**مجله علوم آماری، پاییز و زمستان ۱۳۹۹**
**جلد ۱۴، شماره ۲، ص ۳۸۹ – ۴۰۸**
**DOI:** 10.29252/jss.14.2.389

## معیار نادرستی بر مبنای تابع مفصل بقاء

سیده تکتم حسینی و جعفر احمدی

گروه آمار، دانشکده علوم ریاضی، دانشگاه فردوسی مشهد

تاریخ دریافت: ۱۳۹۸/۰۲/۲۱ تاریخ پذیرش: ۱۳۹۹/۰۱/۱۵

**چکیده:** در این مقاله، با استفاده از ایده اندازه نادرستی در نظریه اطلاع، معیارهای نادرستی باقیمانده و گذشته در حالت دو متغیره به ترتیب بر مبنای تابع مفصل بقاء و مفصل تعریف شده است. تحت فرض تقارن شعاعی، برابری این دو معیار نشان داده شده است. همچنین با استفاده از تساوی بین این دو معیار، مدل‌های متقارن شعاعی مشخص‌سازی شده‌اند. تحت فرض برقراری مدل نرخ خطر متناسب برای توزیع‌های حاشیه‌ای، کران برای معیار معرفی شده، به دست آمده است. همچنین با فرض تناسب بین نادرستی معرفی شده و آنتروپی متناظرش، مدل نرخ خطر متناسب در حالت دو متغیره مشخص‌سازی شده است. بعلاوه، از ترتیب مربعی بالا برای به دست آوردن نابرابری‌هایی استفاده شده است. برای تشریح بیش‌تر نتایج به دست آمده، مثال به همراه روش‌های شبیه‌سازی ارائه شده است.
**واژه‌های کلیدی:** آنتروپی، اندازه اطلاع، اندازه نادرستی، ترتیب‌های مربعی، تقارن شعاعی، مفصل، نرخ خطر (معکوس) متناسب.

## ۱ مقدمه

اندازه نادرستی [1] یکی از معیارهای مهم در نظریه اطلاع به شمار می‌آید که توسط کریج (۱۹۶۱) برای دو متغیر تصادفی گسسته $X_۱$ و $Y_۱$ به ترتیب با تابع‌های جرم احتمال $p_۱(\cdot)$ و $p_۲(\cdot)$، تکیه‌گاه‌های شمارای

آدرس الکترونیکی نویسنده مسئول مقاله: سیده تکتم حسینی، hosseini_toktam@mail.um.ac.ir
کد موضوع‌بندی ریاضی (۲۰۱۰): 94A17, 62B10, 60E15.

[1] Inaccuracy measure

$S_{X_1}$ و $S_{Y_1}$ به‌صورت $I(X_1, Y_1) = -\sum_{x\in S_{X_1}\cap S_{Y_1}} p_1(x)\log p_2(x)$ تعریف شده است. نات (۱۹۶۸) این معیار را برای دو متغیر تصادفی مطلقا پیوسته $X_1$ و $Y_1$ به‌صورت

$$I(X_1, Y_1) = -\int_{x\in S_{X_1}\cap S_{Y_1}} f_1(x)\log g_1(x)dx, \tag{۱}$$

تعریف کرد، که در آن $f_1(\cdot)$ و $g_1(\cdot)$ به‌ترتیب تابع‌های چگالی $X_1$ و $Y_1$ با تکیه‌گاه‌های $S_{X_1}$ و $S_{Y_1}$ هستند. معیار نادرستی، خطایی را که آزمایش‌گر در برآورد احتمال رخداد پیشامدها در اثر استفاده نادرست از توزیع $G_1$ به جای توزیع درست $F_1$ و همچنین در اثر وجود اطلاعات نادرست در مشاهدات مرتکب می‌شود، نشان می‌دهد. شکرانی و خراشادی‌زاده (۱۳۹۷) اندازه نادرستی را براساس تابع چندک برای متغیر تصادفی طول عمر گذشته، مطالعه و نتایج مشخص‌سازی به دست آوردند. اگر در رابطه (۱) متغیر تصادفی $Y_1$ هم‌توزیع با $X_1$ باشد، آن‌گاه $I(X_1, Y_1) = I(X_1, X_1) = H(X_1)$ همان آنتروپی دیفرانسیلی شانون برای متغیر تصادفی $X_1$ است (شانون، ۱۹۴۸). این معیار می‌تواند میزان خطای مرتکب شده در برآورد احتمال رخداد یک پیشامد در اثر وجود اطلاعات نادرست در مشاهدات را نشان دهد. همچنین می‌تواند به عنوان معیاری برای تعیین پراکندگی و تشخیص مدل استفاده شود (زمان‌زاده و ارقامی، ۱۳۸۷؛ کاور و توماس، ۲۰۰۶). فرض کنید $\boldsymbol{X} = (X_1, X_2)$ و $\boldsymbol{Y} = (Y_1, Y_2)$ دو بردار تصادفی پیوسته به ترتیب با توابع چگالی احتمال توأم $f$ و $g$ و تکیه‌گاه‌های توأم $S_{\boldsymbol{X}} = (\ell_{X_1}, r_{X_1})\times(\ell_{X_2}, r_{X_2})$ و $S_{\boldsymbol{Y}} = (\ell_{Y_1}, r_{Y_1})\times(\ell_{Y_2}, r_{Y_2})$ باشند طوری‌که $S_{\boldsymbol{X}} = S_{\boldsymbol{Y}}$ و $-\infty \le \ell_{X_i} < r_{X_i} \le +\infty$ برای $i = 1, 2$. در این‌صورت معیار نادرستی در حالت دو متغیره به‌صورت

$$BI(\boldsymbol{X}, \boldsymbol{Y}) = -\int_{\ell_{X_1}}^{r_{X_1}}\int_{\ell_{X_2}}^{r_{X_2}} f(x,y)\log g(x,y)dydx, \tag{۲}$$

است. جاش و کاندو (۲۰۱۹) نادرستی مانده تجمعی دومتغیره را به‌صورت

$$BCRI(\boldsymbol{X}, \boldsymbol{Y}) = -\int_{\ell_{X_1}}^{r_{X_1}}\int_{\ell_{X_2}}^{r_{X_2}} \overline{F}(x,y)\log \overline{G}(x,y)dydx, \tag{۳}$$

تعریف کردند، که در آن $\overline{F}$ و $\overline{G}$ به ترتیب تابع‌های بقاء توأم $\boldsymbol{X}$ و $\boldsymbol{Y}$ هستند. معیارهای (۲) و (۳) در حالتی که $\boldsymbol{X}$ و $\boldsymbol{Y}$ دارای توزیع‌های یکسانی باشند به ترتیب به آنتروپی دو متغیره و آنتروپی مانده تجمعی دو متغیره تبدیل می‌شود. در سال‌های اخیر استفاده از توابع مفصل در نظریه اطلاع برای به دست آوردن

نتایج جدید، مورد توجه پژوهشگران قرار گرفته است. ولی براساس اطلاعات نویسندگان کاری در موضوع اندازه نادرستی برمبنای نظریه مفصل انجام نشده است. در این مقاله با استفاده از مفاهیم نظریه مفصل معیارهای نادرستی در حالت دو متغیره (۲) و (۳) در نظریه اطلاع گسترش داده می‌شود. بدین منظور در بخش ۲ مفاهیم پایه‌ای توابع مفصل یادآوری می‌شوند. بعد یک معیار نادرستی جدید براساس تابع مفصل بقاء تحت عنوان نادرستی مانده تجمعی مفصلی [1] معرفی می‌شود. در بخش ۳ مفهوم تقارن در حالت دو متغیره بیان می‌شود. سپس نشان داده می‌شود که اگر بردارهای تصادفی متقارن شعاعی [2] باشند، نادرستی مانده تجمعی مفصلی با معیار نادرستی گذشته تجمعی مفصلی [3] برابر می‌شود. همچنین شرایطی بررسی می‌شود که از برابری دو معیار نادرستی مانده و گذشته تجمعی مفصلی بتوان تقارن شعاعی را نتیجه گرفت. در بخش ۴ تحت فرض برقراری مدل نرخ خطر متناسب، کران‌هایی برای نادرستی مانده مفصلی به دست آورده می‌شود. همچنین در کلاسی از توزیع‌های دومتغیره با فرض تناسب بین معیار نادرستی معرفی شده براساس تابع مفصل بقاء و آنتروپی متناظرش نشان داده می‌شود که مدل نرخ خطر متناسب در حالت دو متغیره برقرار می‌شود. در بخش ۵ با استفاده از ترتیب مربعی بالا [4] نابرابری برای معیار معرفی شده، ارائه می‌شود. بعلاوه نشان داده می‌شود با فرض این نوع ترتیب تصادفی بین بردارهای تصادفی، نامساوی مثلث برای این معیار برقرار است. به منظور بررسی بیش‌تر برخی از نتایجی که در این مقاله مطرح می‌شود، در مثال‌هایی بیان و در آنها عدم تقارن RI نیز نشان داده خواهد شد. در بخش ۶ نیز با استفاده از روش‌های شبیه‌سازی و معیار میانگین توان های دوم خطای برآوردگر (MSE) به بررسی تعدادی از نتایج پرداخته می‌شود.

## ۲ تابع مفصل و نتایج مقدماتی

مفصل‌ها توابعی هستند که تابع‌های توزیع چند متغیره را به توزیع‌های کناری آن‌ها ربط می‌دهند. اگر $C(\cdot,\cdot)$ تابع مفصل دومتغیره باشد، آنگاه با استفاده از قضیه اسکلار (۱۹۵۹) تابع توزیع تجمعی $F$ با توابع توزیع کناری $F_۱$ و $F_۲$ می‌تواند به‌صورت

$$F(x,y)=C(F_۱(x),F_۲(y)),\quad x,y\in\overline{\mathbb{R}}=[-\infty,+\infty], \tag{۴}$$

[1] Copula cumulative residual inaccuracy
[2] Radially symmetric
[3] Copula cumulative past inaccuracy
[4] Upper orthant order

مرتبط شود. تابع مفصل بقاء با نماد $\widehat{C}(u,v)$ نشان داده می‌شود که با تابع مفصل به‌صورت

$$\widehat{C}(u,v)=-1+u+v+C(1-u,1-v), \tag{۵}$$

در ارتباط است. فرض کنید $\overline{F}(\cdot,\cdot)$ تابع بقاء یک بردار تصادفی باشد و $\overline{F}_1(\cdot)$ و $\overline{F}_2(\cdot)$ تابع‌های بقاء کناری باشند، در این‌صورت برابری

$$\overline{F}(x,y)=\widehat{C}(\overline{F}_1(x),\overline{F}_2(y)),\quad x,y\in\overline{\mathbb{R}}=[-\infty,+\infty], \tag{۶}$$

برقرار است. با تغییر متغیر $u=\overline{F}_1(x)$ و $v=\overline{F}_2(y)$ در رابطه (۶) نتیجه می‌شود

$$\widehat{C}(u,v)=\overline{F}(\overline{F}_1^{-1}(u),\overline{F}_2^{-1}(v)). \tag{۷}$$

برای جزئیات بیش‌تر درباره نظریه مفصل و کاربردهای آن به نلسن (۲۰۰۷) مراجعه شود.

گاهی به جای برآورد توزیع توأم داده‌ها، تابع مفصل یا تابع مفصل بقاء مناسب با داده‌ها را با توجه به ساختار وابستگی‌شان برآورد می‌کنند. به خاطر وجود داده‌های گمشده و یا داده‌های نادرست ممکن است در تشخیص تابع مفصل یا تابع مفصل بقاء اصلی اشتباه شود. به همین منظور حسینی و احمدی (۱۳۹۷) معیار نادرستی گذشته تجمعی مفصلی را به‌صورت

$$PI(\boldsymbol{X},\boldsymbol{Y})=-\int_0^1\int_0^1 C_{\boldsymbol{X}}(u,v)\log C_{\boldsymbol{Y}}(G_1(F_1^{-1}(u)),G_2(F_2^{-1}(v)))dvdu, \tag{۸}$$

معرفی کردند. این معیار وقتی کاربرد دارد که به جای استفاده از تابع مفصل درست $C_{\boldsymbol{X}}(\cdot,\cdot)$ مفصل $C_{\boldsymbol{Y}}(\cdot,\cdot)$ به کار گرفته می‌شود. حال اگر به جای تابع مفصل، تابع مفصل بقاء مورد توجه باشد، PI کاربرد ندارد. به همین منظور معیاری جدید با عنوان نادرستی مانده تجمعی مفصلی معرفی می‌شود. فرض کنید توابع مفصل بقاء $\widehat{C}_{\boldsymbol{X}}$ و $\widehat{C}_{\boldsymbol{Y}}$ وجود داشته باشد، به طوری که

$$\overline{F}(x,y)=\widehat{C}_{\boldsymbol{X}}(\overline{F}_1(x),\overline{F}_2(y))\quad \text{و}\quad \overline{G}(x,y)=\widehat{C}_{\boldsymbol{Y}}(\overline{G}_1(x),\overline{G}_2(y)).$$

در این‌حالت نادرستی مانده تجمعی مفصلی به‌صورت

$$RI(\boldsymbol{X},\boldsymbol{Y}) = -\int_۰^۱\int_۰^۱ \widehat{C}_{\boldsymbol{X}}(u,v)\log \widehat{C}_{\boldsymbol{Y}}(\overline{G}_۱(\overline{F}_۱^{-۱}(u)),\overline{G}_۲(\overline{F}_۲^{-۱}(v)))dvdu, \quad (۹)$$

تعریف می‌شود. چون $\widehat{C}$ تابع مفصل است، پس $۰ \leq \widehat{C}(\cdot,\cdot) \leq ۱$ و معیار (۹) مانند نادرستی گذشته تجمعی مفصلی همواره نامنفی است. اگر $\boldsymbol{X}$ و $\boldsymbol{Y}$ هم‌توزیع باشند، آن‌گاه $RI(\boldsymbol{X},\boldsymbol{Y}) = RI(\boldsymbol{X},\boldsymbol{X})$ که در آن $RI(\boldsymbol{X},\boldsymbol{X})$ به آنتروپی مانده تجمعی مفصلی معروف است و با نماد $RE(\boldsymbol{X})$ نشان داده می‌شود. فرض کنید متغیرهای تصادفی $Y_۱$ و $Y_۲$ به ترتیب با $X_۱$ و $X_۲$ هم‌توزیع هستند. در این‌صورت معیار (۹) به

$$RI(\boldsymbol{X},\boldsymbol{Y}) = -\int_۰^۱\int_۰^۱ \widehat{C}_{\boldsymbol{X}}(u,v)\log \widehat{C}_{\boldsymbol{Y}}(u,v)dvdu,$$

تبدیل می‌شود، که در بخش‌های بعدی نتایجی برای آن بیان می‌شود.

## ۳ نتایجی تحت مفروضات تقارن

**تعریف ۱** (تقارن شعاعی). *بردار تصادفی $\boldsymbol{X} = (X_۱, X_۲)$ پیرامون نقطه $(a,b)$ متقارن شعاعی است اگر و تنها اگر بردارهای تصادفی $(X_۱ - a, X_۲ - b)$ و $(a - X_۱, b - X_۲)$ تابع‌های توزیع توأم یکسانی داشته باشند. به عبارتی دیگر، برای هر $(x,y) \in S_{\boldsymbol{X}}$، $F(a+x,b+y) = \overline{F}(a-x,b-y)$. برای اطلاعات بیش‌تر درباره تقارن در حالت دو متغیره به امبلارد و جرارد (۲۰۰۲) مراجعه شود.*

**گزاره ۱.** *فرض کنید تکیه‌گاه بردارهای تصادفی $\boldsymbol{X}$ و $\boldsymbol{Y}$ به‌صورت*

$$S = (a-c,a+c)\times(b-d,b+d), \qquad a,b\in\mathbb{R}, \quad c,d>۰,$$

*باشد. اگر $\boldsymbol{X}$ و $\boldsymbol{Y}$ متقارن شعاعی باشند، آن‌گاه $RI(\boldsymbol{X},\boldsymbol{Y}) = PI(\boldsymbol{X},\boldsymbol{Y})$.*

**برهان:** بنابر تقارن شعاعی بردارهای تصادفی،

$$F(۲a-x,۲b-y) = \overline{F}(x,y), \quad G(۲a-x,۲b-y) = \overline{G}(x,y), \quad (x,y)\in S.$$

بنابراین از رابطه‌های (۷) و (۹) نتیجه می‌شود

$$\begin{aligned} RI(\boldsymbol{X},\boldsymbol{Y}) &= -\int_{a-c}^{a+c}\int_{b-d}^{b+d} f_۱(x)f_۲(y)\overline{F}(x,y)\log\overline{G}(x,y)dydx \\ &= -\int_{a-c}^{a+c}\int_{b-d}^{b+d} f_۱(۲a-z)f_۲(۲b-w)F(z,w)\log G(z,w)dwdz. \end{aligned} \quad (۱۰)$$

از طرفی دیگر، چون بردار تصادفی $\boldsymbol{X}$ متقارن شعاعی است با توجه به (۴) و (۶) می‌توان برابری

$$C_{\boldsymbol{X}}(F_۱(۲a-x),F_۲(۲b-y)) = \widehat{C}_{\boldsymbol{X}}(\overline{F}_۱(x),\overline{F}_۲(y)), \quad (۱۱)$$

را به دست آورد. برای بردار تصادفی $\boldsymbol{Y}$ نیز به طور مشابه می‌توان نتیجه گرفت. بنابراین برای $y=b-d$ و $x=a-c$ به ترتیب نتیجه می‌شود که $F_۱(۲a-x)=\overline{F}_۱(x)$ و $F_۲(۲b-y)=\overline{F}_۲(y)$. از (۱۰) با توجه به (۸) برابری $RI(\boldsymbol{X},\boldsymbol{Y})=PI(\boldsymbol{X},\boldsymbol{Y})$ به دست می‌آید.

در حقیقت گزاره ۱ تعمیم نتایج در حالت یک متغیره است که اگر $X$ و $Y$ به ترتیب دارای توزیع‌های متقارن $F$ و $G$ حول صفر باشند، آنگاه آنتروپی تجمعی گذشته برابر آنتروپی باقیمانده است و نادرستی تجمعی گذشته برابر نادرستی تجمعی باقیمانده است. برای عکس نتیجه گزاره ۱، کلاس توزیع‌های دو متغیره با تکیه‌گاه $S=(a-c,a+c)\times(b-d,b+d)$ به‌صورت

$$C_۱ = \{H : \overline{H}(a+x,b+y) \leq \text{ یا } \geq H(a-x,b-y),\}$$

را درنظر بگیرید، که در آن $c,d>۰$، $a,b\in\mathbb{R}$.

**گزاره ۲.** *فرض کنید $RI(\boldsymbol{X},\boldsymbol{Y})=PI(\boldsymbol{X},\boldsymbol{Y})$ و تکیه‌گاه بردارهای تصادفی $\boldsymbol{X}$ و $\boldsymbol{Y}$ برابر با $S$ باشد. اگر $\boldsymbol{X}$ متقارن شعاعی باشد و $G\in C_۱$ آن‌گاه $\boldsymbol{Y}$ نیز متقارن شعاعی است.*

**برهان:** با استفاده از (۶) و (۹) داریم

$$RI(\boldsymbol{X},\boldsymbol{Y}) = -\int_{a-c}^{a+c}\int_{b-d}^{b+d} f_۱(x)f_۲(y)\overline{F}(x,y)\log\overline{G}(x,y)dydx. \quad (۱۲)$$

همچنین با استفاده از (۴) و (۸) داریم

$$PI(\boldsymbol{X},\boldsymbol{Y}) = -\int_{a-c}^{a+c}\int_{b-d}^{b+d} f_1(x)f_2(y)F(x,y)\log G(x,y)dydx. \qquad (۱۳)$$

با توجه به فرض $RI(\boldsymbol{X},\boldsymbol{Y}) = PI(\boldsymbol{X},\boldsymbol{Y})$ از روابط (۱۲) و (۱۳) نتیجه می‌شود

$$\int_{a-c}^{a+c}\int_{b-d}^{b+d} f_1(x)f_2(y)\overline{F}(x,y)\log \overline{G}(x,y)dydx = \int_{a-c}^{a+c}\int_{b-d}^{b+d} f_1(x)f_2(y)F(x,y)\log G(x,y)dydx.$$

چون $\boldsymbol{X}$ متقارن شعاعی است، داریم

$$\int_{-c}^{c}\int_{-d}^{d} f_1(a-z)f_2(b-w)F(a-z,b-w)\log\frac{G(a-z,b-w)}{\overline{G}(a+z,b+w)}dwdz = 0. \qquad (۱۴)$$

طبق فرض $G \in C_1$ از رابطه (۱۴) برابری $\overline{G}(a+x,b+y) = G(a-x,b-y)$ نتیجه می‌شود.

## ۴ نتایجی تحت فرض مدل نرخ خطرات متناسب

اگر برای دو تابع توزیع احتمال $F$ و $F_0$ برابری $\overline{F}(\cdot) = \overline{F}_0^{\alpha}(\cdot)$ برقرار باشد، که در آن $\alpha$ مقدار حقیقی نامنفی و مخالف یک است، در این صورت گویند $F$ متعلق به خانواده مدل نرخ خطر متناسب با تابع توزیع پایه $F_0$ است. این مدل کاربردهای فراوانی در مباحث تحلیل بقاء و نظریه قابلیت اعتماد دارد که شامل تعداد زیادی از توزیع‌های آماری معروف از جمله، نمایی، وایبل یک پارامتری، رایلی، لوماکس و بر نوع ۱۲ است (مارشال و الکین ، ۲۰۰۷). سنجری فارسی‌پور و ریاحی (۱۳۹۲) استنباط درستنمایی و بیزی برای پارامتر تنش-مقاومت در خانواده‌های نرخ خطر متناسب و معکوس براساس داده‌های رکوردی انجام دادند.

**گزاره ۳.** *اگر متغیرهای تصادفی $Y_1$ و $Y_2$ مستقل از یک‌دیگر باشند و برای هر $t > 0$ داشته باشیم*

$\overline{G}_۱(t) = \overline{F}_۱^{\alpha}(t)$ و $\overline{G}_۲(t) = \overline{F}_۲^{\beta}(t)$ که $\alpha$ و $\beta$ مقادیر حقیقی نامنفی و مخالف با یک هستند، آن‌گاه

$$\frac{۲}{۳۶}(\alpha+\beta) \leq RI(\boldsymbol{X},\boldsymbol{Y}) \leq \frac{۷}{۳۶}(\alpha+\beta).$$

**برهان:** با توجه به رابطه (۹) داریم

$$RI(\boldsymbol{X},\boldsymbol{Y}) = -\int_{۰}^{۱}\int_{۰}^{۱} \widehat{C}_{\boldsymbol{X}}(u,v)\log \widehat{C}_{\boldsymbol{Y}}(\overline{G}_۱(\overline{F}_۱^{-۱}(u)), \overline{G}_۲(\overline{F}_۲^{-۱}(v)))dvdu.$$

با استفاده از نامساوی فرشه-هافدینگ[1] $\max\{u+v-۱,۰\} \leq C(u,v) \leq \min\{u,v\}$ نتیجه می‌شود

$$\begin{aligned} RI(\boldsymbol{X},\boldsymbol{Y}) &\leq -\int_{۰}^{۱}\int_{۰}^{۱} \min\{u,v\}\log \widehat{C}_{\boldsymbol{Y}}(\overline{G}_۱(\overline{F}_۱^{-۱}(u)), \overline{G}_۲(\overline{F}_۲^{-۱}(v)))dvdu \\ &= -\int_{۰}^{۱}\int_{۰}^{u} v\log \widehat{C}_{\boldsymbol{Y}}(\overline{G}_۱(\overline{F}_۱^{-۱}(u)), \overline{G}_۲(\overline{F}_۲^{-۱}(v)))dvdu \\ &\quad -\int_{۰}^{۱}\int_{u}^{۱} u\log \widehat{C}_{\boldsymbol{Y}}(\overline{G}_۱(\overline{F}_۱^{-۱}(u)), \overline{G}_۲(\overline{F}_۲^{-۱}(v)))dvdu. \end{aligned}$$

با استفاده از فرض استقلال $Y_۱$ و $Y_۲$ داریم

$$\begin{aligned} RI(\boldsymbol{X},\boldsymbol{Y}) \leq &-\int_{\ell_{X_۱}}^{r_{X_۱}} f_۱(x)\log \overline{G}_۱(x)\int_{F_۲^{-۱}(F_۱(x))}^{r_{X_۲}} f_۲(y)\overline{F}_۲(y)\log \overline{G}_۲(y)dydx \\ &-\int_{\ell_{X_۱}}^{r_{X_۱}} f_۱(x)\int_{F_۲^{-۱}(F_۱(x))}^{r_{X_۲}} f_۲(y)\overline{F}_۲(y)\log \overline{G}_۲(y)dydx \\ &-\int_{\ell_{X_۱}}^{r_{X_۱}} f_۱(x)\overline{F}_۱(x)\log \overline{G}_۱(x)\int_{\ell_{X_۲}}^{F_۲^{-۱}(F_۱(x))} f_۲(y)dydx \\ &-\int_{\ell_{X_۱}}^{r_{X_۱}} f_۱(x)\overline{F}_۱(x)\int_{\ell_{X_۲}}^{F_۲^{-۱}(F_۱(x))} f_۲(y)\log \overline{G}_۲(y)dydx. \end{aligned}$$

[1] Frechet-Hoeffding inequality

با قرار دادن $\overline{G}_1(t)=\overline{F}_1^{\alpha}(t)$ و $\overline{G}_2(t)=\overline{F}_2^{\beta}(t)$ و حل انتگرال با روش جزء به جزء، نابرابری

$$RI(\boldsymbol{X},\boldsymbol{Y})\leq \frac{7}{36}(\alpha+\beta), \tag{15}$$

به دست می‌آید. چون نامساوی $\max\{u+v-1,0\}\leq C(u,v)$ برای هر تابع مفصل $C(\cdot,\cdot)$ از جمله مفصل بقاء $\widehat{C}_{\boldsymbol{X}}(\cdot,\cdot)$ برقرار است، بنابراین از رابطه (۹) نتیجه می‌شود

$$\begin{aligned} RI(\boldsymbol{X},\boldsymbol{Y}) &\geq -\int_0^1\int_0^1 \max\{u+v-1,0\}\log \widehat{C}_{\boldsymbol{Y}}(\overline{G}_1(\overline{F}_1^{-1}(u)),\overline{G}_2(\overline{F}_2^{-1}(v)))dvdu\\ &=-\int_0^1\int_{1-u}^1 (u+v-1)\log \widehat{C}_{\boldsymbol{Y}}(\overline{G}_1(\overline{F}_1^{-1}(u)),\overline{G}_2(\overline{F}_2^{-1}(v)))dvdu. \end{aligned}$$

به طور مشابه با (۱۵) نابرابری $RI(\boldsymbol{X},\boldsymbol{Y})\geq \frac{2}{36}(\alpha+\beta)$ حاصل می‌شود.

**فرع ۱.** *اگر متغیرهای تصادفی $Y_1$ و $Y_2$ مستقل از هم باشند و متغیر تصادفی $Y_i$ هم‌توزیع با $X_i$ باشد، آن‌گاه*

$$\frac{4}{36}\leq RI(\boldsymbol{X},\boldsymbol{Y})\leq \frac{14}{36}.$$

**گزاره ۴.** *اگر $\boldsymbol{X}$ و $\boldsymbol{Y}$ به‌صورت $\overline{G}(x,y)=\overline{F}^{\alpha}(x,y)$ با $\alpha>0$ و $\alpha\neq 1$ در مدل نرخ خطر متناسب صدق کنند، آن‌گاه $RI(\boldsymbol{X},\boldsymbol{Y})=\alpha RE(\boldsymbol{X})$.*

**برهان:** از (۶) و (۹) داریم

$$RI(\boldsymbol{X},\boldsymbol{Y})=-\int_{\ell_{X_1}}^{r_{X_1}}\int_{\ell_{X_2}}^{r_{X_2}} f_1(x)f_2(y)\overline{F}(x,y)\log\overline{G}(x,y)dydx. \tag{16}$$

با قرار دادن $\overline{G}(x,y)=\overline{F}^{\alpha}(x,y)$ در رابطه (۱۶) نتیجه به دست می‌آید.

**گزاره ۵.** *اگر $RI(\boldsymbol{X},\boldsymbol{Y})=\alpha RE(\boldsymbol{X})$ طوری‌که $\alpha>0$، $\alpha\neq 1$ و $G$ متعلق به خانواده $\{H;\overline{H}(x,y)\geq(\leq)\overline{F}^{\alpha}(x,y)\}$ باشد، آن‌گاه $\boldsymbol{X}$ و $\boldsymbol{Y}$ در مدل نرخ خطرات متناسب به‌صورت $\overline{G}(x,y)=\overline{F}^{\alpha}(x,y)$ صدق می‌کنند.*

**برهان:** بنا به فرض $RI(\boldsymbol{X},\boldsymbol{Y})=\alpha RE(\boldsymbol{X})$ و رابطه (۱۶) و با توجه به این‌که

$RE(\boldsymbol{X}) = RI(\boldsymbol{X}, \boldsymbol{X})$ نتیجه می‌شود

$$\int_{\ell_{X_1}}^{r_{X_1}} \int_{\ell_{X_2}}^{r_{X_2}} \overline{F}(x,y) f_1(x) f_2(y) \log \frac{\overline{G}(x,y)}{\overline{F}^{\alpha}(x,y)} dy dx = 0. \tag{۱۷}$$

چون $\overline{G}(x,y) \geq (\leq) \overline{F}^{\alpha}(x,y)$، از (۱۷) نتیجه می‌شود $\overline{G}(x,y) = \overline{F}^{\alpha}(x,y)$.

## ۵ نتایجی براساس ترتیب مربعی بالا برای RI

**تعریف ۲** (ترتیب مربعی بالا). *بردار تصادفی $\boldsymbol{X}$ را در ترتیب مربعی بالا، کوچک‌تر یا مساوی با بردار تصادفی $\boldsymbol{Y}$ گویند اگر و تنها اگر برای هر $(x,y) \in \mathbb{R}_+^2$ نابرابری $\overline{F}(x,y) \leq \overline{G}(x,y)$ برقرار باشد و با نماد $\boldsymbol{X} \leq_{UO} \boldsymbol{Y}$ نشان داده می‌شود.*

برای کسب اطلاعات درباره ترتیب‌های تصادفی به شیکد و شانتی‌کومار (۲۰۰۷) مراجعه شود.

**گزاره ۶.** *اگر $\boldsymbol{X} \leq_{UO} \boldsymbol{Y}$ آن‌گاه $RI(\boldsymbol{Z}, \boldsymbol{Y}) \leq RI(\boldsymbol{Z}, \boldsymbol{X})$.*

**برهان:** با توجه به فرض $\boldsymbol{X} \leq_{UO} \boldsymbol{Y}$ نابرابری $\overline{F}(x,y) \leq \overline{G}(x,y)$ نتیجه می‌شود. با استفاده از رابطه (۹) می‌توان نتیجه گرفت که $RI(\boldsymbol{Z}, \boldsymbol{Y}) \leq RI(\boldsymbol{Z}, \boldsymbol{X})$.

**گزاره ۷.** *اگر $\boldsymbol{X} \leq_{UO} \boldsymbol{Y}$ و متغیرهای تصادفی $Y_1$ و $Y_2$ به‌ترتیب هم‌توزیع با $X_1$ و $X_2$ باشند، آن‌گاه*

$$RI(\boldsymbol{X}, \boldsymbol{Z}) \leq RI(\boldsymbol{Y}, \boldsymbol{Z}).$$

**برهان:** اثبات مشابه برهان گزاره ۶ است.

**گزاره ۸.** *برای بردارهای تصادفی $\boldsymbol{X}$، $\boldsymbol{Y}$ و $\boldsymbol{Z}$ فرض کنید متغیرهای تصادفی $Z_1$ و $Z_2$ به ترتیب هم‌توزیع با $X_1$ و $X_2$ باشند.*

*الف- اگر $\boldsymbol{Z} \leq_{UO} \boldsymbol{Y}$ و $\boldsymbol{Z} \leq_{UO} \boldsymbol{X}$ آن‌گاه $RI(\boldsymbol{Z}, \boldsymbol{Y}) \leq RI(\boldsymbol{X}, \boldsymbol{Y}) \leq RI(\boldsymbol{X}, \boldsymbol{Z})$.*

*ب- اگر $\boldsymbol{Y} \leq_{UO} \boldsymbol{Z} \leq_{UO} \boldsymbol{X}$ آن‌گاه $RI(\boldsymbol{X}, \boldsymbol{Y}) \leq \min\{RI(\boldsymbol{X}, \boldsymbol{Z}), RI(\boldsymbol{Z}, \boldsymbol{Y})\}$.*

*ج- اگر $\boldsymbol{X} \leq_{UO} \boldsymbol{Z}$ و $\boldsymbol{Y} \leq_{UO} \boldsymbol{Z}$ آن‌گاه $RI(\boldsymbol{X}, \boldsymbol{Z}) \leq RI(\boldsymbol{X}, \boldsymbol{Y}) \leq RI(\boldsymbol{Z}, \boldsymbol{Y})$.*

**برهان:** الف) از فرض $\boldsymbol{Y} \leq_{UO} \boldsymbol{Z}$ با توجه به گزاره ۶ نتیجه می‌شود

$$RI(\boldsymbol{X},\boldsymbol{Z}) \leq RI(\boldsymbol{X},\boldsymbol{Y}). \tag{۱۸}$$

از طرفی دیگر، از فرض‌های $\boldsymbol{Z} \leq_{UO} \boldsymbol{X}$ و هم‌توزیعی متغیرهای $Z_۱$ و $Z_۲$ به ترتیب با $X_۱$ و $X_۲$ با توجه به گزاره ۷ می‌توان نتیجه گرفت

$$RI(\boldsymbol{X},\boldsymbol{Y}) \leq RI(\boldsymbol{Z},\boldsymbol{Y}). \tag{۱۹}$$

از نامساوی‌های (۱۸) و (۱۹) داریم $RI(\boldsymbol{X},\boldsymbol{Z}) \leq RI(\boldsymbol{X},\boldsymbol{Y}) \leq RI(\boldsymbol{Z},\boldsymbol{Y})$.
برهان قسمت‌های ب و پ نیز مانند قسمت الف بیان می‌شود.

**گزاره ۹.** *اگر $\boldsymbol{Y} \leq_{UO} \boldsymbol{X}$ آنگاه*
*الف- $RI(\boldsymbol{Z},\boldsymbol{Y}) \leq RI(\boldsymbol{Z},\boldsymbol{X}) + RI(\boldsymbol{X},\boldsymbol{Y})$.*
*ب- اگر $Y_۱$ و $Y_۲$ به‌ترتیب هم‌توزیع با $X_۱$ و $X_۲$ باشند، آن‌گاه*

$$RI(\boldsymbol{X},\boldsymbol{Z}) \leq RI(\boldsymbol{X},\boldsymbol{Y}) + RI(\boldsymbol{Y},\boldsymbol{Z}).$$

**برهان:** اثبات الف با استفاده از گزاره ۶ و نامنفی بودن $RI(\boldsymbol{X},\boldsymbol{Y})$ قابل بیان است. برهان قسمت ب نیز با استفاده از گزاره ۷ و نامنفی بودن $RI(\boldsymbol{X},\boldsymbol{Y})$ به سادگی قابل بیان است.

**مثال ۱.** *فرض کنید $\boldsymbol{X}$ و $\boldsymbol{Y}$ دارای مفصل‌های بقاء $FGM$ و $AMH$ به ترتیب به‌صورت*

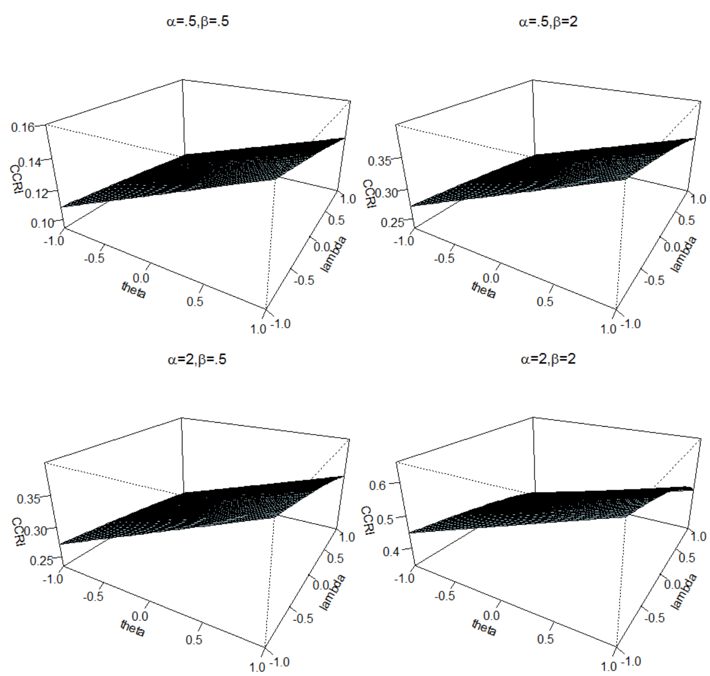
$$\widehat{C}_{\boldsymbol{X}}(u,v) = uv(۱+\theta(۱-u)(۱-v)) \quad \text{و} \quad \widehat{C}_{\boldsymbol{Y}}(u,v) = \frac{uv}{۱-\lambda(۱-u)(۱-v)},$$

*باشند، که در آن $-۱ \leq \theta \leq ۱$ و $-۱ \leq \lambda \leq ۱$ ضرایب وابستگی هستند (نلسن، ۲۰۰۷). همچنین $X_۱$ و $X_۲$ دارای توزیع نمایی استاندارد و $Y_۱$ و $Y_۲$ دارای توزیع نمایی به ترتیب با پارامترهای $\alpha$ و $\beta$*

باشند. با استفاده از رابطه (۹) داریم

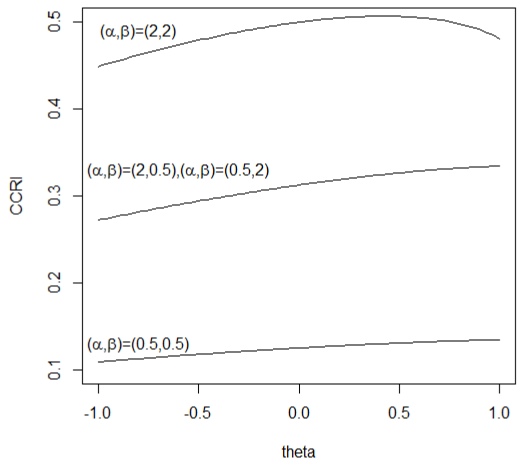
$$RI(\boldsymbol{X},\boldsymbol{Y})= -\int_{0}^{1}\int_{0}^{1} uv(1+\theta(1-u)(1-v))\log\frac{u^{\alpha}v^{\beta}}{1-\lambda(1-u^{\alpha})(1-v^{\beta})}dvdu,$$

که به $\alpha$، $\beta$، $\theta$ و $\lambda$ وابسته است. رفتار معیار بالا نسبت به پارامترها به‌صورت تحلیلی به سادگی قابل بررسی نیست. همان طورکه در شکل ۱ ملاحظه می‌شود $RI(\boldsymbol{X},\boldsymbol{Y})$ نسبت به ضریب وابستگی $\theta$ صعودی و نسبت به ضریب وابستگی $\lambda$ نزولی است. همچنین با افزایش $\alpha+\beta$ مقدار $RI(\boldsymbol{X},\boldsymbol{Y})$ افزایش می‌یابد. شکل ۲ رفتار $RI(\boldsymbol{X},\boldsymbol{Y})$ را نسبت به $\theta$ زمانی‌که هر دو مفصل $FGM$ و $AMH$ دارای ضریب وابستگی برابر $\theta$ هستند، نشان می‌دهد. همان‌طورکه مشاهده می‌شود $RI(\boldsymbol{X},\boldsymbol{Y})$ برای هر $(\alpha,\beta)$ نسبت به $\theta$ لزومی ندارد که یکنوا باشد و با افزایش $\alpha+\beta$ افزایش می‌یابد.




**شکل ۱.** نمودار $RI(\boldsymbol{X},\boldsymbol{Y})$ برحسب ضرایب وابستگی $\theta$ و $\lambda$

**مثال ۲.** اگر متغیرهای تصادفی $Y_1$ و $Y_2$ مستقل از هم باشند و $\overline{G}_1(x)=\overline{F}_1^{\alpha}(x)$، $\overline{G}_2(y)=\overline{F}_2^{\beta}(y)$




**شکل ۲.** نمودار $RI(\boldsymbol{X},\boldsymbol{Y})$ برحسب ضریب وابستگی $\theta$ زمانی‌که $\lambda = \theta$

*و بردار $\boldsymbol{X}$ دارای تابع مفصل $FGM$ باشد، از استقلال $Y_۱$ و $Y_۲$ نتیجه می‌شود $C_{\boldsymbol{Y}}(u,v) = uv$، همچنین با توجه به رابطه‌های (۵) و (۹) داریم*

$$\begin{aligned} RI(\boldsymbol{X},\boldsymbol{Y}) &= -\int_۰^۱\int_۰^۱ \{-۱+u+v+(۱-u)(۱-v)(۱+\theta uv)\}\log u^\alpha v^\beta dudv \\ &= (\frac{۱}{۸}+\frac{۵\theta}{۲۱۶})(\alpha+\beta). \end{aligned} \tag{۲۰}$$

*چون $-۱ \leq \theta \leq ۱$ می‌توان نشان داد $\frac{۷}{۳۶}(\alpha+\beta) \leq (\frac{۱}{۸}+\frac{۵\theta}{۲۱۶})(\alpha+\beta) \leq \frac{۲}{۳۶}(\alpha+\beta)$. بنابراین گزاره ۳ برای این مثال برقرار است.*

**مثال ۳.** *فرض کنید در مثال ۲ توزیع‌های حاشیه‌ای نیز متقارن باشند، در این‌صورت از (۵) و (۸) داریم*

$$\begin{aligned} PI(\boldsymbol{X},\boldsymbol{Y}) &= -\int_۰^۱\int_۰^۱ \{uv(۱+\theta(۱-u)(۱-v))\}\log u^\alpha v^\beta dudv \\ &= (\frac{۱}{۸}+\frac{۵\theta}{۲۱۶})(\alpha+\beta). \end{aligned} \tag{۲۱}$$

*با توجه به (۲۰) و (۲۱) داریم $RI(\boldsymbol{X},\boldsymbol{Y}) = PI(\boldsymbol{X},\boldsymbol{Y})$. بنابراین گزاره ۱ در این مثال برقرار است.*

*با توجه به رابطه‌های (۴) و (۶) و فرض تقارن توزیع‌های حاشیه‌ای، می‌توان نشان داد که شرایط گزاره ۱ برای بردارهای تصادفی $\boldsymbol{X}$ و $\boldsymbol{Y}$ به‌ترتیب با توابع مفصل $C_{\boldsymbol{X}}(u,v)=uv\{1+\theta(1-u)(1-v)\}$ و $C_{\boldsymbol{Y}}(u,v)=uv$ برقرار است.*

**مثال ۴.** *فرض کنید $\widehat{C}_{\boldsymbol{Z}}(u,v)=\min\{u,v\}$ و برای $i=1,2$، $Z_i \leq_{st} X_i$ و $Y_i \leq_{st} Z_i$، در این‌صورت ترتیب‌های $\boldsymbol{Z} \leq_{UO} \boldsymbol{Y}$ و $\boldsymbol{Z} \leq_{UO} \boldsymbol{X}$ برقرار هستند. با توجه به گزاره ۶ نتیجه می‌شود $RI(\boldsymbol{X},\boldsymbol{Y}) \leq RI(\boldsymbol{X},\boldsymbol{Z})$، $RI(\boldsymbol{Y},\boldsymbol{X}) \leq RI(\boldsymbol{Y},\boldsymbol{Z})$ و نامساوی‌های مثلث $RI(\boldsymbol{Y},\boldsymbol{Z}) + RI(\boldsymbol{X},\boldsymbol{Y}) \leq RI(\boldsymbol{X},\boldsymbol{Z})$ و $RI(\boldsymbol{Y},\boldsymbol{X}) \leq RI(\boldsymbol{Y},\boldsymbol{Z})$ $+ RI(\boldsymbol{X},\boldsymbol{Z})$ برقرار هستند. حال اگر متغیرهای تصادفی $Z_1$ و $Z_2$ به‌ترتیب هم‌توزیع با $Y_1$ و $Y_2$ باشند، براساس گزاره ۷ داریم $RI(\boldsymbol{Z},\boldsymbol{X}) \leq RI(\boldsymbol{Y},\boldsymbol{X})$. در این مورد نامساوی مثلث به‌صورت $RI(\boldsymbol{Y},\boldsymbol{X}) \leq RI(\boldsymbol{Y},\boldsymbol{Z}) + RI(\boldsymbol{Z},\boldsymbol{X})$ برقرار می‌شود. اگر $Z_1$ و $Z_2$ به ترتیب هم‌توزیع با $X_1$ و $X_2$ باشند، از گزاره ۷ نتیجه می‌شود $RI(\boldsymbol{Z},\boldsymbol{Y}) \leq RI(\boldsymbol{X},\boldsymbol{Y})$ و در این مورد نامساوی مثلث به‌صورت $RI(\boldsymbol{X},\boldsymbol{Y}) \leq RI(\boldsymbol{X},\boldsymbol{Z}) + RI(\boldsymbol{Z},\boldsymbol{Y})$ برقرار است.*

## ۶ مطالعه شبیه‌سازی

فرض کنید $(X_{1,1},X_{1,2}),\cdots,(X_{n,1},X_{n,2})$ نمونه‌ای تصادفی از توزیع توأم $F(x,y)$ باشد، در این حالت برآوردگر تجربی توزیع توأم به‌صورت

$$F_n(x,y)=\frac{1}{n}\sum_{i=1}^{n} I(X_{i,1}\leq x, X_{i,2}\leq y), \tag{۲۲}$$

است، که در آن $I_A$ تابع نشانگر است. به طور مشابه برآوردگر تجربی برای $\overline{F}(x,y)$ به‌صورت

$$\overline{F}_n(x,y)=\frac{1}{n}\sum_{i=1}^{n} I(X_{i,1}> x, X_{i,2}> y). \tag{۲۳}$$

است. یکی از برآوردگرهای چندک جامعه، آماره‌های ترتیبی نمونه هستند. بنابراین اگر آماره‌های ترتیبی مؤلفه اول نمونه زوجی از $\boldsymbol{X}$ با $X_{1:n,1} \leq X_{2:n,1} \leq \cdots \leq X_{n:n,1}$ نشان داده شود، آن‌گاه

$X_{[nu]+1:n,1}$ و $X_{n-[nu]:n,1}$ به ترتیب برآوردگرهای $F_1^{-1}(u)$ (چندک مرتبه $u$ ام) و $\overline{F}_1^{-1}(u)$ هستند،

$$\widehat{F}_1^{-1}(u) = X_{[nu]+1:n,1}, \qquad \widehat{\overline{F}}_1^{-1}(u) = X_{n-[nu]:n,1}. \tag{۲۴}$$

در معیار نادرستی (۸) و (۹) فرض بر این است که تابع مفصل و توزیع‌های حاشیه‌ای بردار $\boldsymbol{Y}$ معلوم هستند و تابع مفصل و توزیع حاشیه‌ای بردار $\boldsymbol{X}$ از طریق داده مشخص می‌شوند. با توجه به موارد فوق و رابطه‌های (۲۲)، (۲۳) و (۲۴) برآوردگر PI به‌صورت

$$\begin{aligned}
\widehat{PI}(\boldsymbol{X},\boldsymbol{Y}) &= -\int_0^1\int_0^1 F_n(\widehat{F}_1^{-1}(u),\widehat{F}_2^{-1}(v))\log C_{\boldsymbol{Y}}(G_1(\widehat{F}_1^{-1}(u)),G_2(\widehat{F}_2^{-1}(v)))dvdu,\\
&= -\int_0^1\int_0^1 \frac{1}{n}\sum_{i=1}^{n} I(X_{i,1}\le X_{[nu]+1:n,1}, X_{i,2}\le X_{[nv]+1:n,2})\\
&\times \log C_{\boldsymbol{Y}}(G_1(X_{[nu]+1:n,1}),G_2(X_{[nv]+1:n,2}))dvdu.
\end{aligned}$$

به دست آورده می‌شود. حال با افراز $[0,1)$ به $n$ زیر بازه به طول $\frac{1}{n}$ به صورت $[0,1)=\cup_{k=0}^{n-1}[\frac{k}{n},\frac{k+1}{n})$ می‌توان نوشت:

$$\begin{aligned}
\widehat{PI}(\boldsymbol{X},\boldsymbol{Y}) &= -\sum_{k=0}^{n-1}\sum_{j=0}^{n-1}\int_{\frac{k}{n}}^{\frac{k+1}{n}}\int_{\frac{j}{n}}^{\frac{j+1}{n}}\frac{1}{n}\sum_{i=1}^{n} I(X_{i,1}\le X_{k+1:n,1}, X_{i,2}\le X_{j+1:n,2})\\
&\times \log C_{\boldsymbol{Y}}(G_1(X_{k+1:n,1}),G_2(X_{j+1:n,2}))dvdu\\
&= -\frac{1}{n^3}\sum_{k=0}^{n-1}\sum_{j=0}^{n-1}\sum_{i=1}^{n} I(X_{i,1}\le X_{k+1:n,1}, X_{i,2}\le X_{j+1:n,2})\\
&\times \log C_{\boldsymbol{Y}}(G_1(X_{k+1:n,1}),G_2(X_{j+1:n,2})).
\end{aligned} \tag{۲۵}$$

به طور مشابه برای RI داریم

$$\begin{aligned}
\widehat{RI}(\boldsymbol{X},\boldsymbol{Y}) &= -\frac{1}{n^3}\sum_{k=0}^{n-1}\sum_{j=0}^{n-1}\sum_{i=1}^{n} I(X_{i,1}> X_{n-k:n,1}, X_{i,2}> X_{n-j:n,2})\\
&\times \log \widehat{C}_{\boldsymbol{Y}}(\overline{G}_1(X_{n-k:n,1}),\overline{G}_2(X_{n-j:n,2})),
\end{aligned} \tag{۲۶}$$

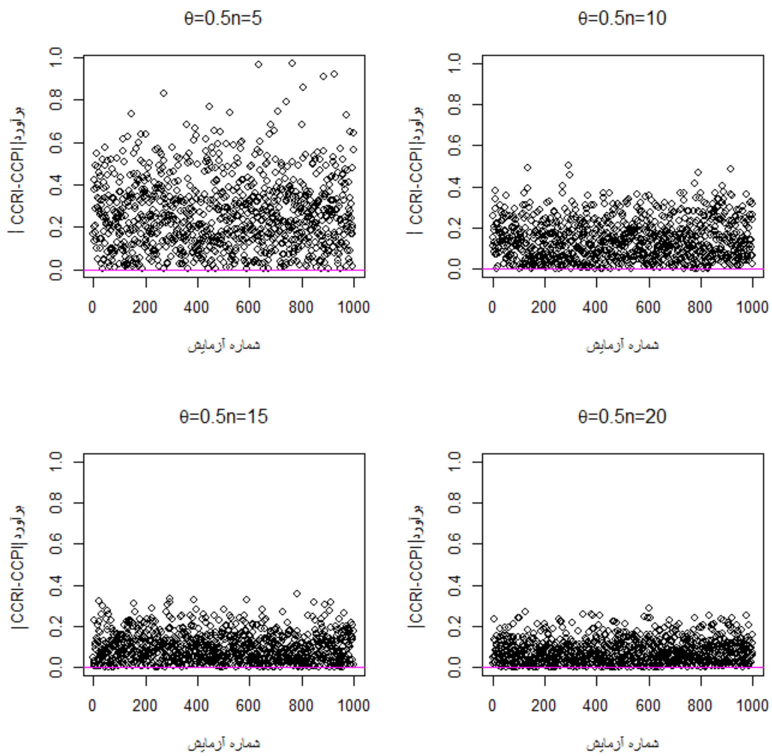
که در آن‌ها $X_{i,\ell}$ $(\ell = 1, 2)$ عضو $i$ام در بین $n$ مؤلفه تصادفی از متغیر تصادفی $X_\ell$ در نمونه‌ای به حجم $n$ از بردار تصادفی $\boldsymbol{X} = (X_1, X_2)$ است و $X_{m:n,l}$ $(\ell = 1, 2)$ آماره مرتب $m$ام در بین $n$ آماره مرتب از متغیر تصادفی $X_\ell$ در نمونه‌ای به حجم $n$ از بردار تصادفی $\boldsymbol{X} = (X_1, X_2)$ است. همچنین $I(X_1 \leq (>)x, X_2 \leq (>)y)$ تابع نشان‌گر در حالت دو متغیره است. لازم به ذکر است که ضریب وابستگی در تابع‌های مفصل و مفصل بقاء مربوط به بردار تصادفی $\boldsymbol{Y}$ در برآوردگرهای (۲۵) و (۲۶) با روش ماکسیمم درستنمایی (MLE) برآورد می‌شود.

**الف- نتیجه حاصل از گزاره ۱:** فرض کنید $C_{\boldsymbol{X}}(u, v) = uv\{1 + \theta(1 - u)(1 - v)\}$ و $C_{\boldsymbol{Y}}(u, v) = \min\{u, v\}$. همچنین متغیرهای تصادفی $X_i$ و $Y_i$ برای $i = 1, 2$ دارای توزیع یکنواخت روی بازه $(0, 1)$ باشند. در اینجا با استفاده از برآوردگرهای معرفی شده در (۲۵) و (۲۶) نتیجه گزاره ۱ را در قالب شبیه‌سازی بررسی می‌کنیم. برای این منظور کدهای برنامه‌نویسی در محیط نرم افزار R نوشته شده است. نتایج به دست آمده با ۱۰۰۰ بار تکرار در جدول ۱ گزارش شده است. این جدول مقادیر MSE را برای برآوردگر $RI(\boldsymbol{X}, \boldsymbol{Y}) - PI(\boldsymbol{X}, \boldsymbol{Y})$ به ازاء مقادیر مختلف ضریب وابستگی $\theta$ و حجم نمونه $n$ نشان می‌دهد. همان‌طور که مشاهده می‌شود برای هر مقدار $\theta_0$ با افزایش حجم نمونه، مقدار MSE کاهش می‌یابد که انتظار نیز همین بود. همچنین به ازاء هر مقدار $n$ با افزایش $\theta_0$ مقدار MSE افزایش می‌یابد که نشان می‌دهد افزایش پارامتر $\theta$ باعث افزایش خطا در این برآورد می‌شود. در شکل ۳ برقراری گزاره ۱ با استفاده از شبیه‌سازی در ۱۰۰۰ بار تکرار آزمایش برای $\theta_0 = 0.5$ و به ازاء مقدارهای متفاوت $n = 5, 10, 15, 20$ بررسی شده است. همان‌طور که مشاهده می‌شود با افزایش حجم نمونه مقدارهای شبیه‌سازی شده $|\widehat{RI} - \widehat{PI}|$ به صفر نزدیک‌تر می‌شوند.

**جدول ۱.** نتایج شبیه‌سازی در ۱۰۰۰ بار تکرار آزمایش برای گزاره ۱

| $n$ | $\theta_0$ | $MSE$ | $n$ | $\theta_0$ | $MSE$ |
|---|---|---|---|---|---|
| ۵ | ۰٫۵ | ۰٫۰۹۷۸۹ | ۵ | ۰٫۱ | ۰٫۰۹۵۲۲ |
| ۱۰ | ۰٫۵ | ۰٫۰۲۶۹۰ | ۱۰ | ۰٫۱ | ۰٫۰۲۴۱۰ |
| ۱۵ | ۰٫۵ | ۰٫۰۱۴۹۵ | ۱۵ | ۰٫۱ | ۰٫۰۱۲۳۲ |
| ۲۰ | ۰٫۵ | ۰٫۰۱۰۲۹ | ۲۰ | ۰٫۱ | ۰٫۰۰۷۷۰ |
| ۵ | −۰٫۵ | ۰٫۰۸۳۶۵ | ۵ | −۰٫۱ | ۰٫۰۹۱۱۴ |
| ۱۰ | −۰٫۵ | ۰٫۰۲۰۱۴ | ۱۰ | −۰٫۱ | ۰٫۰۲۲۷۴ |
| ۱۵ | −۰٫۵ | ۰٫۰۰۹۹۸ | ۱۵ | −۰٫۱ | ۰٫۰۱۱۵۳ |
| ۲۰ | −۰٫۵ | ۰٫۰۰۶۰۸ | ۲۰ | −۰٫۱ | ۰٫۰۰۷۱۵ |
| ۱۵ | ۰ | ۰٫۰۱۱۹۳ | ۵ | ۰ | ۰٫۰۹۳۲۵ |
| ۲۰ | ۰ | ۰٫۰۰۷۴۲ | ۱۰ | ۰ | ۰٫۰۲۳۴۳ |

**ب- نتیجه حاصل از گزاره ۳:** فرض کنید $\overline{G}_1(x) = e^{-\alpha x}$ و $\overline{G}_2(y) = e^{-\beta y}$. جدول ۲




**شکل ۳.** مقادیر شبیه‌سازی شده $|\widehat{RI}-\widehat{PI}|$ در ۱۰۰۰ بار تکرار آزمایش برای گزاره ۱

نتایج شبیه‌سازی ۱۰۰۰ بار تکرار آزمایش با نرم‌افزار R را برای مقدارهای متفاوت $\alpha$، $\beta$، $\theta$ و $n$ نشان می‌دهد. برای هر $(\theta_0, \alpha_0, \beta_0)$ با افزایش حجم نمونه $n$ مطابق انتظار مقدار MSE کاهش می‌یابد. با مقایسه MSE برای $\theta_0 > 0$ و $\theta_0 < 0$، برآوردگر (۲۶) برای RI درحالت $\theta_0 < 0$ نسبت به $\theta_0 > 0$ بهتر است. برای هر $(\alpha_0, \beta_0)$ با افزایش $\theta_0$ مقدار MSE افزایش می‌یابد. همچنین با افزایش هریک از پارامترهای نرخ خطر $\alpha_0$ و $\beta_0$ مقدار MSE افزایش می‌یابد. با توجه به مقدارهای MSE نتیجه می‌شود برآوردگر معرفی شده برای RI مناسب است. در شکل ۴ برقراری گزاره ۳ در ۱۰۰۰ بار تکرار آزمایش برای $n = ۵, ۱۰, ۱۵, ۲۰$ و به ازاء مقدارهای متفاوت $(\theta_0, \alpha_0, \beta_0) = (۰/۵, ۰/۵, ۰/۵)$ بررسی شده است. درصد تعداد نقاط قرارگرفته در بازه $[\frac{۲}{۳۶}, \frac{۷}{۳۶}]$ برای مقدارهای $n = ۵, ۱۰, ۱۵, ۲۰$ محاسبه و به ترتیب مقادیر ۶۹/۵٪، ۹۲/۶٪، ۹۷٪، ۹۸/۷٪ حاصل شده است. ملاحظه می‌شود با افزایش حجم نمونه تعداد بیش‌تری از مقادیر شبیه‌سازی شده $\widehat{RI}$ در بازه $[\frac{۲}{۳۶}, \frac{۷}{۳۶}]$ قرار می‌گیرند.

**جدول ۲.** نتایج شبیه‌سازی در ۱۰۰۰ بار تکرار آزمایش برای گزاره ۳

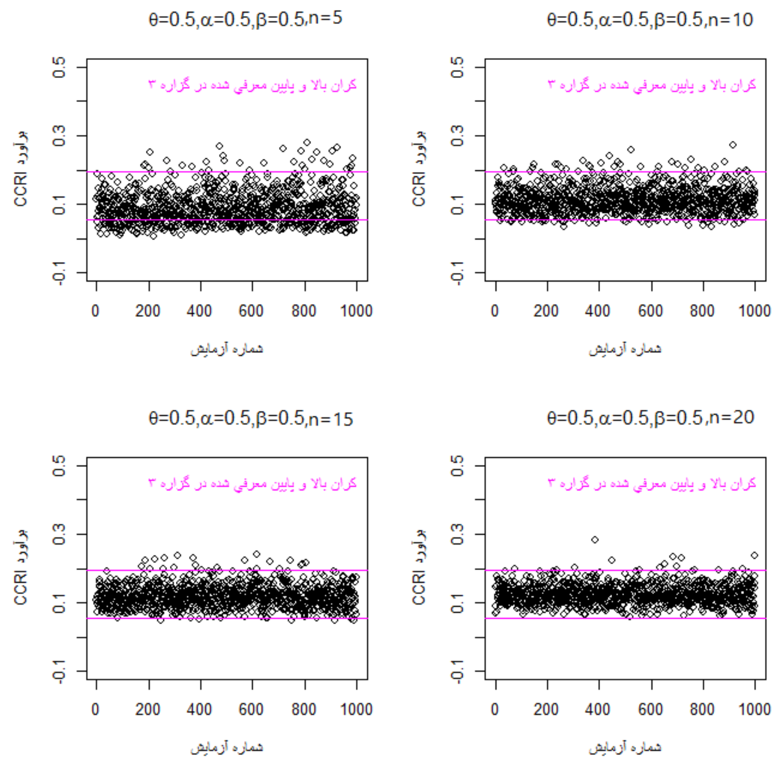
| $MSE$ | $\beta_0$ | $\alpha_0$ | $\theta_0$ | $MSE$ | $\beta_0$ | $\alpha_0$ | $\theta_0$ | $MSE$ | $\beta_0$ | $\alpha_0$ | $\theta_0$ | $n$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ۰/۰۰۳۵۲ | ۰/۵ | ۰/۵ | -۰/۳ | ۰/۰۰۴۱۶ | ۰/۵ | ۰/۵ | ۰/۱ | ۰/۰۰۴۶۶ | ۰/۵ | ۰/۵ | ۰/۵ | ۵ |
| ۰/۰۰۱۷۳ | ۰/۵ | ۰/۵ | -۰/۳ | ۰/۰۰۱۹۰ | ۰/۵ | ۰/۵ | ۰/۱ | ۰/۰۰۱۹۳ | ۰/۵ | ۰/۵ | ۰/۵ | ۱۰ |
| ۰/۰۰۰۹۶ | ۰/۵ | ۰/۵ | -۰/۳ | ۰/۰۰۱۱۸ | ۰/۵ | ۰/۵ | ۰/۱ | ۰/۰۰۱۲۹ | ۰/۵ | ۰/۵ | ۰/۵ | ۱۵ |
| ۰/۰۲۳۳۸ | ۲ | ۰/۵ | -۰/۳ | ۰/۰۲۶۷۶ | ۲ | ۰/۵ | ۰/۱ | ۰/۰۳۰۶۷ | ۲ | ۰/۵ | ۰/۵ | ۵ |
| ۰/۰۱۱۵۳ | ۲ | ۰/۵ | -۰/۳ | ۰/۰۱۳۷۷ | ۲ | ۰/۵ | ۰/۱ | ۰/۰۱۴۵۰ | ۲ | ۰/۵ | ۰/۵ | ۱۰ |
| ۰/۰۰۷۹۲ | ۲ | ۰/۵ | -۰/۳ | ۰/۰۰۸۴۵ | ۲ | ۰/۵ | ۰/۱ | ۰/۰۰۹۷۳ | ۲ | ۰/۵ | ۰/۵ | ۱۵ |
| ۰/۰۲۵۸۲ | ۰/۵ | ۲ | -۰/۳ | ۰/۰۲۸۰۶ | ۰/۵ | ۲ | ۰/۱ | ۰/۰۳۰۶۰ | ۰/۵ | ۲ | ۰/۵ | ۵ |
| ۰/۰۱۱۶۴ | ۰/۵ | ۲ | -۰/۳ | ۰/۰۱۳۸۲ | ۰/۵ | ۲ | ۰/۱ | ۰/۰۱۵۱۰ | ۰/۵ | ۲ | ۰/۵ | ۱۰ |
| ۰/۰۰۷۵۹ | ۰/۵ | ۲ | -۰/۳ | ۰/۰۰۸۵۹ | ۰/۵ | ۲ | ۰/۱ | ۰/۰۰۹۶۴ | ۰/۵ | ۲ | ۰/۵ | ۱۵ |
| ۰/۰۵۵۱۰ | ۲ | ۲ | -۰/۳ | ۰/۰۶۴۸۶ | ۲ | ۲ | ۰/۱ | ۰/۰۶۹۷۸ | ۲ | ۲ | ۰/۵ | ۵ |
| ۰/۰۲۴۳۸ | ۲ | ۲ | -۰/۳ | ۰/۰۲۹۵۳ | ۲ | ۲ | ۰/۱ | ۰/۰۳۳۴۲ | ۲ | ۲ | ۰/۵ | ۱۰ |
| ۰/۰۱۵۳۲ | ۲ | ۲ | -۰/۳ | ۰/۰۱۸۴۳ | ۲ | ۲ | ۰/۱ | ۰/۰۲۲۴۳ | ۲ | ۲ | ۰/۵ | ۱۵ |
| ۰/۰۰۳۰۸ | ۰/۵ | ۰/۵ | -۰/۵ | ۰/۰۰۳۷۶ | ۰/۵ | ۰/۵ | -۰/۱ | ۰/۰۰۴۴۰ | ۰/۵ | ۰/۵ | ۰/۳ | ۵ |
| ۰/۰۰۱۶۰ | ۰/۵ | ۰/۵ | -۰/۵ | ۰/۰۰۱۷۸ | ۰/۵ | ۰/۵ | -۰/۱ | ۰/۰۰۱۹۰ | ۰/۵ | ۰/۵ | ۰/۳ | ۱۰ |
| ۰/۰۰۰۹۱ | ۰/۵ | ۰/۵ | -۰/۵ | ۰/۰۰۱۱۲ | ۰/۵ | ۰/۵ | -۰/۱ | ۰/۰۰۱۲۵ | ۰/۵ | ۰/۵ | ۰/۳ | ۱۵ |
| ۰/۰۲۲۵۰ | ۲ | ۰/۵ | -۰/۵ | ۰/۰۲۵۵۹ | ۲ | ۰/۵ | -۰/۱ | ۰/۰۲۷۸۸ | ۲ | ۰/۵ | ۰/۳ | ۵ |
| ۰/۰۱۰۳۴ | ۲ | ۰/۵ | -۰/۵ | ۰/۰۱۳۳۲ | ۲ | ۰/۵ | -۰/۱ | ۰/۰۱۴۰۹ | ۲ | ۰/۵ | ۰/۳ | ۱۰ |
| ۰/۰۰۷۴۸ | ۲ | ۰/۵ | -۰/۵ | ۰/۰۰۸۴۱ | ۲ | ۰/۵ | -۰/۱ | ۰/۰۰۹۰۱ | ۲ | ۰/۵ | ۰/۳ | ۱۵ |
| ۰/۰۲۳۶۵ | ۰/۵ | ۲ | -۰/۵ | ۰/۰۲۶۶۵ | ۰/۵ | ۲ | -۰/۱ | ۰/۰۲۸۵۰ | ۰/۵ | ۲ | ۰/۳ | ۵ |
| ۰/۰۱۰۸۷ | ۰/۵ | ۲ | -۰/۵ | ۰/۰۱۲۲۶ | ۰/۵ | ۲ | -۰/۱ | ۰/۰۱۳۹۶ | ۰/۵ | ۲ | ۰/۳ | ۱۰ |
| ۰/۰۰۶۹۵ | ۰/۵ | ۲ | -۰/۵ | ۰/۰۰۸۲۷ | ۰/۵ | ۲ | -۰/۱ | ۰/۰۰۹۲۸ | ۰/۵ | ۲ | ۰/۳ | ۱۵ |
| ۰/۰۴۹۶۵ | ۲ | ۲ | -۰/۵ | ۰/۰۶۱۲۶ | ۲ | ۲ | -۰/۱ | ۰/۰۶۷۴۴ | ۲ | ۲ | ۰/۳ | ۵ |
| ۰/۰۲۱۹۸ | ۲ | ۲ | -۰/۵ | ۰/۰۲۸۳۶ | ۲ | ۲ | -۰/۱ | ۰/۰۳۱۸۱ | ۲ | ۲ | ۰/۳ | ۱۰ |
| ۰/۰۱۴۶۲ | ۲ | ۲ | -۰/۵ | ۰/۰۱۶۹۵ | ۲ | ۲ | -۰/۱ | ۰/۰۲۱۲۹ | ۲ | ۲ | ۰/۳ | ۱۵ |

## بحث و نتیجه‌گیری

بررسی معیارهای مختلف برای اندازه‌گیری اطلاعات در داده‌ها یکی از موضوعات مهم و اساسی در زمینه‌های مختلف علوم و مهندسی است. ولی در این کار احتمال رخ دادن اشتباه در ثبت داده‌ها وجود دارد و بنابراین امکان دارد گاهی مواقع داده‌ها دارای اطلاعات نادرست باشند. به همین دلیل آزمایش‌گر در تشخیص مدل اصلی که داده‌ها از آن پیروی می‌کنند، دچار اشتباه می‌شود. معیارهای مختلفی از جمله معیار نادرستی کریج معرفی شده است. در این مقاله با استفاده از تابع مفصل بقاء، یک معیار نادرستی جدید با عنوان نادرستی مانده تجمعی مفصلی معرفی و برخی از ویژگی‌های نظری آن بررسی شد.

## تقدیر و تشکر

نویسندگان از زحمات و پیشنهادات داوران محترم، اعضای هیئت تحریریه و ویراستار مجله که باعث بهبود ارائه مقاله شدند، تشکر می‌کنند.




**شکل ۴.** مقادیر شبیه‌سازی شده $\widehat{RI}$ در ۱۰۰۰ بار تکرار آزمایش برای گزاره ۳

# مراجع


حسینی، ت.، احمدی، ج.، (۱۳۹۷)، نتایجی درباره معیار نادرستی در نظریه اطلاع براساس تابع مفصل، مجموعه مقالات پنجمین سمینار نظریه مفصل و کاربردهای آن، ۲۲-۳۲.

زمان زاده، ا.، ارقامی ن، ر.، (۱۳۸۷)، آزمون نیکویی برازش توزیع‌های نرمال و نمایی برمبنای برآوردگرهای جدید آنتروپی. مجله علوم آماری، **۲**، ۱۷۹-۲۰۰.

سنجری فارسی‌پور، ن، ریاحی، ه.، (۱۳۹۲)، استنباط درستنمایی و بیزی مدل تنش نیرو بر اساس داده های رکوردی در خانواده های نرخ خطر متناسب و معکوس متناسب، مجله علوم آماری، **۷**، ۲۰۷-۲۳۲.

شکرانی، ا.، خراشادی‌زاده، م.، (۱۳۹۷)، اندازه نادرستی پویایی چندکی بین دو متغیر گذشته عمر، مجله

علوم آماری، ۱۲، ۴۴۹–۴۶۸.

Amblard, C., and Girard, S. (2002) Symmetry and Dependence Properties Within a Semiparametric Family of Bivariate Copulas, *Journal of Nonparametric Statistics*, **14**, 715-727.

Cover, T. M., and Thomas, J. A. (2006), *Elements of Information Theory, 2nd edition*, Willey-Interscience: NJ.

Ghosh, A. and Kundu, C. (2019), Bivariate Extension of (dynamic) Cumulative Residual and Past Inaccuracy Measures, *Statistical Papers*, **60**, 2225-2252.

Kerridge, D. F. (1961), Inaccuracy and Inference, *Journal of the Royal Statistical Society.* Series B (Methodological), 184-194.

Marshall, A. W., and Olkin, I. (2007), *Life Distributions*, (Vol. 13), Springer, New York.

Nath, P. (1968), Inaccuracy and Coding Theory, *Metrika*, **13**, 123-135.

Nelsen, R. B. (2007), *An Introduction to Copulas*, Springer, Science and Business Media.

Shaked, M., and Shanthikumar, J. G. (2007), *Stochastic Orders*, Springer Science and Business Media.

Shannon, C. E. (1948), A Mathematical Theory of Communication, *Bell System Technical Journal*, **27**, 379-423.

Sklar, M. (1959), Fonctions de Repartition an Dimensions et Leurs Marges. *Publications de l'Institut de Statistique de l'Université de Paris*, **8**, 229-231.

Journal of Statistical Sciences, Spring and Summer, 2020
Vol. 14, No. 2, pp 389-408
DOI: 10.29252/jss.14.2.389

# Inaccuracy Measure Based on Survival Copula

Hosseini, S. T. and Ahmadi, J.
Department of Statistics, Faculty of Mathematical Sciences, Ferdowsi University of Mashhad, Mashhad, Iran.

**Abstract:** In this paper, using the idea of inaccuracy measure in the information theory, the residual and past inaccuracy measures in the bivariate case are defined based on copula functions. Under the assumption of radial symmetry, the equality of these two criteria is shown, also by the equality between these two criteria, radially symmetrical models are characterized. A useful bound is provided by establishing proportional (inverse) hazard rate models for marginal distributions. Also, the proportional hazard rate model in bivariate mode is characterized by assuming proportionality between the introduced inaccuracy and its corresponding entropy. In addition, orthant orders are used to obtain inequalities. To illustrate the results, some examples and simulations are presented.

**Keywords:** Copula; Entropy; Inaccuracy measure; Orthant orders; Proportional (inverse) hazard rate models; Radially symmetric.

**Mathematics Subject Classification (2010):** 94A17, 62B10, 60E15.