ضریب همبستگی در شرایط نادقیق

رضا زارعی[۱]، طاهره کیانپور[۲]
[۱] استادیار گروه آمار، دانشکده علوم ریاضی، دانشگاه گیلان
[۲] دانش‌آموخته کارشناسی ارشد آمار ریاضی، دانشکده علوم ریاضی، دانشگاه گیلان

چکیده

همبستگی یکی از مهم‌ترین شاخص‌هایی است که به‌طور گسترده در زمینه‌های مختلف به‌کار می‌رود و اندازه‌ای با اهمیت در تحلیل‌داده‌هاست. از آنجایی‌که بسیاری از داده‌های واقعی ممکن است نادقیق باشند، مفهوم همبستگی در محیط نادقیق توسیع داده شد. در این مقاله، ابتدا بر اساس تعریف کلاسیک ضریب همبستگی و اصل توسیع زاده یک معیار برای محاسبه ضریب همبستگی بین اعداد فازی را مطالعه می‌کنیم. رویکرد برنامه‌ریزی ریاضی به‌منظور محاسبه درجه عضویت اندازه‌های فازی بکارگرفته شده است. سپس، روشی ساده برای تعیین و محاسبه دقیق تابع عضویت با استفاده از T-نرم دراستیک و اعمال جبری بر اساس آن مورد بررسی قرارگرفته است که به برنامه‌ریزی ریاضی بستگی ندارد. در پایان برخی روابط برای اندازه‌گیری ضرایب همبستگی بین مجموعه‌های فازی شهودی و مجموعه‌های فازی مردد مطالعه شده است.

کلمات کلیدی: اصل توسیع، برنامه‌ریزی ریاضی، ضریب همبستگی، عدد فازی، T-نرم دراستیک.

۱ مقدمه

بررسی ساختار همبستگی و ارتباط میان متغیرها یکی از مسایل مورد توجه در آمار است. در بیشتر تحلیل‌های آماری یافتن و تعیین نوع رابطه بین متغیرهای تصادفی و مشاهدات حاصل از آن‌ها همواره مورد توجه محققین بوده است. تاکنون اندازه‌های متنوعی تحت عنوان ضرایب همبستگی بدین منظور معرفی شده و روش‌های مختلفی برای بررسی این ارتباط وجود دارد که بسته به ویژگی داده‌های مورد بررسی و هدف تحقیق، می‌توان از هر کدام از آن‌ها استفاده نمود. بر حسب این‌که متغیرها کمی یا کیفی باشند، ضریب همبستگی شاخصی است که میزان رابطه بین متغیرها را نشان می‌دهد. از مهمترین ضرایب همبستگی که تاکنون معرفی شده‌اند می‌توان به ضریب همبستگی پیرسن، ضریب همبستگی رتبه ای اسپیرمن، C ضریب همبستگی توافق پیرسون یا ضریب توافق، ضریب همبستگی کرامر، ضریب همبستگی گاما، ضریب همبستگی فی اشاره کرد. يك ملاك مناسب براي تعيين همبستگي دو متغير كمي ضريب همبستگي پيرسن مي‌باشد كه توسط پيرسون معرفی شد كه از آن زمان تا كنون به

[۲] رضا زارعی, r-zarei@guilan.ac.ir rezazarei.r@gmail.com

منظور تعیین میزان رابطه، نوع و جهت رابطه‌ی بین دو متغیر فاصله‌ای یا نسبی و یا یک متغیر فاصله‌ای و یک متغیر نسبی به کار برده می‌شود.
با وجود کاربرد فراوان ضرایب همبستگی اشاره شده در بسیاری از مسایل واقعی به سبب شرایط حاکم بر مسئله خطای ماشین و یا خطای انسانی امکان اندازه‌گیری دقیق مشاهدات در مسئله تحت بررسی وجود ندارد. در این شرایط داده‌ها به صورت نادقیق گردآوری و ثبت شده، بنابراین استفاده از روش‌های دقیق برای تجزیه و تحلیل این گونه مسائل با خطاهای جبران‌ناپذیری همراه خواهد بود. از این‌رو نیازمندیم تا روش‌های جایگزین به منظور تجزیه و تحلیل همبستگی بین دو مجموعه از داده‌ها یا دو متغیر تصادفی در شرایطی که داده‌های گردآوری شده، نادقیق هستند ارائه دهیم.

در این مقاله با استفاده از ابزارهای مفیدی که نظریه مجموعه‌های فازی برای صورت‌بندی مفاهیم نادقیق در اختیار ما قرار داده است به بررسی ضریب همبستگی تحت شرایط نادقیق خواهیم پرداخت.

۲ ضریب همبستگی بین اعداد فازی

با فرض این‌که $(X_۱, Y_۱), (X_۲, Y_۲), ..., (X_n, Y_n)$ یک نمونه تصادفی از متغیرهای دو جامعه باشند، ضریب همبستگی پیرسن در این مشاهدات به صورت زیر تعریف می‌شود

$$\rho_{X,Y} = \frac{\sum_{i=۱}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=۱}^{n}(X_i - \bar{X})^۲ \sum_{i=۱}^{n}(Y_i - \bar{Y})^۲}}. \qquad (۱)$$

علاوه بر آن درصورتی‌که مقادیر مشاهده شده این نمونه تصادفی را با $(x_۱, y_۱), (x_۲, y_۲), ..., (x_n, y_n)$ نشان دهیم، برآورد ضریب همبستگی آن را با $r_{X,Y}$ نشان می‌دهیم که به صورت زیر محاسبه می‌شود

$$r_{X,Y} = \hat{\rho}_{X,Y} = \frac{\sum_{i=۱}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=۱}^{n}(x_i - \bar{x})^۲ \sum_{i=۱}^{n}(y_i - \bar{y})^۲}}. \qquad (۲)$$

با توجه به رابطه (۲)، به راحتی می‌توان نشان داد که $-۱ \leq r_{X,Y} \leq +۱$.

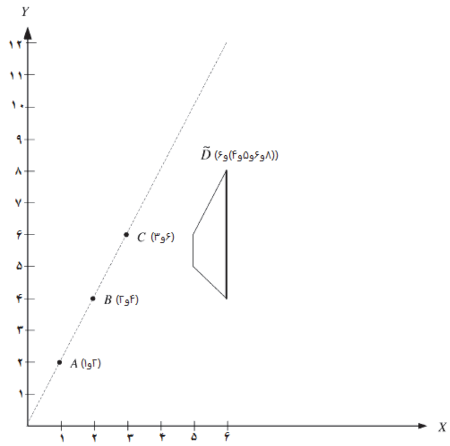
۱.۲ حالت خاص

در این بخش هدف محاسبه ضریب همبستگی بین زوج مشاهدات بر اساس یک نمونه تصادفی n تایی است که در بین آن‌ها یکی از داده‌ها دارای ابهام است و به‌صورت نادقیق مشاهده و ثبت شده است. به منظور سهولت در محاسبات فرض می‌کنیم که نمونه از چهار زوج مشاهده گردآوری و به‌صورت زیر ثبت شده‌اند (شکل ۱ را ببینید)

$$A = (۱, ۲), \quad B = (۲, ۴), \quad C = (۳, ۶), D = (۶, \tilde{Y}_D),$$

که در آن $\tilde{Y}_D = (۴, ۵, ۶, ۸)$ یک عدد فازی ذوزنقه‌ای با تابع عضویت زیر می‌باشد

$$\mu_{\tilde{Y}_D(y)} = \begin{cases} y - ۴ & ۴ \leqslant y \leqslant ۵, \\ ۱ & ۵ \leqslant y \leqslant ۶, \\ \frac{۸-y}{۲} & ۶ \leqslant y \leqslant ۸. \end{cases}$$



شکل ۱: *نمودار زوج مشاهدات* A, B, C, D

برای تعیین مقدار ضریب همبستگی این مشاهدات نیازمند محاسبه $\bar{X}$ و $\bar{Y}$ می‌باشیم. واضح است که مقدار $\bar{X} = ۳$ عددی دقیق است، اما با توجه به این‌که یکی از مشاهدات مربوط به Y_i ها نادقیق است، میانگین مشاهدات ثبت شده مربوط به Y_i ها به‌صورت زیر می‌باشد

$$\tilde{\bar{Y}}_D = ۳ + \frac{\tilde{Y}_D}{۴}.$$

حال با استفاده از رابطه (۲) ضریب همبستگی به‌صورت زیر محاسبه می‌شود

$$\tilde{r}_{X,Y} = \frac{\sum\limits_{i=A,B,C,D} (X_i - ۳)(Y_i - ۳ - \frac{\tilde{Y}_D}{۴})}{\sqrt{\sum\limits_{i=A,B,C,D} (X_i - ۳)^۲ \sum\limits_{i=A,B,C,D} (Y_i - ۳ - \frac{\tilde{Y}_D}{۴})^۲}}. \quad (۳)$$

علی‌رغم این‌که عبارت به‌دست آمده در رابطه (۳) یک کمیت فازی است، لزومی ندارد که شکل تابع عضویت آن با تابع $\tilde{Y}_D$ (به عنوان تنها مشاهده‌ای که ابهام ایجاد کرده است) یکسان باشد، چرا که رابطه (۲) یک ترکیب غیر خطی از $\tilde{Y}_D$ است. برای به‌دست آوردن تابع عضویت $\tilde{r}_{X,Y}$ از عملگرهای محاسباتی که بر اساس α- برش اعداد فازی پایه‌گذاری شده‌اند، استفاده می‌کنیم. به این منظور ابتدا α- برش عدد فازی ذوزنقه‌ای $\tilde{Y}_D$، $\tilde{r}_{X,Y}$ را به‌ترتیب به صورت زیر محاسبه می‌کنیم

$$\begin{aligned}
(\tilde{Y}_D)_\alpha &= [(\tilde{Y}_D)^L_\alpha, (\tilde{Y}_D)^U_\alpha] \\
&= [\min_y\{y \in [۴, ۸] | \mu_{\tilde{Y}_D(y)} \geqslant \alpha\}, \max_y\{y \in [۴, ۸] | \mu_{\tilde{Y}_D(y)} \geqslant \alpha\}], \\
(\tilde{r}_{X,Y})_\alpha &= [(\tilde{r}_{X,Y})^L_\alpha, (\tilde{r}_{X,Y})^U_\alpha] \\
&= [\min_r\{r \in [-۱, ۱] | \mu_{\tilde{r}_{X,Y}(r)} \geqslant \alpha\}, \max_r\{r \in [-۱, ۱] | \mu_{\tilde{r}_{X,Y}(r)} \geqslant \alpha\}].
\end{aligned}$$

حال برای هر مقدار دلخواه $۰ \leqslant \alpha \leqslant ۱$ می‌توانیم α-برش های $(\tilde{Y_D})^L_\alpha, (\tilde{Y_D})^U_\alpha$ را که مقادیری دقیق هستند محاسبه کرد و با جایگذاری آن‌ها در رابطه (۳)، α-های برش $(\tilde{r}_{XY})^L_\alpha, (\tilde{r}_{XY})^U_\alpha$ را تعیین کرد.

برای تعیین تابع عضویت ضریب همبستگی $\tilde{r}$، ابتدا α-برش های عدد فازی ذوزنقه‌ای $\tilde{Y_D}$ را به صورت زیر تعیین می‌کنیم

$$(\tilde{Y_D})_\alpha = [۴ + \alpha, ۸ - ۲\alpha],$$

با قرار دادن این مقادیر در رابطه (۳) α-های برش $\tilde{r}$ به صورت زیر به‌دست خواهد آمد

$$(\tilde{r}_{X,Y})^L_\alpha = \frac{(۴ + ۳\alpha)}{\sqrt{۱۰/۵\alpha^۲ + ۱۱۲}}, \qquad (\tilde{r}_{X,Y})^U_\alpha = \frac{(۱۶ - ۶\alpha)}{\sqrt{۴۲\alpha^۲ - ۱۶۸\alpha + ۲۸۰}}. \quad (۴)$$

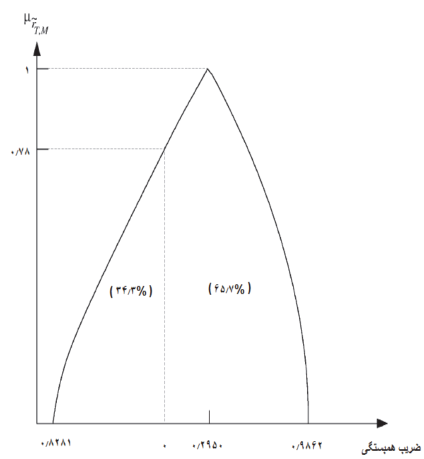
حال برای تعیین تابع عضویت $\tilde{r}$ از دو رابطه به‌دست آمده α- برش‌های آن استفاده می‌کنیم. با اندکی محاسبات جبری تابع عضویت $\tilde{r}$ به‌صورت زیر به‌دست می‌آید

$$\mu_{\tilde{r}_{XY}(r)} = \begin{cases} \frac{-24+28\sqrt{6r^2-6r^4}}{18-21r^2} & \frac{1}{\sqrt{7}} \leqslant r \leqslant \sqrt{\frac{2}{5}}\,, \\ 1 & \sqrt{\frac{2}{5}} \leqslant r \leqslant \frac{10}{\sqrt{154}}\,, \\ \frac{48-42r^2-14\sqrt{6r^2-6r^4}}{18-21r^2} & \frac{10}{\sqrt{154}} \leqslant r \leqslant \frac{8}{\sqrt{70}}. \end{cases} \tag{5}$$

توجه داشته باشید که با توجه به‌اینکه $\tilde{Y}_D$ (تنها مشاهده دارای ابهام در مجموعه مشاهدات است) از نوع ذوزنقه‌ای بوده، برای تعیین مقادیر ابتدایی، وسط و انتهایی به‌صورت زیر عمل می‌کنیم. دامنه سمت چپ و راست عدد فازی ذوزنقه‌ای به‌ترتیب عبارتند از

$$[(\tilde{r}_{X,Y})^L_{\alpha=0}, (\tilde{r}_{X,Y})^L_{\alpha=1}] = [\frac{1}{\sqrt{7}}, \sqrt{\frac{2}{5}}],$$
$$[(\tilde{r}_{X,Y})^U_{\alpha=0}, (\tilde{r}_{X,Y})^U_{\alpha=1}] = [\frac{10}{\sqrt{154}}, \frac{8}{\sqrt{70}}].$$

همان‌طور که انتظار می‌رفت بر اساس این رابطه کمترین مقدار ضریب همبستگی برابر با۰/۳۷۸۰ و بیشترین مقدار ضریب همبستگی ۰/۹۵۶۲ است که دقیقاً با مقادیر به‌دست آمده از رابطه (۴) مطابقت دارد. با این وجود باید توجه نمود که هر چند تنها مشاهده نادقیق در مجموعه مشاهدات موجود، یک عدد فازی ذوزنقه‌ای است، اما شکل حاصل از ضریب همبستگی فازی به‌دست آمده (شکل (۲)راببینید) یک عدد فازی ذوزنقه‌ای نیست.



شکل ۲: *نمودار ضریب همبستگی بین مشاهدات* A, B, C, D

۲.۲ حالت کلی

فرض کنید $(\tilde{X}_1, \tilde{Y}_1), (\tilde{X}_2, \tilde{Y}_2), \ldots, (\tilde{X}_n, \tilde{Y}_n)$ نمونه‌ای از مشاهدات نادقیق جمع‌آوری شده دو متغیر تحت بررسی باشند که هدف محاسبه و تحلیل ضریب همبستگی بین آنها است. مشابه با رابطه (۲)، ضریب همبستگی این مشاهدات نادقیق به‌صورت زیر قابل بازنویسی است

$$\tilde{r}_{X,Y} = \frac{\sum_{i=1}^{n} (\tilde{X}_i - \frac{\sum_{i=1}^{n} \tilde{X}_i}{n})(\tilde{Y}_i - \frac{\sum_{i=1}^{n} \tilde{Y}_i}{n})}{\sqrt{\sum_{i=1}^{n} (\tilde{X}_i - \frac{\sum_{i=1}^{n} \tilde{X}_i}{n})^2 \sum_{i=1}^{n} (\tilde{Y}_i - \frac{\sum_{i=1}^{n} \tilde{Y}_i}{n})^2}}. \tag{6}$$

برخلاف بخش قبل که تنها یک داده نادقیق وجود داشت رابطه (۶) تابعی از اعداد فازی مختلف است، که محاسبه مقدار آن نیازمند انجام محاسبات جبری فراوان و پیچیده روی این اعداد است، که به سادگی نتیجه بخش نخواهد بود. در چنین شرایطی استفاده از اصل گسترش کار را بسیار ساده‌تر خواهد کرد. بر اساس این اصل تابع عضویت ضریب همبستگی این مشاهدات به‌صورت زیر بیان می‌شود.

$$\mu_{\tilde{r}_{X,Y}(r)} = \sup_{X,Y} min\{\mu_{\tilde{X}_i(x_i)}, \mu_{\tilde{Y}_i(y_i)}, \forall i | r = r_{XY}\}, \qquad (۷)$$

که در آن $r_{X,Y}$ در رابطه (۲)تعریف شده است.

با توجه به رابطه (۷)، α-برش های $\tilde{X}_i$ و $\tilde{Y}_i$ را (برای $0 \leqslant \alpha \leqslant 1$ ، $i = 1, 2, ..., n$) به‌صورت زیر در نظر بگیرید

$$\begin{aligned} (\tilde{X}_i)_\alpha &= [(\tilde{X}_i)^L_\alpha, (\tilde{X}_i)^U_\alpha] = [\min_x\{x_i \in X | \mu_{\tilde{X}_i(x_i)} \geqslant \alpha\}, \max_x\{x_i \in X | \mu_{\tilde{X}_i(x_i)} \geqslant \alpha\}], \\ (\tilde{Y}_i)_\alpha &= [(\tilde{Y}_i)^L_\alpha, (\tilde{Y}_i)^U_\alpha] = [\min_y\{y_i \in Y | \mu_{\tilde{Y}_i(y_i)} \geqslant \alpha\}, \max_y\{y_i \in Y | \mu_{\tilde{Y}_i(y_i)} \geqslant \alpha\}]. \end{aligned}$$

برای محاسبه تابع عضویت ضریب همبستگی فازی $\tilde{r}_{X,Y}$ بایستی کران‌های پایین و بالای $\alpha-$برش این کمیت فازی را محاسبه کنیم که با توجه به توضیحات ارائه شده می‌توان این دو کران را در قالب دو برنامه‌ریزی غیرخطی به صورت زیر مدل بندی کرد

$$\begin{aligned} (\tilde{r}_{X,Y})^L_\alpha &= \frac{\min \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}, \\ s.t. \quad & (\tilde{X}_i)^L_\alpha \leqslant x_i \leqslant (\tilde{X}_i)^U_\alpha, \quad \forall i, \qquad (\tilde{Y}_i)^L_\alpha \leqslant y_i \leqslant (\tilde{Y}_i)^U_\alpha, \quad \forall i, \end{aligned} \qquad (۸)$$

$$\begin{aligned} (\tilde{r}_{X,Y})^U_\alpha &= \frac{\max \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}. \\ s.t. \quad & (\tilde{X}_i)^L_\alpha \leqslant x_i \leqslant (\tilde{X}_i)^U_\alpha, \quad \forall i, \qquad (\tilde{Y}_i)^L_\alpha \leqslant y_i \leqslant (\tilde{Y}_i)^U_\alpha, \quad \forall i. \end{aligned} \qquad (۹)$$

برای حل این برنامه‌ریزی‌های غیرخطی از روش گرادیان و نرم‌افزار متلب استفاده خواهیم کرد. برای دو مقدار α_1 و α_2 دلخواه که $1 \leqslant \alpha_1 < \alpha_2 < 0$ هستند، ناحیه جواب‌های شدنی تعریف شده توسط α_1 در برنامه‌های ارائه شده رابطه (۸) کوچکتر از ناحیه تعریف شده توسط α_2 خواهد بود. در نتیجه α-برش‌های $\tilde{r}_{X,Y}$ در سطح α_1 در α-برش‌های ایجاد شده در سطح α_2 قرار خواهد گرفت و به این دلیل تابع عضویت ضریب همبستگی $\tilde{r}_{X,Y}$ محدب خواهد شد. حال با داشتن شرط تحدب مورد بحث می‌توانیم پهنه‌های چپ و راست تابع عضویت $\tilde{r}_{X,Y}$ را با استفاده از α-برش‌های به‌دست آمده محاسبه کنیم و در نهایت به فرم تابع عضویت خواهیم رسید.

نکته قابل ذکر این است که با توجه به تابع عضویت در نظر گرفته‌شده برای داده‌های نادقیق اولیه، ممکن است فرم پهنه‌های چپ و راست تابع عضویت، آنقدر پیچیده به‌دست آیند که امکان پیدا کردن خود تابع عضویت وجود نداشته باشد. در این شرایط از روش‌های عددی برای محاسبه میزان عضویت استفاده خواهیم کرد.

با در نظر گرفتن $L(r)$ و $R(r)$ به عنوان پهناهای چپ و راست تابع عضویت، می‌توان فرم نهایی این تابع را از رابطه زیر به‌دست آورد

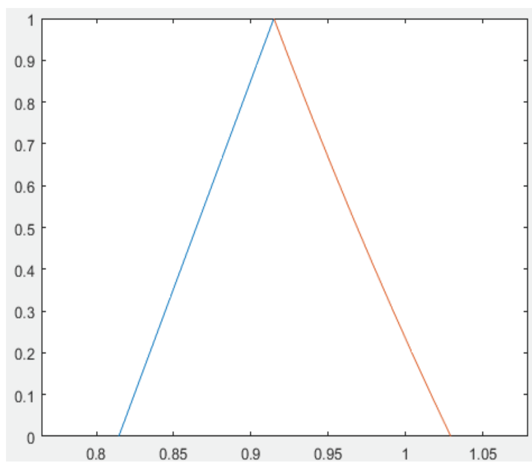
$$\mu_{\tilde{r}_{X,Y}(r)} = \begin{cases} L(r) & (\tilde{r}_{X,Y})^L_{\alpha=0} \leqslant r \leqslant (\tilde{r}_{X,Y})^L_{\alpha=1}, \\ 1 & (\tilde{r}_{X,Y})^L_{\alpha=1} \leqslant r \leqslant (\tilde{r}_{X,Y})^U_{\alpha=1}, \\ R(r) & (\tilde{r}_{X,Y})^U_{\alpha=1} \leqslant r \leqslant (\tilde{r}_{X,Y})^U_{\alpha=0}. \end{cases} \qquad (۱۰)$$

مثال ۱.۲. *رابطه بین شاخص‌های تکنولوژی و مدیریت همواره موضوعی مهم بوده و اندازه‌گیری آن مورد توجه محققین بسیاری قرار گرفته است. (۳) با استفاده از اطلاعات موجود راجع به این دو متغیر که غالباً به شکل متغیرهای زبانی هستند آن‌ها را در قالب اعداد فازی مدل‌بندی و ارائه نموده‌اند. بر این اساس، اطلاعات مربوط به سطوح مختلف تکنولوژی $(\tilde{T}_i)$ و مدیریت $(\tilde{M}_i)$ در یک نمونه ۱۵ تایی از شرکت‌های صنعتی کشور تایوان توسط این محققین گردآوری شد. بر اساس روش مورد مطالعه در این فصل و با استفاده از α-برش‌های موجود برای هر یک از سطوح مختلف برش، می‌توان یک برنامه‌ریزی غیرخطی تشکیل داد و با استفاده از نرم‌افزار متلب آن را حل کرد و جواب نهایی که همان α-برش مقدار ضریب همبستگی می‌باشد را به‌دست آورد. در پایان نتایج مربوط به هریک از سطوح مختلف α در جدول (۱) ارائه شده است. بر اساس مقادیر α- برش بالایی و پایینی $\tilde{r}_{\tilde{T},\tilde{M}}$ در سطوح مختلف، نمودار تقریبی ضریب همبستگی فازی بین مدیریت و تکنولوژی با استفاده از نرم*

جدول ۱: *مقادیرα- برش ضریب همبستگی فازی بین تکنولوژی و مدیریت شرکت‌های منتخب درنمونه تصادفی مثال(۱.۲)*

کران		$\alpha = ۰$	$\alpha = /۱$	$\alpha = ۰/۲$	$\alpha = ۰/۳$	$\alpha = ۰/۴$	$\alpha = ۰/۵$	$\alpha = ۰/۶$	$\alpha = ۰/۷$	$\alpha = ۰/۸$	$\alpha = ۰/۹$	$\alpha = ۱$
	L	-۰/۸۲۸۱	-۰/۷۷۴۸	-۰/۶۹۳۳	-۰/۵۹۴۷	-۰/۴۸۶۱	-۰/۳۶۹۶	-۰/۲۴۲۶	-۰/۱۰۹۱	۰/۰۲۶۷	۰/۱۶۴۹	۰/۲۹۵۰
	U	۰/۹۸۶۲	۰/۹۷۰۴	۰/۹۴۷۵	۰/۹۱۶۳	۰/۸۷۴۳	۰/۸۱۶۷	۰/۷۴۰۹	۰/۶۵۳۷	۰/۵۴۲۵	۰/۴۲۳۳	۰/۲۹۵۰

افزار متلب در شکل (۳) رسم شده است.



شکل ۳: *نمودار تقریبی تابع عضویت ضریب همبستگی فازی بین تکنولوژی و مدیریت در مثال (۱.۲)*

برای تحلیل همبستگی بین این دو متغیر بر اساس شاخص ضریب همبستگی معرفی شده، ابتدا به نکات زیر که می‌توان از نتایج حاصل استخراج کرد توجه می‌کنیم

الف) تکیه‌گاه $\tilde{r}_{\tilde{T},\tilde{M}}$ بسیار گسترده است و دامنه آن از ۰/۸۲۸۱ − تا ۰/۹۸۶۲ می‌باشد. بر اساس این بازه به‌دست آمده، می‌توان چنین نتیجه‌گیری کرد که اگر چه ضریب همبستگی در این شرایط کمیتی نادقیق است، اما نمی‌تواند کمتر از ۰/۸۲۸۱ − و بیشتر از ۰/۹۸۶۲ باشد.

ب) با فرض این‌که تمام داده‌های گزارش شده در این مثال اعدادی دقیق باشند $(\alpha = ۱)$ مقدار ضریب همبستگی بین دو متغیر تحت بررسی برابر با ۰/۲۹۵۰ خواهد بود. از طرفی به سادگی می‌توان دید که میزان عضویت $r = ۰/۲۹۵۰$ در مجموعه فازی به دست آمده برابر با ۱ است، علاوه بر آن می‌توانیم جهت استنباط دقیق تر میزان عضویت $r = ۰$ که مبنی بر ناهمبسته بودن دو متغیر است را نیز بررسی کنیم. بر این اساس میزان عضویت $r = ۰$ در مجموعه فازی به دست آمده برابر ۰/۷۸ است که مقدار قابل توجه است.

پ) با توجه به نمودار به‌دست آمده برای تابع عضویت $\tilde{r}_{\tilde{T},\tilde{M}}$ و با اندکی محاسبات، مساحت قرار گرفته در سمت چپ صفر برابر با ۳۴/۳ درصد و در سمت راست آن ۶۵/۷ درصد بوده است.

بر اساس سه مورد ذکر شده می‌توان در مجموع چنین نتیجه‌گیری کرد که ارتباط بین مدیریت و تکنولوژی در ۱۵ شرکت مورد بررسی نسبتاً ضعیف بوده است.

۳ رویکردی جدید برای ضریب همبستگی بین اعداد فازی تحت T- نرم دراستیک

همان‌طور که در بخش قبل گفته شد، نخستین رویکرد برای تعریف ضریب همبستگی بین اعداد فازی بر پایه استفاده از α-های اعداد فازی بود که منجر به حل یک برنامه‌ریزی ریاضی می‌شود. در این بخش با رویکردی متفاوت که بر پایه استفاده از T نرم‌ها بنا نهاده شده است، به بررسی تعریفی جدید برای ضریب همبستگی بین اعداد فازی می‌پردازیم.

۱.۳ عملیات جبری روی اعداد فازی بر اساس T-نرم دراستیک

در این بخش عملیات جبری چهارگانه شامل جمع، تفریق، ضرب، تقسیم بین اعداد فازی LR را بر اساس T-نرم دراستیک مورد مطالعه قرار می‌دهیم.

تعریف ۱.۳. *فرض کنید $I = [\circ, ۱]$. یک تابع دو متغیره به صورت $T : I \times I \to I$ را یک T- نرم* [۱] *گوییم اگر در خواص زیر صدق کند*

الف) $T(x, ۱) = x$.

ب) یکنوایی: $x_۱ \leqslant x_۲, y_۱ \leqslant y_۲ \Longrightarrow T(x_۱, y_۱) \leqslant T(x_۲, y_۲)$.

پ) جابه‌جایی: $T(x, y) = T(y, x)$.

ت) شرکت‌پذیری: $T[x, T(y, z)] = T[T(x, y), z]$.

T-نرم‌ها و T- همنرم ها (یا S-نرم‌ها) را نرم‌های مثلثی و همنرم‌های مثلثی نیز گویند. این دو رده از اندازه‌ها به خاطر نقشی که در نامساوی‌های مثلثی دارند به این نام خوانده می‌شوند، به دلیل ویژگی‌های بسیار جالب اخیراً بیشتر مورد توجه واقع شده‌اند. در بین عملگرهای T- نرم عملگر دراستیک خواص جالبی دارد و مورد توجه قرارگرفته است و به‌صورت زیر تعریف می‌شود

$$T_W(x, y) = \begin{cases} x & y = ۱\ , \\ y & x = ۱\ , \\ \circ & x, y \neq ۱\ . \end{cases}$$

فرض کنید $T = T_W$ ضعیف‌ترین T-نرم باشد و $\tilde{A} = (a, \alpha_A, \beta_A)_{LR}$، $\tilde{B} = (b, \alpha_b, \beta_B)_{LR}$ دو عدد فازی $L - R$ باشند. دراین‌صورت جمع بین این دو عدد فازی تحت T-نرم دراستیک عبارتست از

$$\tilde{A} \oplus \tilde{B} = (a, \alpha_A, \beta_A)_{LR} \oplus (b, \alpha_B, \beta_B)_{LR} = (a + b, max(\alpha_A, \alpha_B), max(\beta_A, \beta_B))_{LR}. \qquad (۱۱)$$

درصورتی‌که $L = R$ باشد. تفاضل این دو عدد فازی تحت T-نرم دراستیک یک عدد فازی LR به‌صورت زیر است

$$\begin{aligned} \tilde{A} \ominus \tilde{B} &= (a, \alpha_A, \beta_A)_{LR} \ominus (b, \alpha_B, \beta_B)_{LR} \\ &= (a, \alpha_A, \beta_A)_{LR} \oplus (-b, \beta_B, \alpha_B)_{LR} = (a - b, max(\alpha_A, \beta_B), max(\beta_A, \alpha_B))_{LR}. \end{aligned} \qquad (۱۲)$$

برای بررسی ضرب دو عدد فازی LR تحت T-نرم دراستیک بایستی حالت‌های مختلفی را برای مراکز دو عدد فازی تحت بررسی مورد مطالعه قرار دهیم که در ادامه به آن پرداخته شده‌است.

[۱] T-Norm

حالت اول: فرض کنید $a,b > ۰$ دراین‌صورت برای مقادیر $z \leqslant ab$ داریم

$$(\tilde{A}\otimes\tilde{B})(z) = \sup_{x.y=z} T_W(\tilde{A}(x),\tilde{B}(y)) = \max(\tilde{A}(\frac{z}{b}),\tilde{B}(\frac{z}{a})) = \max(L(\frac{a-\frac{z}{b}}{\alpha_A}),L(\frac{b-\frac{z}{a}}{\alpha_B})) = \max(L(\frac{ab-z}{\alpha_A b}),L(\frac{ab-z}{\alpha_B a})), \quad (۱۳)$$

بنابراین می‌توان نوشت

$$(\tilde{A}\otimes\tilde{B})(z) = \begin{cases} L[\frac{(ab-z)}{\max(\alpha_A b,\alpha_B a)}] & z \geqslant ab-\max(\alpha_A b,\alpha_B a)\ , \\ \\ ۰ & \text{.در غیر این صورت} \end{cases} \quad (۱۴)$$

به طور مشابه

$$(\tilde{A}\otimes\tilde{B})(z) = \begin{cases} R[\frac{(z-ab)}{\max(\beta_A b,\beta_B a)}] & z \leqslant ab-\max(\beta_A b,\beta_B a)\ , \\ \\ ۰ & \text{.در غیر این صورت} \end{cases} \quad (۱۵)$$

با جمع‌بندی دو عبارت به‌دست آمده در روابط (۱۴) و (۱۵) می‌توان ضرب دو عدد فازی LR را در حالت $a,b > ۰$ به صورت یک عدد فازی LR با تعریف زیر ارائه نمود

$$\tilde{A}\otimes\tilde{B} = (ab,\max(\alpha_A b,\alpha_B a),\max(\beta_A b,\beta_B a))_{LR}. \quad (۱۶)$$

به روش مشابه با آن‌چه که در حالت اول مورد بررسی قرار گرفت، حالت‌های ممکن دیگر برای a,b به عنوان مراکز دو عدد فازی تحت بررسی در ادامه آورده شده است.

حالت دوم: اگر $a,b < ۰$ آن‌گاه ضرب دو عدد فازی $\tilde{A}$ و $\tilde{B}$ یک عدد فازی LR با تعریف زیر می‌باشد

$$\tilde{A}\otimes\tilde{B} = (ab,\max(\beta_A b,\beta_B a),\max(\alpha_A b,\alpha_B a))_{RL}.$$

حالت سوم: فرض کنید $a = ۰, b > ۰$ دراین‌صورت حاصل ضرب دو عدد فازی $\tilde{A}$ و $\tilde{B}$ عدد فازی LR با تعریف زیر است

$$\tilde{A}\otimes\tilde{B} = (۰,\alpha_A b,\beta_A b)_{LR}.$$

حالت چهارم: برای حالتی که $a = ۰, b < ۰$ باشد، حاصل ضرب یک عدد فازی LR به‌صورت زیر است

$$\tilde{A}\otimes\tilde{B} = (۰,-\beta_A b,-\alpha_A b)_{RL}.$$

حالت پنجم: درصورتی‌که $a = ۰, b = ۰$ باشند آنگاه

$$\tilde{A}\otimes\tilde{B} = (۰,۰,۰)_{LR} = I_{\{۰\}}(x).$$

که در آن

$$I_{\{۰\}}(x) = \begin{cases} ۱ & x = ۰ \\ ۰ & \text{در غیر این صورت} \end{cases} .$$

حالت ششم: در این حالت فرض می‌کنیم $a < ۰, b > ۰$ باشند و علاوه بر آن $L = R$ باشد. دراین‌صورت

$$\tilde{A} \otimes \tilde{B} = (ab, \max(\alpha_A b, -\beta_B a), \max(\beta_A b, -\alpha_B a))_{RR}.$$

درصورتی‌که $\tilde{A}$ و $\tilde{B}$ اعداد فازی متقارن باشند آنگاه حاصل ضرب این دو عدد فازی متقارن یک عدد فازی LR با تعریف زیر می‌باشد

$$(\tilde{A} \otimes \tilde{B})(z) = (ab, \max(\alpha_A \mid b \mid, \alpha_B \mid a \mid), \max(\alpha_A \mid b \mid, \alpha_B \mid a \mid))_{LL}.$$

برای به‌دست آوردن تابع عضویت ضریب همبستگی فازی $\tilde{r}_{\tilde{X},\tilde{Y}}$ نیازمند محاسبه تابع عضویت آماره‌های میانگین و انحراف معیار می‌باشیم. در ادامه بدون کاستن از کلیت مسئله فرض می‌کنیم که $\tilde{A}, \tilde{B}$ دو عدد فازی مثلثی متقارن باشند. در این بخش فرض می‌کنیم که $n \odot \tilde{A} = \tilde{A} + ... + \tilde{A}$ باشد. در این‌صورت اگر $\tilde{A} = (a, \alpha)$ باشد آنگاه $n \odot \tilde{A} = \tilde{A} + ... + \tilde{A} = (na, \alpha)$. بنابراین برای هر مقدار غیر صفر مانند t داریم $t \odot \tilde{A} = (ta, \alpha)$.

اکنون فرض کنید که n مشاهده فازی از متغیرهای X،Y در اختیار داشته باشیم و هدف محاسبه ضریب همبستگی این زوج مشاهدات نادقیق باشد. در چنین شرایطی رابطه (۲) را بر حسب نمادگذاری‌های جدید انجام شده، به‌صورت زیر بازنویسی می‌کنیم (۲)

$$\tilde{r}_{\tilde{X},\tilde{Y}} = \frac{\sum_{j=۱}^{n} (\tilde{X}_j - \frac{۱}{n} \odot \sum_{j=۱}^{n} \tilde{X}_j)(\tilde{Y}_j - \frac{۱}{n} \odot \sum_{j=۱}^{n} \tilde{Y}_j)}{\sqrt{\sum_{j=۱}^{n} (\tilde{X}_j - \frac{۱}{n} \odot \sum_{j=۱}^{n} \tilde{X}_j)^۲ \sum_{j=۱}^{n} (\tilde{Y}_j - \frac{۱}{n} \odot \sum_{j=۱}^{n} \tilde{Y}_j)^۲}}. \qquad (۱۷)$$

فرض کنید $\tilde{X}_j = (x_j, r_j)_L$، $\tilde{Y}_j = (y_j, \delta_j)_L$ نمونه‌های تصادفی از جوامع تحت بررسی می‌باشند که اعداد فازی متقارن در نظر گرفته شده‌اند. $(j = ۱, ..., n)$
با توجه به عملگرهای جبری تعریف شده بر اساس T- نرم دراستیک و نمادگذاری‌های انجام شده می‌توان نوشت

$$\tilde{\bar{X}} = \frac{۱}{n} \odot \sum_{J=۱}^{n} \tilde{X}_j = (\frac{۱}{n} \sum_{J=۱}^{n} x_j, \max_{۱ \leqslant j \leqslant n} r_j)_L, \qquad \tilde{\bar{Y}} = \frac{۱}{n} \odot \sum_{J=۱}^{n} \tilde{Y}_j = (\frac{۱}{n} \sum_{J=۱}^{n} y_j, \max_{۱ \leqslant j \leqslant n} \delta_j)_L$$

دراین‌صورت با استفاده از رابطه (۱۶) خواهیم داشت

$$(\tilde{X}_j - \tilde{\bar{X}})(\tilde{Y}_j - \tilde{\bar{Y}}) = ((x_j - \frac{۱}{n} \sum_{J=۱}^{n} x_j)(y_j - \frac{۱}{n} \sum_{J=۱}^{n} y_j), \max(\mid x_j - \frac{۱}{n} \sum_{J=۱}^{n} x_j \mid \max_{۱ \leqslant j \leqslant n} \delta_j, \mid y_j - \frac{۱}{n} \sum_{J=۱}^{n} y_j \mid \max_{۱ \leqslant j \leqslant n} r_j))_L.$$

با استفاده رابطه (۱۱) داریم

$$\sum_{j=۱}^{n} (\tilde{X}_j - \frac{۱}{n} \odot \sum_{j=۱}^{n} \tilde{X}_j)(\tilde{Y}_j - \frac{۱}{n} \odot \sum_{j=۱}^{n} \tilde{Y}_j) = \qquad (۱۸)$$
$$(\sum_{j=۱}^{n} (x_j - \frac{۱}{n} \sum_{k=۱}^{n} x_k)(y_j - \frac{۱}{n} \sum_{k=۱}^{n} y_k), \max_{۱ \leqslant j \leqslant n} (\mid x_j - \frac{۱}{n} \sum_{k=۱}^{n} x_k \mid \max_{۱ \leqslant k \leqslant n} \delta_k, \mid y_j - \frac{۱}{n} \sum_{k=۱}^{n} y_k \mid \max_{۱ \leqslant k \leqslant n} r_k))_L.$$

از طرفی به‌طور مشابه می‌توان نوشت

$$\sum_{J=۱}^{n} (\tilde{X}_j - \frac{۱}{n} \odot \sum_{J=۱}^{n} \tilde{X}_j)^۲ = (\sum_{J=۱}^{n} (x_j - \frac{۱}{n} \sum_{k=۱}^{n} x_k)^۲, \max_{۱ \leqslant j \leqslant n} \mid x_j - \frac{۱}{n} \sum_{k=۱}^{n} x_k \mid \max_{۱ \leqslant k \leqslant n} r_k)_L$$

بنابراین

$$\sqrt{\sum_{J=1}^{n}(\tilde{X}_j - \frac{1}{n}\odot\sum_{k=1}^{n}\tilde{X}_k)^2\sum_{J=1}^{n}(\tilde{Y}_j - \frac{1}{n}\odot\sum_{k=1}^{n}\tilde{Y}_k)^2} \tag{19}$$

$$= (\sqrt{\sum_{J=1}^{n}(x_j - \frac{1}{n}\sum_{k=1}^{n}x_k)^2\sum_{J=1}^{n}(y_j - \frac{1}{n}\sum_{k=1}^{n}y_k)^2}$$

$$, \frac{1}{\sqrt{\sum_{J=1}^{n}(x_j - \frac{1}{n}\sum_{k=1}^{n}x_k)^2\sum_{J=1}^{n}(y_j - \frac{1}{n}\sum_{k=1}^{n}y_k)^2}}$$

$$\times \max\{\sum_{J=1}^{n}(x_j - \frac{1}{n}\sum_{k=1}^{n}x_k)^2 \max_{1\leqslant j\leqslant n} \mid y_j - \frac{1}{n}\sum_{k=1}^{n}y_k \mid \max_{1\leqslant k\leqslant n}\delta_k,$$

$$\sum_{J=1}^{n}(y_j - \frac{1}{n}\sum_{k=1}^{n}y_k)^2 \max_{1\leqslant j\leqslant n} \mid x_j - \frac{1}{n}\sum_{k=1}^{n}x_k \mid \max_{1\leqslant k\leqslant n} r_k\})_L.$$

در نهایت می‌توان تابع عضویت ضریب همبستگی فازی $\tilde{r}$ را با استفاده از روابط به‌دست آمده برای تقسیم اعداد فازی تحت T- نرم دراستیک محاسبه نمود.

مثال ۲.۳. *یک مجموعه داده شامل یک ۸ زوج مشاهده از اعداد فازی مثلثی متقارن که در جدول (؟؟) آورده شده‌اند را در نظر بگیرید (۱)*

جدول ۲: *مجموعه داده شامل ۸ جفت از اعداد فازی مثلثی متقارن درمثال(۲.۳)*

i	$\tilde{X}_i$	$\tilde{Y}_i$
۱	(۲/۰،۰/۵)	(۴/۰،۰/۵)
۲	(۳/۵،۰/۵)	(۵/۵،۰/۵)
۳	(۵/۵،۱/۰)	(۷/۵،۱/۰)
۴	(۷/۰،۰/۵)	(۶/۵،۰/۵)
۵	(۸/۵،۰/۵)	(۸/۵،۵/۰)
۶	(۱۰/۵،۱/۰)	(۸/۰،۱/۰)
۷	(۱۱/۰،۰/۵)	(۱۰/۵،۰/۵)
۸	(۱۲/۵،۰/۵)	(۹/۵،۰/۵)

برای محاسبه ضریب همبستگی با استفاده از رابطه (؟؟) می‌توان نوشت

$$\sum_{J=1}^{n}(\tilde{X}_j - \tilde{\bar{X}})(\tilde{Y}_j - \tilde{\bar{Y}}) = (5/7500, 5/5625)$$

بنابراین

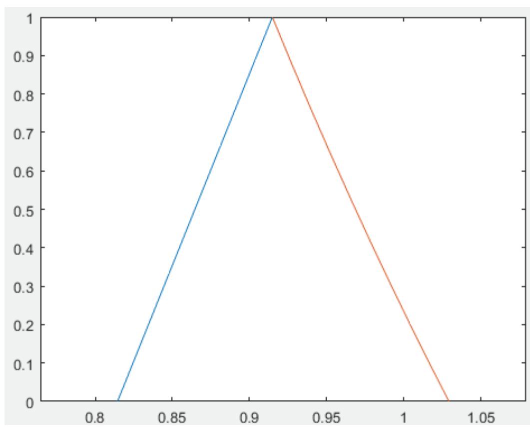
$$\sqrt{\sum_{J=1}^{n}(\tilde{X}_j - \tilde{\bar{X}})^2\sum_{J=1}^{n}(\tilde{Y}_j - \tilde{\bar{Y}})^2} = (55/4810, 6/1646)$$

دراین‌صورت ضریب همبستگی اعداد فازی برابر است

$$\tilde{r}_{\tilde{X},\tilde{Y}}=\frac{(۵/۷۵۰۰,۵/۵۶۲۵)}{(۵۵/۴۸۱۰,۶/۱۶۴۶)}=\begin{cases}۱-\frac{۰/۹۱۴۷-z}{۰/۰۱۸۰\times max(۵/۵۶۲۵,۶/۱۶۴۶\times z)} & ۰/۸۱۴۵\leqslant z\leqslant ۰/۹۱۴۷\,,\\ ۱-\frac{z-۰/۹۱۴۷}{۰/۰۱۸۰\times max(۵/۵۶۲۵,۶/۱۶۴۶\times z)} & ۰/۹۱۴۷\leqslant z\leqslant ۱/۰۲۹۱\,,\\ ۰ & \text{در غیر این صورت}.\end{cases} \tag{۲۰}$$

$$=\begin{cases}۱-\frac{۰/۹۱۴۷-z}{۰/۱۰۰} & ۰/۸۱۴۵\leqslant z\leqslant ۰/۹۰۲۳\,,\\ ۱-\frac{۰/۹۱۴۷-z}{۰/۱۱۱\times z} & ۰/۹۰۲۳\leqslant z\leqslant ۰/۹۱۴۷\,,\\ ۱-\frac{z-۰/۹۱۴۷}{۰/۱۱۱\times z} & ۰/۹۱۴۷\leqslant z\leqslant ۱/۰۲۹۱\,,\\ ۰ & \text{در غیر این صورت}.\end{cases} \tag{۲۱}$$

نمودار مربوط به تابع عضویت ضریب همبستگی فازی مثلثی متقارن در شکل زیر نمایش داده شده است.



شکل ۴: نمودار تقریبی تابع عضویت ضریب همبستگی فازی مثلثی متقارن در مثال(۲.۳)

مراجع

[1] Chachi, J., Taheri, S.M., (2016), Multiple Fuzzy Regression Model for Fuzzy Input Output Data, Iranian Journal of Fuzzy Systems, Vol 13,pp 63-78.

[2] Hong D. H.,(2006), "Fuzzy measures for a correlation coefficient of fuzzy numbers under T_W (the weakest $t-$norm)-based fuzzy arithmetic operations", Information Sciences, Vol 176, pp 150-160.

[3] Liu, S. T., Kao, C.,(2002), "Fuzzy measures for correlation coeffcient of fuzzy numbers", Fuzzy Sets and Systems, Vol 128, pp 267–275.