بررسی میزان قابلیت رگرسیون فازی هیبرید بر مبنای مقایسه با سایر روش‌های رگرسیونی

سیده منا احسانی جوکندان [1]، بهروز فتحی واجارگاه [2]

تاریخ دریافت: ۹۷/۱۱/۲۲
تاریخ پذیرش: ۹۹/۳/۱

چکیده:
در این مقاله تفاوت میان رگرسیون کلاسیک و رگرسیون فازی موردبحث و بررسی قرارگرفته است. در رگرسیون فازی داده‌های غیرفازی و فازی را می‌توان برای مدل‌بندی استفاده کرد. درحالی‌که در رگرسیون کلاسیک فقط از داده‌های غیرفازی استفاده می‌شود. هدف این تحقیق بررسی روش رگرسیون امکانی، روش رگرسیون کمترین مربعات مبتنی بر رگرسیون امکانی و روش هیبرید رگرسیون خطی کمترین مربعات بر اساس حساب فازی وزنی برای ورودی غیرفازی و خروجی فازی با استفاده از اعداد فازی مثلثی متقارن می‌باشد. در ادامه اندازه قابلیت اطمینان، فاصله اطمینان و معیار نیکویی برازش برای انتخاب مدل بهینه ارائه‌شده است. در پایان با ارائه مثال‌هایی رفتار روش‌های مطرح‌شده را موردبررسی قرار داده و بهینگی مدل هیبرید رگرسیون کمترین مربعات خطی فازی نشان داده شد.

واژه‌های کلیدی: رگرسیون هیبرید، اندازه قابلیت اطمینان، رگرسیون امکانی، کمترین مربعات فازی، حساب فازی وزنی.

۱ مقدمه

رگرسیون خطی کلاسیک یک ابزار آماری است که در کلیه علوم استفاده می‌شود. تحلیل رگرسیونی یک روش آماری برای مدل‌بندی رابطه بین متغیرها است. در رگرسیون کلاسیک فرض می‌شود که متغیرها و مشاهدات مربوط به آن‌ها دقیق باشند و درباره جملات خطا و توزیع احتمالی آن‌ها مانند نرمال بودن، ناهمبسته بودن، ثبات واریانس و... مفروضاتی در نظر گرفته می‌شود. حال امکان دارد یک یا چند فرض از مفروضات بالا برقرار نباشد، در این صورت می‌توان رگرسیون فازی را جایگزین رگرسیون کلاسیک نمود [۲]. رگرسیون فازی اولین بار توسط تاناکا و همکاران در سال ۱۹۸۲ معرفی شد. تاناکا با استفاده از یک شاخص و اعداد فازی مثلثی ضرایب رگرسیونی را برآورد نمود که با حداقل کردن میزان ابهام مقدار برآورد شده انجام می‌شود. این روش به رگرسیون امکانی مرسوم است. این‌گونه مسائل را می‌توان از طریق روش‌های برنامه‌ریزی خطی حل نمود و مدل رگرسیونی بهینه را به دست آورد. کلمینس و دایموند روش رگرسیون فازی کمترین مربعات را ارائه کردند [۶] و [۱۰]. اساس این روش استفاده از یک فاصله روی مجموعه‌های فازی بود، که با استفاده از آن مجموع توان‌های دوم مقادیر خروجی فازی مشاهده‌شده و مقادیر برآورد شده فازی کمینه می‌شوند. تا به امروز افراد بسیاری حالت‌های مختلف رگرسیون امکانی و رگرسیون کمترین مربعات را بررسی کردند. [۱]. ساویک و پدریچ ترکیبی از رگرسیون کمترین مربعات فازی و معیار کمینه سازی میزان فازی بودن، ارائه دادند . [۱۵] در ادامه یک روش هیبریدی از رگرسیون کمترین مربعات بر اساس حساب وزنی فازی توسط چانگ در سال ۲۰۰۱ گسترش یافت [۷]. با توجه برخی اشکالات موجود در روش‌های معمول منطق فازی رویکردهای جدیدی مانند حساب فازی وزنی ارائه‌شده است. سپس توتماز و برومان نشان دادند می‌توان از فاصله اطمینان برای محاسبه ضرایب در روش هیبرید خطی به عنوان روشی جایگزین برآورد نقطه‌ای پارامترها به برآورد فاصله‌ای برای ارزیابی میزان دقت آن استفاده نمود [۵] و [۱۸].

۲ رگرسیون کلاسیک و فازی

در حالت کلی مدل‌های رگرسیونی به‌گونه‌ای هستند که می‌توان به کمک آن‌ها رابطه بین مجموعه‌ای از متغیرها را بررسی نمود که شامل متغیرهای وابسته و مستقل هستند. رگرسیون خطی کلاسیک آماری به صورت زیر می‌باشد:

$$Y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_k x_{ik} + \varepsilon_i, \qquad i = 1, \ldots, m. \qquad (1)$$

[1] دانشجوی کارشناسی ارشد، گروه آمار، دانشکده علوم ریاضی، دانشگاه گیلان
[2] استاد گروه آمار، دانشکده علوم ریاضی، دانشگاه گیلان

که در آن Y_i متغیر وابسته (پاسخ)، x_{ij} متغیرهای مستقل و β_j ضرایب (پارامترها) غیر فازی و ε_i خطای تصادفی غیرفازی با میانگین $E(\varepsilon_i) = 0$، واریانس $\sigma^2(\varepsilon_i) = \sigma^2$ و کوواریانس $cov(\varepsilon_i, \varepsilon_j) = 0 \forall i, j, i \neq j$.
در رگرسیون کلاسیک متغیرها و مشاهدات دقیق هستند و درباره خطاها و توزیع احتمالی آن‌ها مفروضاتی در نظر گرفته می‌شود. امکان دارد که در ساختن مدل با مشکلاتی مانند ناکافی بودن یا نادقیق بودن مشاهدات، ابهام درباره ارتباط متغیرهای وابسته و مستقل وجود داشته باشند، یکی از شیوه‌های جایگزین و مناسب رگرسیون کلاسیک، رگرسیون فازی است. در حالت کلی می‌توان رگرسیون فازی را به چهار حالت تقسیم‌بندی کرد [۳]:

۱) زمانی که ضرایب مدل رگرسیونی فازی در نظر گرفته می‌شود، این مدل را رگرسیون امکانی می‌نامند.

۲) زمانی که متغیرهای مستقل و وابسته (ورودی و خروجی) فازی باشند و ضرایب مدل غیرفازی باشند.

۳) زمانی که ضرایب و خروجی مدل فازی و متغیرهای ورودی غیرفازی باشند.

۴) زمانی که هم متغیرها و هم ضرایب مدل فازی باشند.

در تابع عضویت رگرسیون فازی می‌توان از اعداد فازی مثلثی، سهموی، ذوزنقه و نرمال استفاده کرد این عامل به تنوع رگرسیون فازی می‌افزاید.

دو روش مهم و اصلی رگرسیون فازی، رگرسیون امکانی و رگرسیون کمترین مربعات هستند که توسط افراد متعدد بسط داده‌شده است.

۱.۲ رگرسیون امکانی

رگرسیون امکانی یکی از روش‌های رگرسیون فازی است که نخستین بار توسط تاناکا و همکاران در سال ۱۹۸۲ مطرح شد [۱۷]. مدل رگرسیون خطی آن‌ها، مدل رگرسیونی با ضرایب فازی مثلثی، متغیرهای مستقل دقیق، متغیر پاسخ دقیق و فازی مثلثی است. با توجه به این روش ضرایب فازی، اعدادی هستند که می‌توان با اعداد بازه‌ای که دارای درجه عضویت هستند نمایش داد. در صورتی که ضرایب رگرسیونی اعداد فازی باشند، برآورد متغیر وابسته نیز عدد فازی است. مدل رگرسیونی چند چندگانه خطی به صورت (۲) می‌باشد.

$$\tilde{y} = \tilde{A}_0 + \tilde{A}_1 x_i + \cdots + \tilde{A}_p x_p. \tag{۲}$$

که در آن $\tilde{y}_i$ خروجی فازی (متغیر وابسته)، $\tilde{A}_0$ عرض از مبدأ فازی، $\tilde{A}_i$ ضرایب فازی و $x_1, \ldots, x_n$ ورودی غیرفازی (متغیر مستقل) است. در صورتی که ضرایب رگرسیونی اعداد فازی باشند، برآورد متغیر پاسخ نیز عدد فازی است. اما پاسخ‌های مشاهده‌شده را می‌توان به دو صورت غیرفازی و دقیق در نظر گرفت. در صورتی که ضرایب مدل رگرسیونی، اعداد فازی مثلثی متقارن باشند به صورت $\tilde{A}_i = (m_i, c_i)$ نشان داده می‌شود که در آن m_i مرکز فازی و c_iها پهناهای فازی هستند و تابع عضویت آن به صورت

$$\mu_{\tilde{A}_i}(y) = \begin{cases} 1 - \frac{m_i - m}{c_i} & m_i - c_i \leq y \leq m_i \\ 1 - \frac{m - m_i}{c_i} & m_i \leq y \leq m_i + c_i \\ 0 & o.w \end{cases} \tag{۳}$$

می‌باشد. خروجی فازی مشاهده‌شده به صورت $\tilde{Y}_i = (Y_i, e_i)$ می‌باشد که در آن Y_i مرکز و e_i پهنای فازی است و تابع عضویت آن به صورت

$$\mu_{\tilde{Y}_i}(y) = \begin{cases} 1 - \frac{m_i - y}{e_i} & m_i - e_i \leq y \leq m_i \\ 1 - \frac{y - m_i}{e_i} & m_i \leq y \leq m_i + e_i \end{cases} \tag{۴}$$

می‌باشد. در صورتی که مقدار مشاهده‌شده دقیق باشد، مقدار e_i صفر می‌شود.

پارامترهای فازی را باید طوری برآورد کرد که مدل بهترین برازش را به داده‌ها داشته باشد. ضرایب مدل باید طوری تعیین شوند که ابهام مدل کمینه شود و درجه عضویت خروجی مشاهده‌شده $\tilde{Y}$ برای تمامی مقادیر y_i دارای درجه عضویتی حداقل به بزرگی h باشد. برای به دست آوردن ضرایب فازی می‌توان از مسئله برنامه‌ریزی خطی زیر که تاناکا و همکاران آن را برای به دست آوردن پارامترهای مدل در حالت متغیر مستقل دقیق و خروجی فازی معرفی کردند [۱۷].

$$\begin{aligned} \min \quad & S = nc_0 + \sum_{j=1}^{n} c_j |x_{ij}| \qquad c_j \geq 0, \qquad (۵) \\ & m_0 + \sum_{j=1}^{n} (m_j x_{ij}) + (1-h)(c_0 + \sum_{j=1}^{n} c_j x_{ij}) \\ & \quad \geq Y_i + (1-h) e_i, \\ & m_0 + \sum_{j=1}^{n} (m_j x_{ij}) - (1-h)(c_0 + \sum_{j=1}^{n} c_j x_{ij}) \\ & \quad \leq Y_i - (1-h) e_i, \\ & c_j \geq 0, \quad i = 1, \ldots, m, \quad j = 1, \ldots, n. \end{aligned}$$

از هر مشاهده دو محدودیت به وجود می‌آید، با حل مسئله برنامه‌ریزی خطی (۵) تابع هدف S با توجه به $2n$ محدودیت به وجود آمده توسط n مشاهده کمینه می‌شود.

۲.۲ رگرسیون کمترین مربعات مبتنی بر رگرسیون امکانی

روش رگرسیون کمترین مربعات مبتنی بر رگرسیون امکانی توسط ساویک و پدریچ در سال ۱۹۹۱ مطرح شد [۱۵]. آن‌ها با این ادعا که مدل رگرسیونی تاناکا مرکزهای داده‌ها را در محاسبه ضرایب رگرسیونی به‌حساب نمی‌آورد، درحالی‌که این مرکزها بیش‌ترین درجه عضویت را دارند، در نتیجه از اهمیت بیشتری برخوردارند، روش خود را ارائه دادند. آن‌ها برای حل این مشکل با

به‌کارگیری از روش رگرسیون کمترین مربعات کلاسیک روش خود را در دو مرحله بیان کردند.

در مرحله اول با استفاده از داده‌های موجود مراکز پارامترهای مدل توسط رگرسیون کمترین مربعات کلاسیک همان‌طور که در بخش ۱.۲ توضیح داده شد، به دست می‌آیند. سپس پارامترهای به‌دست‌آمده از مدل رگرسیون کلاسیک را به عنوان مرکز پارامترهای مدل رگرسیون فازی در نظر گرفته و به عنوان ورودی‌های مرحله دوم بکار می‌روند.

در مرحله دوم پهناهای آن را با روش تاناکا با استفاده از مراکز برآورد شده توسط رگرسیون کمترین مربعات کلاسیک برآورد کردند. برای انجام روش ساویک و پدریچ، ابتدا باید روش رگرسیون کمترین مربعات کلین‌بوم و کوپر (۱۹۷۸) را بکار برد[۱۲]. برآورد مرکز پارامترها، از طریق رابطه (۶) به دست می‌آید.

$$m^* = (X'X)^{-1}X'Y. \tag{۶}$$

X ماتریس متغیرهای ورودی و بردار Y مرکز پاسخ‌های مشاهده‌شده فازی است. بعد از به دست آوردن m^* وارد مرحله دوم شده یعنی حل یک مسئله برنامه‌ریزی خطی به روش تاناکا و همکارانش اما به جای بردار m از بردار m^* که در مرحله یک به دست آمد، استفاده می‌شود. در این مسئله برنامه‌ریزی c_j متغیرهای مجهولی هستند که از حل مسئله برنامه‌ریزی به دست آیند [۸].

$$\begin{aligned}
\min \quad & S = nc_0 + \sum_{j=1}^{n} c_j|x_{ij}| \qquad c_j \geq 0, \\
& m_0^* + \sum_{j=1}^{n}(m_j^* x_{ij}) + (1-h)(c_0 + \sum_{j=1}^{n} c_j x_{ij}) \\
& \quad \geq Y_i + (1-h)e_i, \\
& m_0^* + \sum_{j=1}^{n}(m_j^* x_{ij}) - (1-h)(c_0 + \sum_{j=1}^{n} c_j x_{ij}) \\
& \quad \leq Y_i - (1-h)e_i, \\
& c_j \geq 0, \quad i = 1, \dots, m, \quad j = 1, \dots, n.
\end{aligned} \tag{۷}$$

برای هر مشاهده دو محدودیت به وجود آمده، با حل مسئله برنامه‌ریزی (۷) تابع هدف S با توجه به $2n$ محدودیت به وجود آمده توسط n مشاهده کمینه می‌گردد و مقادیر c_j به دست می‌آید.

۳ هیبرید رگرسیون کمترین مربعات خطی فازی

در این بخش رگرسیون کمترین مربعات با استفاده از حساب فازی وزنی معرفی‌شده است که برای فرمول‌بندی مربعات خطای پیش‌بینی‌شده و اندازه‌گیری شده متغیرهای فازی، بکار می‌رود[۷]. اعداد فازی به صورت توابع عضویت مثلثی متقارن بیان می‌شوند. بنابراین مدل رگرسیون دو متغیره را می‌توان به صورت

$$\hat{Y}_i = \tilde{A}_0 + \tilde{A}_1 X_i = (a_0, c_0) + (a_1, c_1)X_i \tag{۸}$$

بیان نمود و هر مقدار پیش‌بینی‌شده را می‌توان به صورت اعداد مثلثی متقارن زیر نمایش داد.

$$\hat{Y}_i = (a_0 + a_1 X_i, c_0 + c_1 X_i) \tag{۹}$$

که در آن $i = 1, 2, \dots, n$ است. هر مقدار مشاهده‌شده $\tilde{Y}_i$ را می‌توان به صورت $\tilde{Y}_i = (Y_i, e_i)$ نمایش داد. $^{\mu}\hat{Y}_i$ کران $\hat{Y}_i$ در سطح μ است. درحالی‌که $^{\mu}\tilde{Y}_i$ کران $\tilde{Y}_i$ در سطح عضویت μ است که از طریق رابطه زیر به دست می‌آیند.

$$\begin{aligned}
^{\mu}\hat{Y}_i &= [a_0 - (1-\mu)c_0] + [a_1 - (1-\mu)c_1]X_i \\
&= (a_0 + a_1 X_i) - (1-\mu)(c_0 + c_1)X_i,
\end{aligned} \tag{۱۰}$$

$$^{\mu}\tilde{Y}_i = Y_i - (1-\mu)e_i \tag{۱۱}$$

با استفاده از تعریف حساب فازی وزنی، مجموع باقیمانده خطا بین مقدار پیش‌بینی‌شده $\hat{Y}_i$ و مشاهده‌شده $\tilde{Y}_i$ به صورت (۱۲) فرمول‌بندی می‌شود.

$$\begin{aligned}
\sum (e)^2 &= \sum_{i=1}^{n} (\hat{Y}_i - \tilde{Y}_i)^2 \\
&= \sum_{i=1}^{n} \frac{[\int_0^1 [(^{\mu}\hat{Y}_i - {}^{\mu}\tilde{Y}_i)^2 \mu d\mu]}{\int \mu d\mu}
\end{aligned} \tag{۱۲}$$

که مخرج آن می‌تواند به صورت (۱۳) محاسبه شود.

$$\int \mu d\mu = 2\int_0^1 \mu d\mu = 2[\frac{1}{2}\mu^2]_0^1 = 2(\frac{1}{2}) = 1. \tag{۱۳}$$

و برای صورت می‌توان نشان داد:

$$\begin{aligned}
F &= \left[\int_0^1 [(^{\mu}\hat{Y}_i - {}^{\mu}\tilde{Y}_i)^2 \mu d\mu\right] \\
&= \int_0^1 [(a_0 + a_1 X_i - Y_i) - (1-\mu)(c_0 + c_1 X_i - e_i)]^2 \mu d\mu \\
&= \int_0^1 [(a_0 + a_1 X_i - Y_i)^2 \mu - 2\mu(1-\mu)(a_0 + a_1 X_i - Y_i) \\
&\quad (c_0 + c_1 X_i - e_i) + (c_0 + c_1 X_i - e_i)^2 \mu (1-\mu)^2] d\mu \\
&= \frac{1}{2}(a_0 + a_1 X_i - Y_i)^2 - \frac{1}{3}(a_0 + a_1 X_i - Y_i) \\
&\quad (c_0 + c_1 X_i - e_i) + \frac{1}{12}(c_0 + c_1 X_i - e_i)^2
\end{aligned} \tag{۱۴}$$

تابع هدف برای روش کمترین مربعات از حداقل کردن مجموع مربعات باقیمانده خطا (تابع F) به دست می‌آید. معادله (۱۴) دارای ۴ مجهول (a_0, a_1, c_0, c_1) است. برای به دست آوردن ضرایب مجهول مدل رگرسیونی بر اساس حداقل کردن F، باید از معادله (۱۴) بر حسب مجهول‌ها

مشتق گرفته و عبارت‌های به‌دست‌آمده را مساوی صفر قرار داد. از این طریق می‌توان ۴ مجهول را به دست آورد.

$$\frac{\partial F}{\partial a_0} = 2\sum_{i=1}^{n}(a_0 + a_1 X_i - Y_i) + \frac{1}{3}\sum_{i=1}^{n}[(c_0 + c_1 X_i - e_i)] = 0 \quad (15)$$

$$\frac{\partial F}{\partial a_1} = 2\sum_{i=1}^{n}(a_0 + a_1 X_i - Y_i)X_i + \frac{1}{3}\sum_{i=1}^{n}[(c_0 + c_1 X_i - e_i)]X_i = 0 \quad (16)$$

$$\frac{\partial F}{\partial c_0} = \frac{1}{3}(a_0 + a_1 X_i - Y_i) + 2(\frac{1}{12})(c_0 + c_1 X_i - e_i) = 0 \quad (17)$$

$$\frac{\partial F}{\partial c_1} = \frac{1}{3}(a_0 + a_1 X_i - Y_i)X_i + 2(\frac{1}{12})(c_0 + c_1 X_i - e_i)X_i = 0 \quad (18)$$

بعد از مرتب کردن معادله‌ها، مجموعه معادله‌های نرمال به دست می‌آیند. معادله چند مجهولی ۴×۴ بالا را می‌توان با استفاده از عملیات سطری ماتریس حل نمود [۱۳]. بعد از عملیات ماتریسی و بازآرایی معادله‌های (۱۵) تا (۱۸) می‌توان آن‌ها را به صورت دو دستگاه دو مجهولی ساده کرد.

$$\begin{cases} na_0 + \sum_{i=1}^{n}(X_i)a_1 = \sum_{i=1}^{n} Y_i \\ \sum_{i=1}^{n}(X_i)a_0 + \sum_{i=1}^{n}(X_i^2)a_1 = \sum_{i=1}^{n}(X_i Y_i) \end{cases} \quad (19)$$

با حل دستگاه (۱۹) می‌توان a_0 و a_1 را به دست آورد.

$$\begin{cases} nc_{0,L} + \sum_{i=1}^{n}(X_i)c_{1,L} = \sum_{i=1}^{n} e_{i,L} \\ \sum_{i=1}^{n}(X_i)c_0 + \sum_{i=1}^{n}(X_i^2)c_1 = \sum_{i=1}^{n}(X_i e_i) \end{cases} \quad (20)$$

از حل دستگاه (۲۰) c_0 و c_1 به دست می‌آیند.

۱.۳ اندازه قابلیت اطمینان در رگرسیون هیبرید

بعد از یافتن ضرایب رگرسیونی نوبت به ارزیابی اندازه قابلیت اطمینان معادله رگرسیون هیبرید می‌رسد. برای مقادیر مشاهده‌شده میانگین $\bar{\tilde{Y}}$ با استفاده از روش حساب فازی وزنی به عنوان اعداد غیرفازی محاسبه می‌شود و میزان انحراف معیار به صورت (۲۱) محاسبه می‌شود.

$$S_{\tilde{y}} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(\tilde{Y}_i - \bar{\tilde{Y}})^2} \quad (21)$$

مقدار $S_{\tilde{y}}$ نشان‌دهنده میزان پراکندگی یا پهنای داده‌ها می‌باشد. برای ارزیابی قابلیت اطمینان همبستگی ضریب هیبریدی از HR برای ارزیابی اثر هم خطی مدل رگرسیون خطی هیبرید استفاده می‌شود. خطای استاندارد هیبرید هم برای اندازه‌گیری میزان نیکویی برازش بین مدل رگرسیون هیبرید و داده‌های فازی مشاهده‌شده استفاده می‌شود. HS_e خطای استاندارد برآورد هیبرید هم برای اندازه‌گیری میزان نیکویی برازش بین مدل رگرسیون هیبرید و داده‌ها فازی مشاهده‌شده استفاده می‌شود. با توجه به‌اندازه قابلیت اطمینان برای مدل رگرسیون کلاسیک با توجه به [۴] HR و HS_e با استفاده از حساب فازی وزنی به صورت زیر تعریف می‌شود.

الف) ضریب همبستگی هیبرید

$$(HR)^2 = \frac{\sum_{i=1}^{n}(\hat{Y}_i - \bar{\tilde{Y}})^2}{\sum_{i=1}^{n}(\tilde{Y}_i - \bar{\tilde{Y}})^2} \quad (22)$$

ب) خطای استاندارد برآورد

$$HS_e = \sqrt{\frac{1}{n-p-1}\sum_{i=1}^{n}(\hat{Y}_i - \tilde{Y})^2} \quad (23)$$

که در آن $n-p-1$ درجه آزادی برای مدل دو متغیره p برابر یک می‌باشد. مقدار HS_e بین صفر و $S_{\tilde{y}}$ تغییر می‌کند. هرچقدر میزان HS_e بیشتر باشد، نیکویی برازش بهتر و در نتیجه دقت در بیش بینی بهتر می‌شود. اگر HS_e نزدیک و یا بیشتر از $S_{\tilde{y}}$ باشد مدل رگرسیونی موفق نخواهد بود. از آنجایی که $S_{\tilde{y}}$ یک مقدار ثابت و مستقل از مدل رگرسیونی است می‌توان از HS_e و نسبت $HS_e/S_{\tilde{y}}$ برای ارزیابی میزان اثربخشی مدل‌های رگرسیونی مختلف استفاده کرد.

۲.۳ فاصله اطمینان ضرایب رگرسیونی در رگرسیون هیبرید

همان‌طور که بحث شد ضرایب رگرسیون فازی به دست آمد و از این پارامترها در معیار حداقل سازی میزان فازی بودن و تحلیل بازه‌ای استفاده شد. به‌طور مشابه می‌توان از فاصله اطمینان برای محاسبه ضرایب در روش هیبرید خطی به عنوان روشی جایگزین برآورد نقطه‌ای پارامترها به برآورد فاصله‌ای برای ارزیابی میزان دقت آن باشد [۵] و [۱۸]. برای ساختن برآورد فاصله اطمینان برای ضرایب رگرسیون B_j بایستی فرض نرمال بودن توزیع خطای ε_i و دارای میانگین صفر و واریانس ثابت σ^2 و همچنین مستقل بودن را برای آن‌ها پذیرفت. بنابراین مشاهدات y_i دارای توزیع‌های نرمال مستقل با میانگین $\beta_0 + \sum_{i=1}^{k}\beta_j x_{ij}$ و واریانس σ^2 می‌باشند. چون برآوردگر حداقل مربعات $\hat{\beta}$ دارای توزیع نرمال با بردار میانگین β و ماتریس کوواریانس $\sigma^2(X'X)^{-1}$ می‌باشد، لذا توزیع حاشیه‌ای هر یک از ضرایب رگرسیونی $\hat{\beta}_j$ نرمال با میانگین β_j و واریانس $\sigma^2 C_{jj}$ است، که C_{jj} عضو j ام قطر ماتریس $(X'X)^{-1}$ می‌باشد. در نتیجه هر یک از آماره‌های

$$\frac{\hat{\beta}_j - \beta_j}{\sqrt{\hat{\sigma}^2 C_{jj}}}, \qquad j = 0, 1, \ldots, k \quad (24)$$

دارای توزیع t با $n-p-1$ درجه آزادی است. که $\hat{\sigma}^2$ برآورد واریانس خطا و MS_E میانگین مربعات باقیمانده‌ها می‌باشند که به صورت (۲۵) به دست می‌آیند.

$$MS_E = \frac{SS_E}{n-p-1} = \hat{\sigma}^2 \qquad (۲۵)$$

بنابراین یک فاصله اطمینان $(1-\alpha)100\%$ برای ضریب رگرسیون به (۲۶) صورت است [۱۴].

$$\hat{\beta}_j \pm t_{\frac{\alpha}{2},n-p-1}\sqrt{\hat{\sigma}^2 C_{jj}} \qquad (۲۶)$$

یا

$$\hat{\beta}_j \pm t_{\frac{\alpha}{2},n-p-1}\sqrt{\frac{SS_E}{n-p-1}C_{jj}} \qquad (۲۷)$$

۴ نتایج عددی

در این بخش با کمک مثال‌های عددی و با استفاده از مطالبی که در بخش‌های قبل ارائه شد رگرسیون فازی خطی موردبحث و بررسی قرارگرفته است. همان‌طور که در (۲۸) مشاهده می‌شود جفت داده‌های $(X_i : Y_i; i = 1,\ldots,8)$ که در آن X ها غیرفازی و Y ها فازی‌اند، ارائه‌شده است [۱۱].

$$[X_i, Y_i] = [(2:14),(4:11),(6:17),(8:15),(10:19),(12:22),(14:18),(16:30)]. \qquad (۲۸)$$

با اضافه کردن ± 1 به Y_i ها داده‌ها فازی می‌شوند و به صورت X_i غیرفازی و Y_i درمی‌آیند.

$$[X_i, \tilde{Y}_i] = [(2:(14,1)),(4:(11,1)),(6:(17,1)),(8:(15,1)),(10:(19,1)),(12:(22,1)),(14,(18,1)),(16,(30,1))]. \qquad (۲۹)$$

در این قسمت با توجه به جفت داده‌های ارائه‌شده رگرسیون فازی معرفی می‌شود که در هر مرحله نتایج به‌دست‌آمده از رگرسیون کلاسیک، رگرسیون بازه‌ای فازی، رگرسیون هیبرید کمترین مربعات فازی، اندازه قابلیت اطمینان و رگرسیون هیبرید و در نهایت فاصله اطمینان مقایسه می‌شود.

۱.۴ رگرسیون کلاسیک

در این قسمت از رگرسیون کمترین مربعات به عنوان معیاری برای مدل رگرسیون فازی و رگرسیون کمترین مربعات فازی و ... استفاده می‌شود. برای داده‌های X_i غیرفازی و Y_i فازی با استفاده از نرم‌افزار R ضرایب رگرسیون خطی کلاسیک به دست آورده شده است.

$$\hat{Y} = A_0 + A_1 X = 9.36 + 0.99X. \qquad (۳۰)$$

که در آن $A_0 = 9.36$ عرض از مبدأ و $A_1 = 0.99$ شیب خط رگرسیونی می‌باشند. در کنار معادله رگرسیونی، اندازه قابلیت اطمینان نیز ابزاری مهم برای تحلیل رگرسیونی می‌باشد.

انحراف استاندارد مشاهدات

$$S_y = \sqrt{\frac{1}{(n-1)}\sum_{i=1}^{n}(Y_i - \bar{Y})^2} = 5.8, \qquad (۳۱)$$

خطای استاندارد برآورد

$$S_e = \sqrt{\frac{1}{(n-2)}\sum_{i=1}^{n}(\hat{Y}_i - Y_i)^2} = 3.45, \qquad (۳۲)$$

ضریب همبستگی

$$R^2 = \frac{\sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2} = 0.70. \qquad (۳۳)$$

بنابراین

$$R = \sqrt{R^2} = 0.83. \qquad (۳۴)$$

۲.۴ رگرسیون امکانی

ر این روش با توجه به معادلات (۷)، طبق داده ارائه‌شده در (۳۱) برای هر مشاهده دو محدودیت تولید می‌شود. بنابراین برای این مثال ۱۶ محدودیت به وجود می‌آید. معادله تابع هدف برای حداقل کردن پهنای کل فازی ضرایب رگرسیونی بکار می‌رود که در سه حالت $h = 0$ ، $h = 0.5$ و $h = 0.7$ موردبررسی قرارگرفته است. محدودیت‌های زیر بر حسب $h = 0$ نوشته‌شده‌اند.

$$\begin{aligned}
& min \quad S = 8s_0 + 72s_1 \qquad (۳۵)\\
& m_0 + 2m_1 + s_0 + 2s_1 \geq 14\\
& m_0 + 4m_1 + s_0 + 4s_1 \geq 11\\
& m_0 + 6m_1 + s_0 + 6s_1 \geq 17\\
& m_0 + 8m_1 + s_0 + 8s_1 \geq 15\\
& m_0 + 10m_1 + s_0 + 10s_1 \geq 19\\
& m_0 + 12m_1 + s_0 + 12s_1 \geq 22\\
& m_0 + 14m_1 + s_0 + 14s_1 \geq 18\\
& m_0 + 16m_1 + s_0 + 16s_1 \geq 30\\
& m_0 + 2m_1 - s_0 - 2s_1 \leq 14\\
& m_0 + 4m_1 - s_0 - 4s_1 \leq 11\\
& m_0 + 6m_1 - s_0 - 6s_1 \leq 17\\
& m_0 + 8m_1 - s_0 - 8s_1 \leq 15
\end{aligned}$$

جدول ۱: رگرسیون امکانی

h	S	$m_۱$	$m_۰$	$s_۱$	$s_۰$
۰/۰۰	۳۸	۰/۹۲	۹/۹۳	۰/۲۲	۲/۷۶
۰/۵۰	۶۸	۰/۹۲	۹/۹۳	۰/۴۴	۴/۵۱
۰/۷۰	۱۰۸	۰/۹۲	۹/۹۳	۰/۷۴	۶/۸۶

$$m_۰ + ۱۰m_۱ - s_۰ - ۱۰s_۱ \leq ۱۹$$
$$m_۰ + ۱۲m_۱ - s_۰ - ۱۲s_۱ \leq ۲۲$$
$$m_۰ + ۱۴m_۱ - s_۰ - ۱۴s_۱ \leq ۱۸$$
$$m_۰ + ۱۶m_۱ - s_۰ - ۱۶s_۱ \leq ۳۰$$

برای $h = ۰$ جواب‌های زیر به دست می‌آید:

$$m_۰ = ۹/۳۶ \quad m_۱ = ۰/۹۲ \quad s_۰ = ۱/۷۶ \quad s_۱ = ۰/۲۲ \qquad (۳۶)$$

مقدار پهنای کل (تابع هدف) $s = ۳۰$ می‌باشد. بنابراین ضرایب فازی مدل را می‌توان به صورت

$$\hat{Y} = \tilde{A}_۰ + \tilde{A}_۱ X = (۹/۳۶, ۱/۷۶) + (۰/۹۲, ۰/۲۲)X \qquad (۳۷)$$

همان‌طور که در جدول ۱ مشاهده می‌شود، این مسئله برای مقادیر مختلف h موردبررسی قرارگرفته است. هدف انجام این کار تأثیر h بر مرکز و پهنای اعداد فازی $\tilde{A}$ است. در اثر تغییر h مرکز $\tilde{A}$ ها تغییر نکردند اما $s_۰$، $s_۱$ و S افزایش‌یافته‌اند. این یعنی با افزایش h میزان ابهام مدل بیشتر می‌شود. اندازه قابلیت اطمینان رگرسیون امکانی بر اساس اطلاعات موجود در جدول ۱ با استفاده از مراکز $\tilde{Y}_i$ محاسبه‌شده‌اند و نتایج آن $R = ۰/۷۴$، $S_e = ۴/۱۵$ و میزان $S_e/S_y = ۰/۷۲$ می‌باشند.

۳.۴ رگرسیون کمترین مربعات مبتنی بر رگرسیون امکانی

در این روش مرکز $\tilde{Y}$ فازی به عنوان مقادیر غیرفازی در نظر گرفته می‌شوند و با استفاده از روش رگرسیون کمترین مربعات کلاسیک ضرایب مدل رگرسیون کمترین مربعات به دست می‌آید که آن نتایج مشابه (۳۲) می‌باشد. سپس باید پهنای ضرایب رگرسیونی را به دست آورد که می‌توان با حل مسئله برنامه‌ریزی خطی (۳۸) این ضرایب را به دست آورد. نتایج به‌دست‌آمده از جدول ۲ را می‌توان با رگرسیون کمترین مربعات کلاسیک مقایسه کرد. همان‌طور که مشاهده می‌شود با افزایش h میزان پهنای مدل افزایش می‌یابد. محدودیت‌های زیر بر حسب $h = ۰/۷$ به‌دست‌آمده‌اند.

$$min = ۸c_۰ + ۷۲c_۱; \qquad (۳۸)$$
$$۹/۳۶ + ۲(۰/۹۹) + ۰/۳(s_۰ + ۲s_۱) \geq ۱۴/۳;$$
$$۹/۳۶ + ۴(۰/۹۹) + ۰/۳(s_۰ + ۴s_۱) \geq ۱۱/۳;$$
$$۹/۳۶ + ۶(۰/۹۹) + ۰/۳(s_۰ + ۶s_۱) \geq ۱۷/۳;$$
$$۹/۳۶ + ۸(۰/۹۹) + ۰/۳(s_۰ + ۸s_۱) \geq ۱۵/۳;$$
$$۹/۳۶ + ۱۰(۰/۹۹) + ۰/۳(s_۰ + ۱۰s_۱) \geq ۱۹/۳;$$
$$۹/۳۶ + ۱۲(۰/۹۹) + ۰/۳(s_۰ + ۱۲s_۱) \geq ۲۲/۳;$$
$$۹/۳۶ + ۱۴(۰/۹۹) + ۰/۳(s_۰ + ۱۴s_۱) \geq ۱۸/۳;$$
$$۹/۳۶ + ۱۶(۰/۹۹) + ۰/۳(s_۰ + ۱۶s_۱) \geq ۳۰/۳;$$
$$۹/۳۶ + ۲(۰/۹۹) - ۰/۳(s_۰ + ۲s_۱) \leq ۱۳/۷;$$
$$۹/۳۶ + ۴(۰/۹۹) - ۰/۳(s_۰ + ۴s_۱) \leq ۱۰/۷;$$
$$۹/۳۶ + ۶(۰/۹۹) - ۰/۳(s_۰ + ۶s_۱) \leq ۱۶/۷;$$
$$۹/۳۶ + ۸(۰/۹۹) - ۰/۳(s_۰ + ۸s_۱) \leq ۱۴/۷;$$
$$۹/۳۶ + ۱۰(۰/۹۹) - ۰/۳(s_۰ + ۱۰s_۱) \leq ۱۷/۷;$$
$$۹/۳۶ + ۱۲(۰/۹۹) - ۰/۳(s_۰ + ۱۲s_۱) \leq ۲۱/۷;$$
$$۹/۳۶ + ۱۴(۰/۹۹) - ۰/۳(s_۰ + ۱۴s_۱) \leq ۱۷/۷;$$
$$۹/۳۶ + ۱۶(۰/۹۹) - ۰/۳(s_۰ + ۱۶s_۱) \leq ۲۹/۷;$$

می‌توان مشاهده کرد که خط مرکزی رگرسیون کمترین مربعات فازی برابر خط رگرسیون کمترین مربعات کلاسیک $\hat{Y} = A_۰ + A_۱X = ۹/۹۳ + ۰/۹۹X$ است. با افزایش h میزان ابهام مدل بیشتر می‌شود. این مسئله برای مقادیر مختلف h موردبررسی قرارگرفته است. هدف انجام این کار تأثیر h بر مرکز و پهنای اعداد فازی $\tilde{A}_i$ است. در اثر تغییر h مرکز $\tilde{A}_i$ ها تغییر نکردند اما $s_۰$، $s_۱$ و S افزایش‌یافته‌اند.

جدول ۲: مدل رگرسیون کمترین مربعات مبتنی بر رگرسیون امکانی

h	S	m_1	m_0	s_1	s_0
۰/۰۰	۳۳/۱۱	۰/۹۹	۹/۳۶	۰/۲۱	۲/۲۵
۰/۵۰	۶۶/۲۲	۰/۹۹	۹/۳۶	۰/۴۲	۴/۴۹
۰/۷۰	۱۱۰/۳۷	۰/۹۹	۹/۳۶	۰/۷	۷/۴۹

۴.۴ هیبرید رگرسیون کمترین مربعات خطی فازی

از داده‌های (۳۱) به عنوان اعداد فازی مثلثی متقارن برای نشان دادن روش رگرسیون کمترین مربعات هیبریدی از طریق حساب فازی وزنی استفاده می‌شود. معادلات نرمال (۲۱) و (۲۲) می‌توان دستگاه معادلات (۳۹) و (۴۰) را محاسبه کرد.

$$\begin{cases} 8a_0 + 72a_1 = 142, \\ 72a_0 + 816a_1 = 1368 \end{cases} \tag{۳۹}$$

و

$$\begin{cases} 8c_0 + 72c_1 = 8, \\ 72c_0 + 816c_1 = 72. \end{cases} \tag{۴۰}$$

با حل معادله (۳۹) مقادیر $a_0 = 9{,}36$ و $a_1 = 0{,}99$ و حل معادله (۴۰) مقادیر $c_0 = 1$ و $c_1 = 0$ به دست می‌آیند. بنابراین مدل رگرسیون فازی هیبرید برای $\mu = 0$ به صورت (۴۱) می‌باشد.

$$\hat{Y} = A_0 + A_1 X = (9{,}36, 1) + (0{,}99, 0)X \tag{۴۱}$$

همان‌طور که در معادله (۴۱) مشاهده می‌شود پهنای فازی برای عرض از مبدأ ۱ و پهنای فازی برای شیب ۰ می‌باشد. مراکز فازی برابر ضرایب رگرسیون کمترین مربعات کلاسیک است. پهناهای فازی را می‌توان برای هر سطح عضویت به دست آورد. مثلاً معادله رگرسیون هیبرید برای $\mu = 0{,}7$ را می‌توان به صورت (۴۲) به دست آورد.

$$\hat{Y} = A_0 + A_1 X = (9{,}36, 0{,}3) + (0{,}99, 0)X \tag{۴۲}$$

با جایگذاری مقادیر در فرمول‌های متناظر، مقادیر قابلیت اطمینان هیبرید به صورت زیر محاسبه می‌شود.

الف) ضریب همبستگی هیبرید

$$(HR)^2 = \frac{\sum_{i=1}^{n}(\hat{Y}_i - \bar{\tilde{Y}})^2}{\sum_{i=1}^{n}(\tilde{Y}_i - \bar{\tilde{Y}})^2} \tag{۴۳}$$

$$= \frac{\sum_{i=1}^{n}[(a_0 + a_1 X_i - \bar{\tilde{Y}})^2 + \frac{1}{6}(c_0 + c_1 X_i)^2]}{\sum_{i=1}^{n}[(Y_i - \bar{\tilde{Y}})^2 + \frac{1}{6}(e_i)^2]}$$

$$= 0{,}71 \tag{۴۴}$$

بنابراین

$$HR = \sqrt{0{,}71} = 0{,}8484 \simeq 0{,}85 \tag{۴۵}$$

ضریب همبستگی رگرسیون کلاسیک $R = 0{,}835$ می‌باشد، این مقدار تقریباً با ضریب همبستگی رگرسیون فازی هیبرید برابر است.

ب) خطای استاندارد برآورد

$$\begin{aligned} HS_e &= \sqrt{\frac{1}{n-p-1}\sum_{i=1}^{n}(\hat{Y}_i - \tilde{Y}_i)^2} \\ &= \sqrt{\frac{1}{n-p-1}\sum_{i=1}^{n}[(a_0 + a_1 X_i - Y_i)^2} \\ &\quad - \sqrt{\frac{1}{n-p-1}(c_0 + c_1 x_i - e_i)^2} \\ &= 1{,}79 \end{aligned} \tag{۴۶}$$

که مقدار $(c_0 + c_1 x_i - e_i)$ صفر می‌شود. می‌توان نسبت $HS_e/S_{\tilde{y}}$ را محاسبه کرد.

$$\frac{HS_e}{S_{\tilde{y}}} = \frac{1{,}79}{3{,}14} = 0{,}59. \tag{۴۷}$$

مقایسه بین معادله رگرسیون و اندازه قابلیت اطمینان آن در جدول ۳ آورده شده است. در این قسمت اندازه قابلیت اطمینان و معادله‌های رگرسیونی و روش محاسبه آن‌ها مشاهده شد. می‌توان به این نتایج رسید که:

۱) مرکز فازی معادله (۴۱) با ضرایب روش رگرسیون کلاسیک برابر است و تغییر مقدار h تأثیری بر مرکز فازی ندارد.

۲) همان‌طور که در جدول ۳ مشاهده می‌شود معادله (۴۱) کمترین نیم پهنا را نسبت به سایر مدل‌های رگرسیون فازی دارد.

جدول ۳: اندازه قابلیت اطمینان برای مدل‌های ارائه‌شده

مدل رگرسیونی	معادله رگرسیونی	HR	HS_e	$HS_e/S_{\tilde{y}}$
رگرسیون کمترین مربعات هیبرید فازی $\mu = ۰$	$\hat{Y} = (۹/۳۶, ۱) + (۰/۹۹, ۰)X$	۰/۸۴	۳/۴۵	۰/۵۹
رگرسیون امکانی $h = ۰$	$\hat{Y} = (۹/۹۶, ۱/۷۶) + (۰/۹۲, ۰/۲۲)X$	۰/۷۵	۴/۱۵	۰/۷۱
رگرسیون کمترین مربعات فازی $h = ۰$	$\hat{Y} = (۹/۳۶, ۳/۲۵) + (۰/۹۹, ۰/۲۱)X$	۰/۷۳۵	۴/۲۴	۰/۷۳
رگرسیون کمترین مربعات کلاسیک	$\hat{Y} = ۹/۳۶ + ۰/۹۹X$	۰/۸۳	۳/۴۵	۰/۶۰
فاصله اطمینان ۹۵ درصد برای رگرسیون کمترین مربعات هیبرید	$\hat{Y} = (۹/۳۶, ۶/۵۸) + (۰/۹۹, ۰/۶۵)X$			
فاصله اطمینان ۹۰ درصد برای رگرسیون کمترین مربعات هیبرید	$\hat{Y} = (۹/۳۶, ۵/۲۳) + (۰/۹۹, ۰/۵۲)X$			
فاصله اطمینان ۸۰ درصد برای رگرسیون کمترین مربعات هیبرید	$\hat{Y} = (۹/۳۶, ۳/۸۷) + (۰/۹۹, ۰/۳۸)X$			

۳) همان‌طور که در جدول ۳ مشاهده می‌شود مقدار HS_e معادله (۴۱) با S_e رگرسیون فازی برابر است و در مقایسه با سایر مدل‌های رگرسیون فازی معادله (۴۱) کمترین HS_e را دارد و نسبت $HS_e/S_{\tilde{y}}$ نیز با رگرسیون کلاسیک تقریباً برابر است و کمترین مقدار را نسبت به سایر مدل‌های بررسی‌شده در جدول ۳ را داراست. می‌توان نتیجه گرفت مدل رگرسیون هیبرید بهترین نیکویی برازش را نسبت به سایر مدل‌های رگرسیونی جدول ۳ برخوردار است.
۴) علاوه بر ارزیابی اندازه قابلیت اطمینان روش رگرسیون هیبرید نسبت به سایر روش‌ها ساده‌تر است. به ویژه آنجایی که برنامه‌های رگرسیون کلاسیک قابل‌استفاده می‌باشند.

۵.۴ فاصله اطمینان ضرایب رگرسیونی در رگرسیون هیبرید

در این بخش با استفاده از فاصله اطمینان فازی عبارت‌های رگرسیون در سطح‌های مختلف ۸۰٪، ۹۰٪ و ۹۵٪ ارائه‌شده است که این فاصله‌ها در جدول آورده شده‌اند. معادله‌ها بر حسب مقادیر مرکز و پهنا نوشته‌شده‌اند. می‌توان دید که رگرسیون کمترین مربعات هیبرید در اطراف رگرسیون کمترین مربعات کلاسیک متمرکز هستند. در جدول ۳ مشاهده می‌شود که میزان فازی بودن با افزایش سطح اطمینان افزایش می‌یابد بنابراین برای اطمینان بیشتر باید عدم اطمینان بیشتر را پذیرفت.

همان‌طور که توضیح داده شد فاصله اطمینان را می‌توان به صورت (۴۸) به دست آورد.

$$\beta_۰ \pm t_{\frac{\alpha}{۲}, n-p-۱}\sqrt{\frac{SS_E}{(n-p-۱)}c_{۰۰}}, \qquad (۴۸)$$

$$\beta_۱ \pm t_{\frac{\alpha}{۲}, n-p-۱}\sqrt{\frac{SS_E}{(n-p-۱)}c_{۱۱}} \qquad (۴۹)$$

۵ بحث و نتیجه‌گیری

در این مقاله بررسی میان رگرسیون با ضرایب فازی (رگرسیون امکانی) و رگرسیون کمترین مربعات فازی مبتنی بر رگرسیون امکانی انجام شد و با مثال‌های عددی موردبررسی قرار گرفتند. برای تمامی روش‌ها به جز روش رگرسیون کمترین مربعات هیبرید فازی از برنامه‌ریزی خطی برای برآورد ضرایب فازی استفاده شد. که به ازای هر مشاهده دو محدودیت تولید شد، که با استفاده از نرم‌افزارهای حل مسائل برنامه‌ریزی خطی مانند *lingo* تابع هدف‌های موردنظر با توجه به محدودیت‌های تولیدشده توسط مشاهدات به حداقل رسانده شد. رگرسیون کمترین مربعات فازی هیبرید نوعی رفتار محدودکننده دارد که باعث کاهش میزان ابهام مدل می‌شود نتایج حاصل از رگرسیون هیبرید بسیار نزدیک به رگرسیون کمترین مربعات کلاسیک است. زمانی که از داده‌های غیرفازی استفاده می‌شود نتایج به‌دست‌آمده از رگرسیون کمترین مربعات فازی هیبرید با رگرسیون معمولی برابر است. همین طور می‌توان از برنامه‌های رگرسیون کلاسیک برای رگرسیون کمترین مربعات هیبرید استفاده نمود. به عبارت دیگر می‌توان گفت اندازه قابلیت اطمینان کلاسیک حالت خاصی از قابلیت اطمینان هیبرید و میزان عدم قطعیت است. در مقایسه با سایر روش‌های رگرسیون فازی، رگرسیون کمترین مربعات هیبرید بهترین اندازه قابلیت اطمینان را ارائه می‌کند. در نتیجه می‌تواند مدل پیش‌بینی بهتری را نسبت به سایر مدل‌های رگرسیون فازی ارائه دهد. تحلیل رگرسیون کمترین مربعات هیبرید و تحلیل اندازه قابلیت اطمینان آن یک روش کامل رگرسیونی شامل اعداد فازی است. روش‌های ارائه‌شده فقط برای اعداد فازی مثلثی اعمال شدند یعنی تمام داده‌ها دارای تابع عضویت مثلثی نرمال بودند.

مراجع

[۱] طاهری،سید محمود، (۱۳۹۶)، رویکردی ابتکاری در رگرسیون فازی، مجله اندیشه آماری، سال(۲۳)، شماره دوم پاییز و زمستان، ۴۳-۵۲.

[۲] طاهری، سید محمود، ماشین چی ماشاءالله، (۱۳۸۷)، مقدمه‌ای بر احتمال و آمار فازی، انتشارات دانشگاه شهید باهنر کرمان، کرمان.

[۳] میرزایی یگانه، شهره، ارقامی، ناصر رضا (۱۳۸۶). رگرسیون فازی: مروری بر چند رویکرد. اندیشه آماری، سال ۱۲ ،شماره ۱ (شماره پیاپی ۲۳) (شماره ویژه دومین کارگاه آمار و احتمال فازی)، ۳۵ -۴۷.

[4] Ayyub, B.M. and McCuen, R. (1996). *Numerical Methods for Engineers*, Prentice-Hall, New York.

[5] Bowermann, B.L.and O'Connell, R. (2000). *Linear Statistical Models: an applied approach*, Duxbury Press, Boston, USA.

[6] Celmins, A. (1987). Least squares model fitting to fuzzy vector data, *Fuzzy Sets and Systems,* **22**, 245–269.

[7] Chang, Y.H.O. (2001). Hybrid fuzzy least-squares regression analysis and its reliability measures, *Fuzzy Set and System,* **119**, 225–246.

[8] Chang, Y.H.O. and Ayyub, B.M. (2001). Fuzzy regression methods-a comparative assessment, *Fuzzy Set and System,* **119**, 187–203.

[9] Chukhrova, N. and Johannssen, A. (2019). Fuzzy regression analysis: Systematic review and bibliography, *Applied Soft Computing,* **84**, 105708

[10] Diamond, P. (1987). Fuzzy least squares, *Inform. Sci.,* **46**, 141–157.

[11] Kacprzyk, J. and Fedrizzi, M. (1991). *Fuzzy Regression Analysis*, Physica-Verlag, Heidelberg, .

[12] Kleinbaum, D.G. and Kupper, L.L. (1978). *Applied Regression Analysis and other Multivariable Methods*, Wadsworth, Belmont, CA.

[13] Kreyzig, E. (1993). *Advanced Engineering Mathematics*, seventh ed, Wiley, New York.

[14] Milton, J.S. amd Arnold, J.C. (1995). *Introduction to Probability and Statistics*, McGraw-Hill,Singapore.

[15] Savic, D. and Pedrycz, W. (1991). Evaluation of fuzzy regression models, *Fuzzy Sets and Systems,* **39**, 51–63.

[16] Seung ,H.C. and Jin, H.Y. (2010). General fuzzy regression using least squares method, *International Journal of Systems Science,* **41**, 477–485.

[17] Tanaka, H., Uejima, S. and Asai, K. (1982). Linear regression analysis with fuzzy model, *IEEE Transaction on Systems, Man and Cybernetics,* **12**(6), 903–907.

[18] Tutmez, B., Kaymak, V., Borgelt, C. , Ángeles Gil, M., Sousa, J. M. C. and Verleysen, M. (2013). *"Hybrid least-squares regression modelling using confidence bounds" in Towards Advanced Data Analysis by Combining Soft Computing and Statistics*, Berlin, Germany:Springer, 53–63.