \mathbb{H}(\theta) + \mathbb{C}(\theta) \neq 0 \end{cases}$$

است. هوآنگ و پروخوروف [۷] برای ساخت آماره آزمون از مؤلفه‌های ماتریس $(\hat{\mathbb{H}}_t(\theta) + \hat{\mathbb{C}}_t(\theta))$ به طریق زیر استفاده کردند:

$$d_t(\theta) = \text{vech}(\mathbb{H}_t(\theta) + \mathbb{C}_t(\theta)),$$

$$\hat{d}_t(\theta) = \text{vech}(\hat{\mathbb{H}}_t(\theta) + \hat{\mathbb{C}}_t(\theta)),$$

که vech مؤلفه‌های مثلث پایینی ماتریس را در یک بردار قائم^{۱۴} قرار می‌دهد. باید توجه داشت که $d_t(\theta)$ به حاشیه‌های ناشناخته وابسته است، در حالی که در $\hat{d}_t(\theta)$ از حاشیه‌های تجربی $(\hat{F}_{n,k})$ ، $k = 1, \dots, d$ استفاده می‌شود. میانگین $\hat{d}_t(\theta)$ به صورت

$$\bar{D}_{\theta} \equiv \bar{D}(\theta) \equiv n^{-1} \sum_{t=1}^n \hat{d}_t(\theta)$$

تعریف می‌شوند. فرض کنید $\bar{D}_{\hat{\theta}} = \bar{D}(\hat{\theta})$ و $D_{\theta} = E d_t(\theta)$ که $\hat{\theta}$ بردار پارامتر θ است. تحت تشخیص درست تابع مفصل خواهیم داشت:

$$D_{\theta} \equiv E d_t(\theta) = 0.$$

با استفاده از $\bar{D}_{\hat{\theta}}$ هوآنگ و پروخوروف [۷] آماره آزمون نیکویی برازش برای تابع مفصل را به صورت زیر ارائه کردند:

$$j = T \bar{D}_{\hat{\theta}}' V_{\theta}^{-1} \bar{D}_{\hat{\theta}},$$

که به طور مجانبی دارای توزیع $\chi_{p(p+1)/2}^2$ است و در آن V_{θ}

به صورت

$$\begin{aligned} V_{\theta} = & E \left\{ d_t(\theta) + \nabla D_{\theta} B^{-1} \left[\nabla_{\theta} \ln c(F_{1t}, F_{2t}, \dots, F_{dt}; \theta) \right. \right. \\ & \left. \left. + \sum_{k=1}^d W_k(F_{kt}) \right] + \sum_{k=1}^d M_k(F_{kt}) \right\} \\ & \times \left\{ d_t(\theta) + \nabla D_{\theta} B^{-1} \left[\nabla_{\theta} \ln c(F_{1t}, F_{2t}, \dots, F_{dt}; \theta) \right. \right. \\ & \left. \left. + \sum_{k=1}^d W_k(F_{kt}) \right] + \sum_{k=1}^d M_k(F_{kt}) \right\} \end{aligned}$$

تعریف می‌شود که در آن برای $k = 1, 2, \dots, d$ داریم:

$$\begin{aligned} W_k(F_{kt}) = & \int_0^1 \int_0^1 \dots \int_0^1 [I\{F_{kt} \leq u_k\} - u_k] \\ & \times \nabla_{\theta, u_k}^{\top} \ln c(u_1, u_2, \dots, u_d; \theta) \\ & \times c(u_1, u_2, \dots, u_d; \theta) du_1 du_2 \dots du_d \\ M_k(F_{kt}) = & \int_0^1 \int_0^1 \dots \int_0^1 [I\{F_{kt} \leq u_k\} - u_k] \\ & \times \nabla_{u_k} \text{vech}[\nabla_{\theta}^{\top} \ln c(u_1, u_2, \dots, u_d; \theta) \\ & + \nabla_{\theta} \ln c(u_1, u_2, \dots, u_d; \theta) \\ & \times \nabla_{\theta}^{\top} \ln c(u_1, u_2, \dots, u_d; \theta)] \\ & \times c(u_1, u_2, \dots, u_d; \theta) du_1 du_2 \dots du_d \end{aligned}$$

A۳: آزمون نیکویی برازش براساس فاصله ماهالانویس بین لگاریتم درست‌نمایی واقعی و شبیه‌سازی

فرایند ارائه شده در [۱۳] برای تابع مفصل به شرح زیر است:

در ابتدا توابع مفصل C_1, C_2, \dots, C_m را به داده‌های اصلی برازش می‌دهیم و برآورد پارامترها را محاسبه می‌کنیم و لگاریتم درست‌نمایی هر کدام از آن‌ها را با λ_i ، $i = 1, \dots, m$ نشان می‌دهیم و برای داده‌های اصلی (مشاهده شده)، بردار $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_m)'$ را به دست می‌آوریم. در مرحله دوم از هر تابع مفصل برازش داده شده C_k ، $k = 1, \dots, m$ نمونه‌ای شبیه‌سازی می‌کنیم و خانواده‌های مفصل C_1, C_2, \dots, C_m را به داده‌های شبیه‌سازی شده نیز برازش می‌دهیم و لگاریتم درست‌نمایی هر یک را محاسبه می‌کنیم و بردار لگاریتم درست‌نمایی را در این مرحله با $\lambda_k = (\lambda_{1k}, \lambda_{2k}, \dots, \lambda_{mk})'$ نشان می‌دهیم. مرحله ۲ را B بار تکرار می‌کنیم و در نتیجه برای هر نمونه شبیه‌سازی شده در مرحله ۲ بردار لگاریتم‌های درست‌نمایی $\lambda_k = (\lambda_{1k_i}, \lambda_{2k_i}, \dots, \lambda_{mk_i})'$ به ازای $i = 1, \dots, B$ لگاریتم درست‌نمایی ارزیابی شده در

^{۱۴}vertical vectorization

توابع مفصل مختلف شبیه‌سازی می‌شود، در تولید اعداد تصادفی از این توابع مفصل، پارامترها را طوری تعیین می‌کنیم که ضریب همبستگی τ کندال، به‌عنوان معیار هماهنگی توابع مفصل دارای مقدار یکسانی باشد. آماره‌های روش‌های مختلف نیکویی برازش برای توابع مفصل تحت فرضیه H_0 محاسبه شده و p -مقدارها به‌دست آمده‌اند. فرایند آزمون با تکرار $k = 100$ انجام شده و درصد موارد رد فرضیه صفر در ۱۰۰ بار تکرار و میانگین p -مقدارها به‌ازای n و τ مختلف در جداول ۲ تا ۵ ملاحظه می‌شود. بوت‌استرپ برای دو روش A_1 و A_2 ، صد بار تکرار شده و پهنای باندها در روش A_1 ، برابر با ۰/۰۵ در نظر گرفته شده است. آزمون‌ها در سطح $\alpha = 0.05$ انجام شده‌اند. همه محاسبات توسط برنامه‌ای در نرم‌افزار R و با استفاده از بسته نرم‌افزاری Copula نسخه ۰/۹۹۹ انجام شده است.

۱.۲.۴ آزمون فرضیه تابع مفصل گوسی

نتایج حاصل از آزمون‌ها تحت فرضیه صفر نرمال در مقابل این که داده‌ها از چهار تابع مفصل مختلف نرمال، فرانک، گامبل و جو تولید شده‌اند، در جدول ۲ آمده است. نتایج جدول ۲ نشان می‌دهد که روش A_1 در خصوص عدم رد تابع مفصل درست، به‌خوبی عمل می‌کند (درصد ردها وقتی فرضیه صفر درست است بسیار کم شده‌اند)، متوسط p -مقدارها نیز نشان می‌دهد که آزمون A_1 در این حالت، خوب عمل می‌کند. البته نتایج نشان می‌دهد که برای تعداد زیاد نمونه تصادفی، آزمون A_1 می‌تواند نتایج بهتری را به‌دست دهد و درصد موارد رد فرضیه صفر در حالتی که تابع مفصل نادرست است، به‌جز در مورد رد تابع مفصل نرمال وقتی تابع مفصل واقعی فرانک است، به‌ازای $\tau = 0.7$ افزایش یافته است. همین‌طور با افزایش τ نیز درصد ردها در روش A_1 افزایش یافته است. روش A_2 در مقایسه با دو روش دیگر، درصد ردهای کمتری نتیجه داده است. به‌خصوص وقتی اندازه نمونه ۳۰۰ است درصد موارد رد کمتر می‌باشد. روش A_2 در مقایسه با روش A_1 تنها در حالتی که اندازه نمونه ۵۰۰ و تابع مفصل واقعی فرانک است نتیجه بهتری داده است و فرضیه صفر مفصل نرمال را به تعداد بیشتری رد کرده است. نتایج جدول ۲ نشان می‌دهد که روش A_3 در خصوص رد توابع مفصل نادرست، بسیار خوب عمل کرده است و مخصوصاً در مورد توابع مفصل واقعی گامبل و جو درصد موارد رد بسیار زیاد است. اما این روش در مورد عدم رد تابع مفصل درست (وقتی که تابع مفصل واقعی نرمال است)

داده‌های اصلی را با لگاریتم درست‌نمایی ارزیابی شده در داده‌های شبیه‌سازی شده مقایسه می‌کنیم. انتظار می‌رود بهترین مدل، مدلی باشد که لگاریتم درست‌نمایی آن به لگاریتم درست‌نمایی مشاهدات متناظر نزدیک‌تر باشد. بنا بر این یک آزمون نیکویی برازش بر مبنای اندازه فاصله بین لگاریتم درست‌نمایی‌های مشاهده شده و مقادیر شبیه‌سازی شده به‌دست می‌آید. توزیع توأم لگاریتم‌های درست‌نمایی $\lambda_k = (\lambda_{1k_i}, \lambda_{2k_i}, \dots, \lambda_{mk_i})'$ ، $i = 1, \dots, B$ براساس قضیه حدی مرکزی تقریباً نرمال چندمتغیره است. بنا بر این فاصله مربع ماهالانویس برای انجام دادن مقایسه‌های بالا، مناسب می‌باشد. میانگین لگاریتم‌های درست‌نمایی نسبت به B بار تکرار را محاسبه می‌کنیم و با $\bar{\lambda}_k = (\bar{\lambda}_{k_1}, \bar{\lambda}_{k_2}, \dots, \bar{\lambda}_{k_m})'$ نشان می‌دهیم. فاصله ماهالانویس بین بردار لگاریتم درست‌نمایی λ در داده‌های اصلی و بردار میانگین لگاریتم درست‌نمایی، $\bar{\lambda}_k$ در داده‌های شبیه‌سازی شده از k امین تابع مفصل به‌صورت

$$D_k^* = (\lambda - \bar{\lambda}_k)S^{-1}(\lambda - \bar{\lambda}_k)', k = 1, \dots, m$$

به‌دست می‌آید، که در آن S ماتریس کوواریانس نمونه لگاریتم‌های درست‌نمایی است. براساس فرض نرمال بودن لگاریتم‌های درست‌نمایی، ماردیا و همکاران [۱۰] نشان دادند، تحت فرض صفری که k امین توزیع مفصل صحیح است، آماره mD_k^* دارای توزیع $F_{m, B-1}$ است.

۲.۴ فرایند آزمون

الگوریتم آزمون ارائه شده را بر روی چهار تابع مفصل برای اندازه‌های نمونه و سطوح وابستگی مختلف به کار می‌بریم و با استفاده از فرایند شبیه‌سازی، حالت‌های زیر را مورد آزمون قرار می‌دهیم:

- H_0 : توابع مفصل (چهار انتخاب: نرمال، فرانک، گامبل و جو)
- H_1 : توابع مفصل (چهار انتخاب: نرمال، فرانک، گامبل و جو)
- τ کندال (دو انتخاب: $\{0.7, 0.3\}$)
- اندازه نمونه (دو انتخاب: $\{500, 300\}$)

در همه این ۶۴ حالت، یک نمونه به اندازه n و با بعد $d = 2$ از هر یک از توابع مفصل H_1 تولید می‌کنیم. به‌دلیل این که از

برای $n = 300$ و $\tau = 0.3$ درصد موارد رد بیشتری داشته است. روش A_3 نیز همانند نتایج دو فرضیه صفر نرمال و فرانک درصد موارد رد زیادی به خصوص برای مقادیر زیاد n و τ نتیجه داده است.

۴.۲.۴ آزمون فرضیه تابع مفصل جو

نتایج جدول ۵ نشان می‌دهد که برای مقادیر بزرگ n و τ هر سه آزمون در مورد رد تابع مفصل نادرست به خوبی عمل کرده‌اند. برای اندازه نمونه ۵۰۰ روش A_2 در خصوص رد توابع مفصل واقعی فرانک و گامبل بهتر از روش A_1 عمل کرده و روش A_1 وقتی تابع مفصل واقعی جو است به خوبی عمل کرده است.

۵ مثال کاربردی

مقایسه میان خانواده توابع مفصل مختلف را می‌توان برای ارزیابی نیکویی برازش مدل، مورد استفاده قرار داد. در این بخش، سه روش آزمون نیکویی برازش ذکر شده را برای انتخاب تابع مفصل مناسب با استفاده از داده‌های سری‌های زمانی دو شاخص بورس، فراورده‌های نفتی و محصولات شیمیایی به کار می‌بریم. داده‌ها روزانه و از تاریخ ۱۶ آذر ۱۳۹۰ تا ۱۵ آذر ۱۳۹۱ و به طول ۲۹۳ می‌باشند. جدول ۶، p -مقدارهای به دست آمده را به ازای $\alpha = 0.05$ ارائه می‌کند. نتایج حاصل در جدول ۶ نشان می‌دهد مفصل گوسی به دلیل ویژگی متقارن بودن برای تشخیص بین این سری‌های مالی ناکارآمد می‌باشد، اما تابع مفصل گامبل می‌تواند به طور مناسب‌تر وابستگی بین سری‌ها را شرح دهد. همچنین در بین سه روش ارائه شده، روش‌های A_1 و A_3 نسبت به روش A_2 مناسب بودن تابع مفصل گامبل را بهتر نشان داده‌اند.

۶ نتیجه گیری

در این مقاله با استفاده از شبیه‌سازی، فرایند سه روش آزمون نیکویی برازش پیشنهاد شده در [۱۴]، [۷] و [۱۳] بر روی چهار تابع مفصل مختلف نرمال، فرانک، گامبل و جو به ازای مقادیر همبستگی و اندازه‌های نمونه مختلف مورد آزمون قرار گرفتند. روش پیشنهاد شده در [۱۴] در عدم رد تابع مفصل درست (تشخیص درست وقتی است که تابع مفصل صحیح همان تابع مفصل مورد آزمون باشد) دارای عملکرد مناسب است و میانگین p -مقدارها نشان از عملکرد خوب آزمون در این زمینه است و

به خوبی عمل نکرده است. هر سه روش وقتی تابع مفصل واقعی جو می‌باشد، به خوبی عمل کرده‌اند. در هر دو روش A_2 و A_3 نیز با افزایش n و τ تعداد رد فرضیه صفر افزایش یافته است. شکل ۱ رفتار سه روش آزمون ارائه شده را به ازای n ها و τ های مختلف و وقتی که تابع مفصل واقعی فرانک و گامبل است نشان می‌دهد. به طور مشابه می‌توان نتایج جدول ۲ را با توجه به نمودار مشاهده کرد. روش A_2 در مقایسه با دو روش دیگر توان کمتری دارد و همچنین هر سه روش با افزایش n و τ ، نتایج بهتری در مورد رد تابع مفصل نادرست می‌دهند.

۲.۲.۴ آزمون فرضیه تابع مفصل فرانک

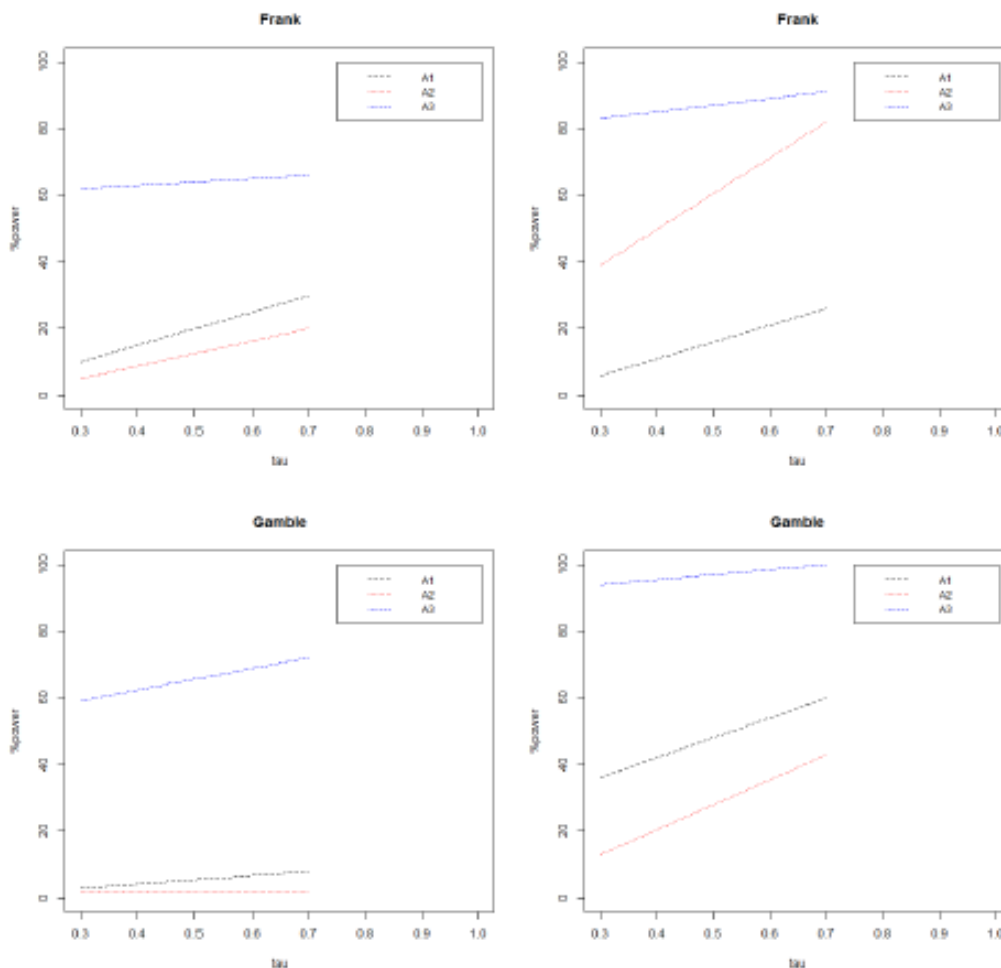
همان‌طور که نتایج جدول ۳ برای آزمون فرض صفر تابع مفصل فرانک نشان می‌دهد، روش A_1 در عدم رد تابع مفصل فرانک وقتی تابع مفصل واقعی فرانک است، به خوبی عمل کرده است. در روش A_1 با افزایش τ به جز در اندازه نمونه ۳۰۰ وقتی تابع مفصل واقعی گامبل است، درصد موارد رد افزایش یافته است. روش A_2 نسبت به حالتی که فرضیه صفر تابع مفصل نرمال بود، درصد موارد رد کمتری نتیجه داده است و در مقایسه با دو آزمون دیگر در مورد رد تابع مفصل نادرست، ضعیف‌تر عمل کرده و فقط به ازای $\tau = 0.7$ وقتی تابع مفصل واقعی گامبل است، نسبت به آزمون A_1 بهتر عمل کرده است. در روش A_2 همان‌طور که در جدول ۳ مشاهده می‌شود، با افزایش τ در حالتی که اندازه نمونه ۵۰۰ است و تابع مفصل واقعی نرمال است، درصد موارد رد افزایش یافته است. روش A_3 نیز درصد موارد رد زیادی را نتیجه داده و در مورد عدم رد تابع مفصل فرانک وقتی تابع مفصل واقعی فرانک است نسبت به دو آزمون دیگر ضعیف‌تر عمل کرده است. هر سه روش در مورد رد تابع مفصل فرانک وقتی تابع مفصل واقعی گامبل و جو است به خوبی عمل کرده‌اند. نتایج جدول ۳ نشان می‌دهد که با افزایش اندازه نمونه، هر سه روش نتایج بهتری در مورد رد تابع مفصل نادرست می‌دهند.

۳.۲.۴ آزمون فرضیه تابع مفصل گامبل

نتایج جدول ۴ نشان می‌دهد که روش A_2 در مقایسه با دو روش دیگر در مورد عدم رد تابع مفصل گامبل وقتی تابع مفصل واقعی گامبل است بهتر عمل کرده است و همچنین در مقایسه با روش A_1 در خصوص رد توابع مفصل نرمال و فرانک نیز عملکرد بهتری داشته است و به جز وقتی تابع مفصل واقعی فرانک است

شده در [۱۳] سادگی در محاسبه آماره آزمون آن است. با توجه به نتایج حاصل از شبیه سازی، تعداد دفعات انتخاب درست تابع مفصل توسط این سه روش به شدت به تعداد نمونه ها و میزان وابستگی بین متغیرها ارتباط دارد است. با مقایسه نتایج جداول ۲ تا ۵ با افزایش n و τ نتایج حاصل از انتخاب نادرست تابع مفصل بهبود می یابد و درصد موارد رد تابع مفصل نادرست افزایش می یابد.

برعکس، روش [۱۳] در عدم رد تابع مفصل درست عملکرد خوبی ندارد. همچنین تعداد رد تابع مفصل نادرست توسط روش [۱۳] بسیار زیاد است، ولی روش [۱۴] در این خصوص، عملکردی قوی نداشته است. بنا بر این دو روش تقریباً عکس همدیگر عمل می کنند. یک تفاوت دیگر بین این دو روش، مدت زمان لازم برای اجرای محاسبات است. روش [۱۳] در شرایط یکسان با روش [۱۴] بیش از چهار برابر زمان نیاز دارد که نقطه ضعف برای این روش به حساب می آید. اما یک مزیت روش پیشنهاد



شکل ۱. توان در مقابل τ -کنندال برای سه روش آزمون نیکویی برازش A_1 ، A_2 و A_3 برای تابع مفصل فرانک و گامبل تحت فرضیه صفر مفصل نرمال. نمودارهای سمت چپ به ازای $n = 300$ و نمودارهای سمت راست به ازای $n = 500$ می باشند.

جدول ۶. نتایج روش های آزمون نیکویی برازش برای چهار تابع مفصل برای شاخص های بورس

مقدار P	نرمال	فرانک	گامبل	جو
A_1	۰/۰۱۵	۰/۶۲۷	۰/۷۹۲	۰/۵۴۲
A_2	۰/۰۱۹	۰/۴۹۳	۰/۵۰۲	۰/۴۹۸
A_3	۰/۰۰۹	۰/۵۱۲	۰/۶۸۳	۰/۵۳۹

مراجع

- [1] Berg, D. and Bakken, H. (2007). A goodness-of-fit test for copulae based on the conditional probability integral transform. *Preprint Series. Statistical Research Report* <http://urn.nb.no/URN:NBN:no-23420>.
- [2] Diks, C., Van Zwet, W. R., Takens, F., and DeGoede, J. (1996). Detecting differences between delay vector distributions. *Physical Review E*, **53**, 2169–2176.
- [3] Fermanian, J.-D. (2005). Goodness of Fit Tests for Copulas. *Journal of Multivariate Analysis*, **95**, 119-152.
- [4] Genest, C., Quessy, J.-F., and R'emillard, B. (2006). Goodness-of-fit procedures for copula models based on the probability integral transform. *Scandinavian Journal of Statistics*, **33**.
- [5] Genest, C., R'emillard, B., and Beaudoin, D. (2009). Goodness-of-fit tests for copulas: a review and a power study. *Insurance: Mathematics and Economics*, **44**, 199–213.
- [6] Gumbel, E. J. (1960). Bivariate exponential distributions. *Journal of the American Statistical Association*, **55(292)**, 698-707.
- [7] Huang, W. and Prokhorov, A. (2011). A Goodness- of- fit Test for Copulas. *Econometric Reviews*, **33(7)**, 751-771.
- [8] Joe, H. (1993). Parametric families of multivariate distributions with given margins. *Journal of Multivariate Analysis*, **46**, 262-282.
- [9] Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika*, **30**, 81–93.
- [10] Mardia, K., Kent, J., and Bibby, J. (1979). *Multivariate Analysis*. London: Academic Press.
- [11] Nelsen, R. B. (1991). Copulas and association. In: Dall'Aglio G, Kotz S, Salinetti G (eds) *Advances in Probability Distributions with Given Marginals*. Kluwer, Dordrecht, pp 51-74.
- [12] Nelsen, R. B. (2006). *An Introduction to Copulas*. New York: Springer Series in Statistics. second ed. Springer.
- [13] Nikoloulopoulos, A. K. and Karlis, D. (2008). Copula model evaluation based on parametric bootstrap. *Computational Statistics and Data Analysis*, **52**, 3342-3353.
- [14] Panchenko, V. (2005). Goodness-of-fit test for copulas. *Physica A: Statistical Mechanics and its Applications*, **355(1)**, 176–182.
- [15] Sklar, A. (1959). Fonctions de repartition a n dimensions et leurs marges. *Publications de l'Institut de statistique de l'Université de Paris*, **8**, 229-231.
- [16] White, H. (1982). Maximum likelihood estimation in misspecified models. *Econometrica*, **50**, 1–26.