مجله مدل‌سازی پیشرفته ریاضی ، دوره ۹، شماره، ۲، ۱۳۹۸

بهبود برآورد، آزمون فرضیه و فاصله اطمینان برای ضرایب فازی مدل رگرسیون خطی ساده با ورودی دقیق و خروجی فازی

محمد قاسم اکبری[۱]، غلامرضا حسامیان

*گروه آمار، دانشگاه بیرجند

**گروه آمار، دانشگاه پیام نور، تهران

تاریخ دریافت: ۱۳۹۶/۸/۱۴ تاریخ پذیرش: ۱۳۹۸/۵/۱۱

چکیده: در این مقاله، ابتدا یک روش موجود برای برآورد، انجام آزمون فرضیه فازی و فاصله اطمینان فازی برای ضرایب فازی یک مدل رگرسیون خطی ساده با ورودی‌های دقیق و خروجی‌های فازی معرفی می‌شود. سپس، معایب و نقایص این روش با تحلیل چند مثال عددی مورد بررسی و نقد قرار می‌گیرد. با معرفی یک رویکرد جایگزین مناسب در برآورد ضرایب و به‌کارگیری فرضیه‌های فازی متعارف در محیط فازی، سعی می‌کنیم روش مذکور در ساختار و اتخاذ تصمیم‌گیری برای پذیرش یا رد فرضیه‌های فازی را بهبود ببخشیم. برای این منظور، به کمک یک متر فازی متداول و تکنیک بوت استرپ، آماره‌های آزمون موردنیاز به‌صورت معیارهای غیر فازی تعریف می‌شوند. سپس، با مقایسه‌ی مقدار احتمال و سطح معنی‌داری داده شده، همانند روش کلاسیک، فرض صفر فازی قبول یا رد می‌شود. همچنین با ارائه چند مثال کاربردی، تأثیر و ارجحیت رویکرد پیشنهادی در آزمون فرضیه و فاصله اطمینان برای ضرایب فازی مدل رگرسیون خطی در مقایسه با روش موجود را مورد بررسی و تحلیل قرار می‌دهیم.

واژه‌های کلیدی: مشاهدات فازی، فرضیه فازی، ناحیه اطمینان فازی، بوت استرپ، متر بین دو عدد فازی.

رده‌بندی ریاضی (۲۰۱۰): ۹۴D۰۵، ۶۲G۰۸.

---

[۱]– آدرس الکترونیک مسئول مقاله: محمد قاسم اکبری، g_z_akbari@birjand.ac.ir

۱- مقدمه و پیش‌نیازها

مدل‌های رگرسیونی برای بررسی ارتباط بین یک متغیر پاسخ و تعدادی متغیر پیش‌بینی به کار می‌روند. در رگرسیون کلاسیک فرض می‌شود که متغیرها و مشاهدات مربوط به آن‌ها دقیق هستند. همچنین یک خطای تصادفی در چنین مدل‌هایی در نظر گرفته می‌شود. برای این خطای تصادفی و توزیع احتمالی آن، مفروضاتی مانند نرمال بودن، ناهمبسته بودن، ثبات واریانس و ... در نظر گرفته می‌شود؛ اما در بسیاری اوقات ممکن است یک یا چند فرضیه از فرضیه‌هایی که در مورد رگرسیون کلاسیک در نظر می‌گیریم برقرار نبوده و حتی بررسی صحت آن‌ها نیز مقدور نباشد. همچنین ممکن است متغیرهای تحت مطالعه دارای ارتباطی نادقیق باشند. در چنین شرایطی ابزارهای کلاسیک نمی‌توانند معیارهای مناسبی را برای مدل‌سازی داده‌ها ارائه دهند.

یکی از راه‌های ممکن، استفاده از نظریه‌ی مجموعه‌های فازی برای مدل‌سازی داده‌ها در چنین شرایطی است. رگرسیون آماری در محیط فازی توسط محققین مختلفی مورد بررسی قرارگرفته است. دیاموند و کرنر [۱] با معرفی متری در فضای اعداد فازی و با کمک روش کمترین مربعات فازی، به برازش مدل رگرسیون خطی با ورودی دقیق و خروجی فازی پرداختند. محمدی و طاهری [۲] روش کمترین مربعات برای برازش یک مدل رگرسیون خطی فازی را در یک کاربرد عملی به کاربردند. کائو و چیو [۳] روش کمترین مربعات را برای رگرسیون فازی با ورودی دقیق (و فازی) و خروجی فازی ارائه کردند. همچنین کائو و چیو [۴] یک روش دومرحله‌ای برای برآورد ضرایب مدل رگرسیون خطی فازی پیشنهاد نمودند. در مرحله اول، با استفاده از روش‌های غیرفازی ساز، ابتدا مشاهدات فازی را به اعداد معمولی تبدیل و سپس به کمک روش کمترین مربعات، ضرایب غیر فازی مدل را برآورد کردند. در مرحله دوم، به‌منظور افزایش قدرت توضیح دهندگی داده‌های فازی، یک عبارت خطای فازی به مدل برآورد شده اضافه و سپس عبارت خطا را به‌وسیله‌ی یک مسئله برنامه‌ریزی خطی برآورد نمودند. چوی و باکلی [۵] دو شیوه مختلف انجام رگرسیون خطی فازی ارائه نمودند. در مدل با ورودی‌های غیر فازی و ضرایب فازی، قدر مطلق کران‌های پایین (و بالای) آلفا-برش‌های اعداد فازی مشاهده شده را مینیمم کردند. برای مدل با ورودی فازی و ضرایب دقیق، ابتدا با غیرفازی سازی داده‌های فازی، رگرسیون کمترین قدر مطلق خطا را اجرا کردند. سپس با تعریف یک عبارت خطا به‌صورت یک عدد فازی، قدر مطلق تفاضل میان کران‌های پایین (و بالای) اعداد فازی مشاهده شده و برآورد شده را مینیمم کردند. فرارو و همکاران [۶] یک مدل رگرسیون خطی با ورودی و خروجی فازی پیشنهاد کردند که مبتنی بر روش کمترین مربعات است. چاچی و طاهری [۷] ابتدا با استفاده از متر هاسدوف که بر اساس آلفا-برش‌های اعداد فازی تعریف می‌شود، مجموع قدر مطلق خطای فازی را برای مدل رگرسیون خطی با متغیر ورودی غیر فازی، متغیر خروجی فازی و ضرایب فازی به دست

آوردند و سپس به کمک روش برنامه‌ریزی غیرخطی، پارامترها را برآورد کردند. لی و همکاران [۸] برای مدل رگرسیون خطی با ورودی غیر فازی، خروجی فازی و ضرایب رگرسیونی فازی روش کمترین مربعات را برای برآورد ضرایب فازی به کاربردند. همچنین آن‌ها روشی را برای انجام آزمون فرضیه و تعیین فاصله اطمینان برای ضرایب رگرسیونی فازی با استفاده از روش بوت استرپ پیشنهاد نمودند.

در این مقاله، روش لی و همکاران [۸] در برآورد ضرایب فازی مدل رگرسیون خطی با ورودی غیرفازی و خروجی فازی مرور و بررسی می‌شود. همچنین روش آن‌ها را در مسئله آزمون فرضیه و فاصله اطمینان برای ضرایب رگرسیون فازی، مورد بررسی و نقد قرار می‌دهیم. سپس با ذکر معایب روش لی و همکاران، یک رویکرد جایگزین مناسب معرفی و مزیت‌های آن نسبت به روش ایشان را بیان می‌کنیم. این مقاله شامل سه بخش است که در زیر خلاصه‌ای از مطالب هر بخش آمده است.

در بخش یک شرح مختصری از نظریه مجموعه‌های فازی، آلفا-برش، متر d_C و آزمون فرضیه‌های فازی ذکر می‌شود. در بخش دو روش لی و همکاران در برآورد ضرایب فازی مدل رگرسیون خطی فازی (متغیرهای ورودی غیرفازی و خروجی فازی) با استفاده از روش کمترین مربعات خطا بیان می‌گردد. همچنین رویکرد ایشان در آزمون فرضیه و فاصله اطمینان برای ضرایب مدل به‌تفصیل مرور و با چند مثال عددی توضیح می‌دهیم. سپس در هر مثال به بیان نقاط ضعف این روش می‌پردازیم. در بخش سه با کمک بوت‌استرپ در محیط فازی روش لی و همکاران بهبود بخشیده می‌شود و با ذکر چند مثال عددی عملکرد رویکرد پیشنهادی را مورد بررسی قرار می‌دهیم.

۱-۱- اعداد فازی

نظریه مجموعه‌های فازی، بسیاری از مفاهیم و متغیرهای مبهم را به شکل ریاضی صورت‌بندی می‌کند و زمینه را برای استدلال و استنتاج در شرایط عدم اطمینان فراهم می‌آورد. مجموعه فازی $\widetilde{A}$ (لطفی عسگر زاده [۹]) از مجموعه مرجع X به‌صورت $\{(x,\mu_{\widetilde{A}}(x)) : x\in X\}$ تعریف می‌شود که $\mu_{\widetilde{A}}(x) : X\to[0,1]$ تابع عضویت x در $\widetilde{A}$ است. مجموعه‌ی همه اعضایی از X را که درجه‌ی عضویت آن‌ها در مجموعه‌ی فازی $\widetilde{A}$، حداقل به بزرگی α باشد، الفا-برش $\widetilde{A}$ (مجموعه تراز αام $\widetilde{A}$) نامیده و با نماد $\widetilde{A}_{[a]}=\{x\in X : \mu_{\widetilde{A}}(x)\ge\alpha\}$ نشان داده می‌شود. مجموعه‌ی فازی $\widetilde{A}$ از $\mathbb{R}$ را یک عدد فازی گوییم، اگر:

۱) $\widetilde{A}$ تک نمائی باشد؛ یعنی دقیقاً یک $x_0\in\mathbb{R}$ وجود داشته باشد که $\mu_{\widetilde{A}}(x_0)=1$.

۲) الفا-برش‌های $\widetilde{A}$ بازه‌های بسته، کران‌دار و محدب به ازای هر $a \in (0,1]$ باشند. مجموعه همه اعداد فازی را با $F(\mathbb{R})$ نشان می‌دهیم.

نوع خاصی از اعداد فازی، اعداد فازی LR می‌باشند که علاوه بر ساختار ویژه، برخی اعمال حسابی برای آن‌ها از قواعد ساده‌تری پیروی می‌کنند. فرض کنید $\widetilde{A}$ یک عدد فازی روی مجموعه مرجع X باشد، آنگاه $\widetilde{A}$ را یک عدد فازی LR گوییم اگر تابع عضویت آن به‌صورت زیر باشد:

$$\mu_{\widetilde{A}}(x)=\begin{cases} L(\dfrac{a-x}{l_a}) & x \leq a, \\ \\ R(\dfrac{x-a}{r_a}) & x > a, \end{cases}$$

که در آن L و R توابعی غیر صعودی از $\mathbb{R}^+$ به $[0,1]$ هستند و $L(0)=R(0)=1$. یک عدد فازی LR را با نماد $\widetilde{A}=(a;l_a,r_a)_{LR}$ نشان می‌دهیم. عدد a را مقدار نمایی (میانه) و اعداد مثبت l_a و r_a را به ترتیب پهنای چپ و پهنای راست $\widetilde{A}$ می‌نامیم. در حالت خاص عدد فازی مثلثی $\widetilde{A}$ را با نماد $\widetilde{A}=(a,l_a,r_a)_T$ نشان می‌دهیم، اگر

$$R(x)=L(x)=\max\{0,1-x\}.$$

در اینجا برخی از روابط مهم در حساب اعداد فازی که در طول مقاله مورد استفاده قرار می‌گیرند به‌اختصار بیان می‌شوند.

برای دو عدد فازی $\widetilde{A}=(a;l_a,r_a)_{LR}$ و $\widetilde{B}=(b;l_b,r_b)_{LR}$ و عدد حقیقی $\lambda \in \mathbb{R}$، جمع و ضرب عدد ثابت بر اساس اصل گسترش به‌صورت زیر تعریف می‌شوند:

$$\widetilde{A}\oplus\widetilde{B}=(a;l_a,r_a)_{LR}\oplus(b;l_b,r_b)_{LR}=(a+b;l_a+l_b,r_a+r_b)_{LR},$$

$$\lambda\oplus\widetilde{A}=\begin{cases}(\lambda_a,\lambda l_a,\lambda r_a)_{LR} & \lambda>0, \\ (\lambda_a,-\lambda r_a,-\lambda l_a)_{RL} & \lambda<0.\end{cases}$$

همچنین، اعمال جبری اعداد فازی را می‌توان برحسب الفا-برش‌های آن‌ها نیز انجام داد. فرض کنید آلفا-برش‌های دو عدد فازی $\widetilde{A}$ و $\widetilde{B}$ به‌صورت زیر باشند:

$$\widetilde{A}_{[\alpha]}=[a_\alpha^L,a_\alpha^U]\ ,\qquad \widetilde{B}_{[\alpha]}=[b_\alpha^L,b_\alpha^U]$$

آنگاه داریم:

$$(\tilde{A}\oplus\tilde{B})_{[\alpha]}=\tilde{A}_{[\alpha]}+\tilde{B}_{[\alpha]}=[a_\alpha^L+b_\alpha^L, a_\alpha^U+b_\alpha^U],$$
$$(\tilde{A}\oplus\tilde{B})_{[\alpha]}=\tilde{A}_{[\alpha]}-\tilde{B}_{[\alpha]}=[a_\alpha^L-b_\alpha^L, a_\alpha^U-b_\alpha^U],$$
$$(\tilde{A}\otimes\tilde{B})_{[\alpha]}=\tilde{A}_{[\alpha]}\times\tilde{B}_{[\alpha]}$$
$$=[\min\{a_\alpha^L b_\alpha^L, a_\alpha^L b_\alpha^U, a_\alpha^U b_\alpha^L, a_\alpha^U b_\alpha^U\}, \max\{a_\alpha^L b_\alpha^L, a_\alpha^L b_\alpha^U, a_\alpha^U b_\alpha^L, a_\alpha^U b_\alpha^U\}],$$

که $a_\alpha^L(a_\alpha^U)$ و $b_\alpha^L(b_\alpha^U)$ کران‌های پایین و بالای الفا-برش‌های $\tilde{A}_{[\alpha]}$ $(\tilde{B}_{[\alpha]})$ هستند.

در مواردی که با تفاضل دو مجموعه‌ی فازی سرکار داریم، برای دست‌یابی به یک عدد فازی با میزان ابهام کمتر می‌توان از رابطه‌ی زیر که در استفانی [۱۰] به‌عنوان تفاضل عمومی هاکوهارا بیان شده است، استفاده نمود. اگر $\tilde{A}_{[\alpha]}=[A_\alpha^L, A_\alpha^U]$ و $\tilde{B}_{[\alpha]}=[B_\alpha^L, B_\alpha^U]$ به ترتیب آلفا ـ برش دو اعداد فازی باشند، آنگاه تفاضل بین آن‌ها بر اساس تفاضل عمومی هاکوهارا برابر است با:

$$\tilde{A}_{[\alpha]} *_G \tilde{B}_{[\alpha]}=[A_\alpha^L, A_\alpha^U] *_G [B_\alpha^L, B_\alpha^U]=[C_\alpha^-, C_\alpha^+], \qquad (۱)$$

که در آن $C_\alpha^-=\min\{A_\alpha^L-B_\alpha^L, A_\alpha^U-B_\alpha^U\}$ و $C_\alpha^+=\max\{A_\alpha^L-B_\alpha^L, A_\alpha^U-B_\alpha^U\}$ است.

یکی از مهم‌ترین جنبه‌های تحلیل داده‌های فازی، استفاده از یک‌فاصله مناسب برای اعداد فازی است. برای این منظور، مترها و فاصله‌های گوناگونی توسط نویسندگان مختلف مورد استفاده قرار گرفته‌اند (برای مثال، رجوع شود به دیاموند و کرنر [۱]). ازآنجاکه هدف این مقاله پیشنهاد یک رویکرد اصلاحی برای روش لی و همکاران [۸] در آزمون فرضیه‌ی ضرایب یک مدل رگرسیون خطی فازی است، ابتدا لازم است متر پیشنهادی آن‌ها را بیان و بررسی نماییم. با تعمیم متر اقلیدسی، لی و همکاران برای هر سطح از آلفا-برش‌های دو عدد فازی $\tilde{A}$ و $\tilde{B}$، متر زیر را پیشنهاد کردند.

تعریف ۱: فرض کنید $\tilde{A}_{[\alpha]}=[A_\alpha^L, A_\alpha^U]$، $\tilde{B}_{[\alpha]}=[B_\alpha^L, B_\alpha^U]$ نمایانگر آلفاـبرش‌های دو عدد فازی $\tilde{A}$ و $\tilde{B}$ باشند. فاصله بین آلفاـبرش‌های $\tilde{A}$، $\tilde{B}$ به‌صورت زیر تعریف می‌شود (هیلپرن [۱۱]):

$$d_c(\tilde{A}_{[\alpha]}, \tilde{B}_{[\alpha]})=\sqrt{(A_\alpha^C-B_\alpha^C)^۲+(A_\alpha^W-B_\alpha^W)^۲}, \qquad (۲)$$

که در آن

$A_\alpha^C = \frac{۱}{۲}(A_\alpha^L + A_\alpha^U)$ ، $B_\alpha^C = \frac{۱}{۲}(B_\alpha^L + B_\alpha^U)$ ، $A_\alpha^W = \frac{۱}{۲}(A_\alpha^U - A_\alpha^L)$ و $B_\alpha^W = \frac{۱}{۲}(B_\alpha^U - B_\alpha^L)$.

قضیه ۱: تابع $d_c : \mathbb{R}\times\mathbb{R} \to [۰,\infty]$ در شرایط متر صدق می‌کند؛ به عبارت دیگر برای هر $x, y, z \in R$ ، داریم:

۱) $d_c(x, y) = ۰$ اگر و تنها اگر $x = y$.

۲) $d_c(x, y) = d_c(y, x)$.

۳) $d_c(x, z) \le d_c(x, y) + d_c(y, z)$.

اثبات: شرایط ۱ و ۲ به‌سادگی قابل اثبات‌اند. برای اثبات شرط ۳، ابتدا توجه کنید که:

$$\begin{aligned}
d_c^۲(\widetilde{A}_{[\alpha]}, \widetilde{B}_{[\alpha]}) &= (A_\alpha^C - B_\alpha^C)^۲ + (A_\alpha^W - B_\alpha^W)^۲ \\
&= (A_\alpha^C + C_\alpha^C - C_\alpha^C - B_\alpha^C)^۲ + (A_\alpha^W - C_\alpha^W + C_\alpha^W - B_\alpha^W)^۲ \\
&\quad (A_\alpha^C - C_\alpha^C)^۲ + (C_\alpha^C - B_\alpha^C)^۲ + (A_\alpha^W - C_\alpha^W)^۲ + (C_\alpha^W - B_\alpha^W)^۲ \\
&\quad + ۲((A_\alpha^C - C_\alpha^C)(C_\alpha^C - B_\alpha^C) + (A_\alpha^W - C_\alpha^W)(C_\alpha^W - B_\alpha^W)).
\end{aligned}$$

اما طبق نامساوی کشی-شوارتز داریم:

$$\begin{aligned}
&(A_\alpha^C - C_\alpha^C)(C_\alpha^C - B_\alpha^C) + (A_\alpha^W - C_\alpha^W)(C_\alpha^W - B_\alpha^W) \\
&\le \sqrt{(A_\alpha^C - C_\alpha^C)^۲ + (A_\alpha^W - C_\alpha^W)^۲} \times \sqrt{(C_\alpha^C - B_\alpha^C)^۲ + (C_\alpha^W - B_\alpha^W)^۲},
\end{aligned}$$

در نتیجه

$$\begin{aligned}
d_c^۲(\widetilde{A}_{[\alpha]}, \widetilde{B}_{[\alpha]}) &\le (A_\alpha^C - C_\alpha^C)^۲ + (C_\alpha^C - B_\alpha^C)^۲ + (A_\alpha^W - C_\alpha^W)^۲ + (C_\alpha^W - B_\alpha^W)^۲ \\
&\quad + ۲\sqrt{(A_\alpha^C - C_\alpha^C)^۲ + (A_\alpha^W - C_\alpha^W)^۲} \times \sqrt{(C_\alpha^C - B_\alpha^C)^۲ + (C_\alpha^W - B_\alpha^W)^۲} \\
&= (d_c(\widetilde{A}_{[\alpha]}, \widetilde{C}_{[\alpha]}) + d_c(\widetilde{C}_{[\alpha]}, \widetilde{B}_{[\alpha]}))^۲.
\end{aligned}$$

بنابراین، شرط ۳ نیز اثبات می‌شود.

همان‌گونه که بیان شد، فاصله‌ی لی و همکاران وابسته به سطوح مختلف آلفا‌ـبرش‌های دو عدد فازی است؛ بنابراین، متر لی و همکاران برای کاربر درنهایت مشخص نمی‌کند که فاصله بین دو عدد فازی را در کدام سطح از آلفاـبرش می‌بایست تفسیر کرد. برای رفع چنین نقصی، متوسط فواصل آلفاـبرش‌های دو عدد فازی در روش لی و همکاران را به‌صورت زیر تعمیم می‌دهیم.

تعریف ۲: فرض کنید $\tilde{A}_{[\alpha]}=[A_\alpha^L, A_\alpha^U]$ و $\tilde{B}_{[\alpha]}=[B_\alpha^L, B_\alpha^U]$ نشانگر آلفاـبرش‌های دو عدد فازی $\tilde{A}$ و $\tilde{B}$ باشند. فاصله بین دو عدد فازی $\tilde{A}$ و $\tilde{B}$ به‌صورت زیر تعریف می‌شود:

$$d(\tilde{A},\tilde{B})=\sqrt{\int_۰^۱ d_c^۲(\tilde{A}_{[\alpha]},\tilde{B}_{[\alpha]})\,d\alpha.} \qquad (۳)$$

همانند قضیه ۱ می‌توان نشان داد که d در شرایط یک متر در محیط فازی صدق می‌کند.

قضیه ۲: تابع $d: F(\mathbb{R})\times F(\mathbb{R})\to F(۰,\infty)$ در شرایط زیر برای هر $\tilde{A},\tilde{B},\tilde{C}\in F(\mathbb{R})$ صدق می‌کند:

۱) $d(\tilde{A},\tilde{B})=۰$ اگر و تنها اگر $\tilde{A}=\tilde{B}$.

۲) $d(\tilde{B},\tilde{A})=d(\tilde{A},\tilde{B})$.

۳) $d(\tilde{A},\tilde{B})\le(\tilde{A},\tilde{C})+d(\tilde{C},\tilde{B})$.

اثبات. برای اثبات رابطه‌ی ۱، ابتدا فرض کنید $\tilde{A}=\tilde{B}$. در این صورت طبق رابطه‌ی ۱ از قضیه ۱، برای هر $\alpha\in[۰,۱]$داریم $d_c^۲(\tilde{A}_{[\alpha]},\tilde{A}_{[\alpha]})=۰$ و لذا $d(\tilde{A},\tilde{A})=۰$ حال از $d(\tilde{A},\tilde{B})=۰$ به‌سادگی نتیجه می‌شود که $d_c^۲(\tilde{A}_{[\alpha]},\tilde{B}_{[\alpha]})=۰$ برای $\alpha\in[۰,۱]$؛ بنابراین طبق رابطه‌ی ۱ از قضیه ۱ داریم $\tilde{A}=\tilde{B}$. رابطه‌ی ۲ نیز از رابطه‌ی ۲ قضیه ۱ به‌سادگی حاصل می‌شود. برای اثبات رابطه‌ی ۳، سه عدد فازی $\tilde{A},\tilde{B},\tilde{C}\in F(\mathbb{R})$ را در نظر بگیرید. با توجه به اثبات رابطه‌ی۳ قضیه داریم:

$$d^۲(\tilde{A},\tilde{B})=\int_۰^۱ d_c^۲(\tilde{A}_{[\alpha]},\tilde{B}_{[\alpha]})d\alpha\le\int_۰^۱(d_c(\tilde{A}_{[\alpha]},\tilde{C}_{[\alpha]})+d_c(\tilde{C}_{[\alpha]},\tilde{B}_{[\alpha]}))^۲d\alpha.$$

از طرفی بنا بر نامساوی مینکوفسکی (هاردی و همکاران [۱۲]) رابطه‌ی برقرار است:

$$\sqrt{\int_۰^۱(d_c(\tilde{A}_{[\alpha]},\tilde{C}_{[\alpha]})+d_c(\tilde{C}_{[\alpha]},\tilde{B}_{[\alpha]}))^۲\,d\alpha}$$
$$\le\sqrt{\int_۰^۱ d_c^۲(\tilde{A}_{[\alpha]},\tilde{C}_{[\alpha]})d\alpha}+\sqrt{\int_۰^۱ d_c^۲(\tilde{A}_{[\alpha]},\tilde{C}_{[\alpha]})d\alpha.}$$

بنابراین از روابط فوق نتیجه می‌شود که $d(\widetilde{A},\widetilde{B}) \le d(\widetilde{A},\widetilde{C}) + d(\widetilde{C},\widetilde{B})$.

توجه داشته باشید که برای بررسی فاصله‌ی بین دو عدد فازی، مترهای متعددی در متون فازی معرفی و بررسی شده‌اند. ازآنجاکه هدف مقاله حاضر، رفع نقطه ضعف‌های روش لی و همکاران است، لذا تمرکز ما در این مقاله بر متر اصلاحی فوق به‌عنوان یک رویکرد اصلاحی بدیهی و منطقی است.

مثال ۱. دو عدد فازی مثلثی $\widetilde{B} = (۵;۲,۱)_T, \widetilde{A}(۲;۱,۲)_T$ را در نظر بگیرید. ابتدا بنا بر تعریف ۱ داریم:

$$\begin{cases} A_\alpha^L = ۲-(۱-\alpha), \\ A_\alpha^U = ۲+۲(۱-\alpha), \end{cases} \qquad \begin{cases} B_\alpha^L = ۵-۲(۱-\alpha), \\ B_\alpha^U = ۵+(۱-\alpha), \end{cases}$$

$$\begin{cases} A_\alpha^C = ۲+\dfrac{(۱-\alpha)}{۲}, \\ A_\alpha^W = \dfrac{۳}{۲}(۱-\alpha), \end{cases} \qquad \begin{cases} B_\alpha^C = ۵-\dfrac{(۱-\alpha)}{۲}, \\ B_\alpha^W = \dfrac{۳}{۲}(۱-\alpha). \end{cases}$$

بنابراین از رابطه‌ی ۲، فاصله‌ی بین دو عدد فازی $\widetilde{B}$, $\widetilde{A}$ به‌صورت زیر حاصل می‌شود:

$$d(A,B) = \sqrt{\int_۰^۱ d_c^۲(A_{[\alpha]}, B_{[\alpha]})d\alpha} = \sqrt{\int_۰^۱ (-۳+(۱-\alpha))^۲ d\alpha} = ۲/۵۱۷.$$

۲-۱- فرضیه‌های فازی

روش‌های کلاسیک در آزمون فرضیه، مبتنی بر دقیق بودن فرضیه‌ها است. این دیدگاه باعث ایجاد محدودیت‌هایی در تصمیم‌گیری می‌شود. برای مثال فرض کنید θ پارامتری باشد که فرضیه $H_۰ : \theta = \theta_۰$ را برای آن می‌خواهیم بررسی کنیم. در این فرضیه به‌طور دقیق برابری θ با $\theta_۰$ آزمون می‌شود. اگر $\theta_۰$ کمی کمتر یا بیشتر در نظر گرفته شود ممکن است نتایج متفاوتی به دست آید و لذا دقیق بودن فرضیه با مشکل مواجه می‌شود؛ اما در بسیاری موارد، فردی که آزمون را انجام می‌دهد حساسیتی در مورد مقدار $\theta_۰$ ندارد و می‌خواهد بداند θ تقریباً برابر $\theta_۰$ است یا نه. در اینجا است که آزمون فرضیه فازی، ما را قادر به بررسی فرضیه‌هایی همچون «تقریباً برابر بودن»، «کوچک بودن»، «خیلی بزرگ بودن» و ... می‌سازد. لذا با چنین دیدگاهی دیگر شاهد محدودیت‌هایی که برای فرضیه‌های کلاسیک داشتیم، نیستیم.

در حالت غیر فازی، فرضیه‌ها به‌صورت $H_۰ : \theta = \theta_۰$ در مقابل $H_۱ : \theta \ne \theta_۰$، $H_۱ : \theta \ge \theta_۰$ در مقابل $H_۱ : \theta < \theta_۰$ و $H_۱ : \theta \le \theta_۰$ در مقابل $H_۱ : \theta > \theta_۰$ می‌باشند؛ اما بر اساس فرضیه‌های

فازی، فرضیه H_0 را می‌توان "θتقریباً برابر با θ_0"، "θتقریباً بزرگ‌تر از θ_0" و "θ تقریباً کوچک‌تر از θ_0" در نظر گرفت. بر این اساس می‌توان آزمون فرضیه‌های دوطرفه و یک‌طرفه را به‌صورت زیر بیان کرد:

الف) $\theta\,(\theta \approx \theta_0)$ تقریباً برابر θ_0 است: H_0 در مقابل $\theta\,(\theta \neq \theta_0)$ تقریباً برابر θ_0 نیست: H_1.

ب) $\theta\,(\theta \geq \theta_0)$ تقریباً بزرگ‌تر از θ_0 است: H_0 در مقابل $\theta\,(\theta \geq \theta_0)$ تقریباً بزرگ‌تر از θ_0 نیست: H_1.

ج) $\theta\,(\theta \leq \theta_0)$ تقریباً کوچک‌تر از θ_0 است: H_0 در مقابل $\theta\,(\theta \leq \theta_0)$ تقریباً کوچک‌تر از θ_0 نیست: H_1.

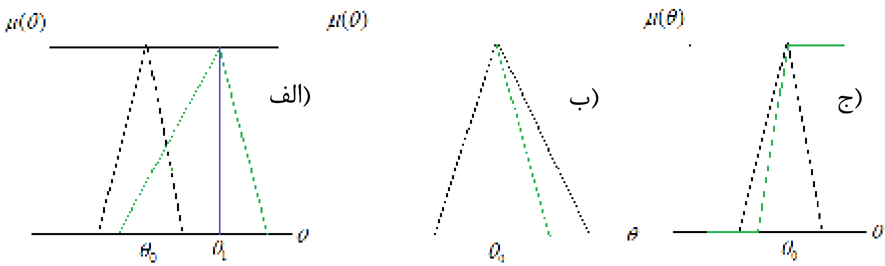
برای بیان فرضیه‌های فازی بالا به‌صورت مشخص‌تر اگر $\mu_{\tilde{\theta}_0}$ نمایانگر تابع عضویت $\tilde{\theta}_0$ باشد، آنگاه H_0 در مقابل H_1 به‌صورت زیر بیان می‌شود:

$\tilde{\theta}$ دارای عضویت $\mu_{\tilde{\theta}_0}(\theta)$ است: H_0

در مقابل

$\tilde{\theta}$ دارای تابع عضویت $\mu_{\tilde{\theta}}(\theta)$ است: H_1.

تابع عضویت در نظر گرفته شده برای فرضیه‌ی الف مانند تابع عضویت‌های معمولی فازی است؛ اما در حالت‌های ب و ج توابع عضویت به ترتیب صعودی و نزولی است که در شکل ۱ این توابع عضویت نشان داده شده است.



شکل (۱): نمودارهای مربوط به تابع عضویت‌های فرضیه‌های فازی الف، ب، ج

۲ – مدل رگرسیون خطی فازی بر اساس روش لی و همکاران

مدل رگرسیون خطی فازی در روش لی و همکاران به‌صورت زیر است:

$$\tilde{y} = \tilde{\beta}_{\circ} \oplus \tilde{\beta}_{1} x \oplus \tilde{\varepsilon}, \qquad (۳)$$

که در آن $\tilde{y}$ متغیر خروجی فازی، $x \geq \circ$ متغیر ورودی غیر فازی، $\tilde{\beta}_{\circ}$, $\tilde{\beta}_{1}$ عرض از مبدأ، شیب فازی خط رگرسیون و $\tilde{\varepsilon}$ خطای فازی هستند. توجه شود که مدل رگرسیون خطی بالا برحسب حساب بازه‌ای به‌صورت زیر تبدیل می‌شود:

$$\tilde{y}_{[\alpha]} = \tilde{\beta}_{\circ[\alpha]} + \tilde{\beta}_{1[\alpha]} x + \varepsilon_{[\alpha]}, \qquad \alpha \in [\circ,1], \qquad (۴)$$

که $\tilde{y}_{[\alpha]} = [y_{\alpha}^{L}, y_{\alpha}^{U}]$ آلفابرش متغیر خروجی، $\tilde{\beta}_{\circ[\alpha]} = [\beta_{\circ\alpha}^{L}, \beta_{\circ\alpha}^{L}]$, $\tilde{\beta}_{1[\alpha]} = [\beta_{1\alpha}^{L}, \beta_{1\alpha}^{U}]$ آلفابرش ضرایب رگرسیونی و $\tilde{\varepsilon}_{[\alpha]} = [\varepsilon_{\alpha}^{L}, \varepsilon_{\alpha}^{U}]$ آلفابرش خطا است. ضرایب رگرسیونی مدل (۴) را می‌توان با استفاده از تعمیم روش کمترین مربعات اعداد فازی و با استفاده از متر (۲) به دست آورد. برای این منظور، قرار می‌دهیم:

$$Q(\beta_{\circ\alpha}^{L}, \beta_{\circ\alpha}^{L}, \beta_{1\alpha}^{L}, \beta_{1\alpha}^{U}) = \sum_{i=1}^{n} d_c^{2}(\tilde{y}_{i\alpha}, \hat{\tilde{y}}_{i\alpha}) = \sum_{i=1}^{n} [(y_{i\alpha}^{C} - \hat{y}_{i\alpha}^{W})^{2} + (y_{i\alpha}^{C} - \hat{y}_{i\alpha}^{W})^{2}] =$$
$$\sum_{i=1}^{n} \Bigg[(\frac{y_{i\alpha}^{L} + y_{i\alpha}^{U} - \beta_{\circ\alpha}^{L} - \beta_{1\alpha}^{L} x_i - \beta_{\circ\alpha}^{U} - \beta_{1\alpha}^{U} x_i}{2})^{2} +$$
$$(\frac{y_{i\alpha}^{L} - y_{i\alpha}^{U} - \beta_{\circ\alpha}^{L} - \beta_{1\alpha}^{L} x_i + \beta_{\circ\alpha}^{U} - \beta_{1\alpha}^{U} x_i}{2})^{2}.$$

بنابراین، با مشتق‌گیری از Q نسبت به پارامترهای ضرایب رگرسیونی فازی:

$$\frac{\partial Q}{\beta_{j\alpha}^{L}} = \circ, \quad \frac{\partial Q}{\beta_{j\alpha}^{U}} = \circ, \quad j = \circ,1.$$

ضرایب $\beta_{\circ a}^{L}$ و β_{1a}^{L} به‌صورت زیر حاصل می‌شوند:

$$\hat{\tilde{\beta}}_{1\alpha}^{L} = \frac{\sum_{i=1}^{n} (y_{i\alpha}^{L} - \overline{(y_{\alpha}^{L})})(x_i - \overline{x})}{\sum_{i=1}^{n} (x_i - \overline{x})^{2}}, \qquad \hat{\tilde{\beta}}_{\circ\alpha}^{L} = \overline{(y_{\alpha}^{L})} - \hat{\beta}_{1\alpha}^{L} \overline{x}. \qquad (۵)$$

همچنین، $\beta_{\circ a}^{U}$ و $\beta_{\circ a}^{U}$ به‌صورت زیر محاسبه می‌شوند:

$$\hat{\tilde{\beta}}_{1\alpha}^{L} = \frac{\sum_{i=1}^{n} (y_{i\alpha}^{U} - \overline{(y_{\alpha}^{U})})(x_i - \overline{x})}{\sum_{i=1}^{n} (x_i - \overline{x})^{2}}, \qquad \hat{\tilde{\beta}}_{\circ\alpha}^{U} = \overline{(y_{\alpha}^{U})} - \hat{\beta}_{1\alpha}^{U} \overline{x}. \qquad (۶)$$

ازآنجاکه ضرایب رگرسیون مدل (۴) به‌صورت آلفا‌ـبرش هستند، بر اساس تفاضل هاکوهارا که در رابطه (۱) تعریف شد، برآورد ضرایب فازی به‌صورت زیر به دست می‌آیند:

$$\hat{\tilde{\beta}}_{0[\alpha]} = [\hat{\beta}^{-}_{0\alpha}, \hat{\beta}^{+}_{0\alpha}], \quad \hat{\tilde{\beta}}_{1[\alpha]} = [\hat{\beta}^{-}_{1\alpha}, \hat{\beta}^{+}_{1\alpha}], \tag{۷}$$

که $\beta^{-}_{j\alpha} = \min\{\hat{\beta}^{L}_{j\alpha}, \hat{\beta}^{U}_{j\alpha}\}$ و $\hat{\beta}^{+}_{j\alpha} = \max\{\hat{\beta}^{L}_{j\alpha}, \hat{\beta}^{U}_{j\alpha}\}$ درنهایت، برآورد ضرایب مدل رگرسیونی (۳) را می‌توان بر اساس برآورد آلفا‌ـبرش‌های این ضرایب با استفاده از اتحاد تجزیه به‌صورت زیر به دست آورد:

$$\hat{\tilde{\beta}}_{j} = \bigcup_{\alpha \in [0,1]} \alpha[\hat{\beta}^{-}_{j\alpha}, \hat{\beta}^{+}_{j\alpha}], \qquad j = 0,1. \tag{۸}$$

اما ایراد این روش این است که تضمینی وجود ندارد که $\hat{\beta}^{-}_{j\alpha}$, $\hat{\beta}^{+}_{j\alpha}$ به ترتیب دنباله‌های صعودی و نزولی از اعداد باشند که شرط اساسی برای تشکیل یک عدد فازی است.

۲-۱- آزمون فرضیه و فاصله اطمینان برای ضرایب رگرسیون خطی فازی بر اساس روش لی و همکاران

در این بخش، روش لی و همکاران در آزمون فرضیه و فاصله اطمینان برای ضرایب رگرسیون خطی فازی ساده موردبحث و کنکاش قرار می‌گیرد. این روش اساساً مبتنی بر تکنیک بوت استرپ است. توجه نمایید که تحلیل رگرسیون با استفاده از روش بوت استرپ معمولاً به دو روش انجام می‌پذیرد. روش اول بر پایه بازنمونه‌گیری از مشاهدات و روش دوم بر اساس بازنمونه‌گیری از خطاهاست. برای حالتی که X (متغیر ورودی) غیر تصادفی (قطعی) باشد، روش دوم مناسب‌تر است و در حالتی که X تصادفی باشد، روش بوت بر اساس نمونه تصادفی مشاهدات زوجی کاربرد بیشتری دارد. هرچند که این روش را وقتی‌که X غیر تصادفی هم باشد نیز می‌توان به کاربرد. برای دانستن مطالب بیشتر در این زمینه به افرون [۱۳] مراجعه گردد.

اما لی و همکاران برای آزمون ضرایب مدل رگرسیون خطی فازی (۴)، از روش بازنمونه‌گیری از خطاها استفاده کردند و فرضیه‌ها را به‌صورت زیر در نظر گرفتند:

$$H_0 : \tilde{\beta}_{j[\alpha]} = \tilde{\beta}_{j0[\alpha]} \quad \text{در مقابل} \quad H_1 : \tilde{\beta}_{j[\alpha]} \neq \tilde{\beta}_{j0[\alpha]}$$

که $\beta_{j0[\alpha]}$ یک بازه یا عدد مشخص است. آمار آزمونی که لی و همکاران برای انجام آزمون فوق در نظر گرفتند، یک مدل تعمیم‌یافته از آماره آزمون در حالت کلاسیک است که به‌صورت زیر تعریف می‌شود:

برای حالت $j=0$:

$$T_{1\alpha}=\sqrt{d_c^2\left(\frac{\hat{\tilde{\beta}}_{1[\alpha]} *_G \hat{\tilde{\beta}}_{1[\alpha]}}{s_{\hat{\beta}_{1\alpha}}},\{0\}\right)},\qquad s^2_{\hat{\beta}_{1\alpha}}=\frac{\sum_{i=1}^n d_c^2(\tilde{y}_{i[\alpha]},\hat{\tilde{y}}_{i[\alpha]})}{(n-2)\sum_{i=1}^n (x_i-\bar{x})^2}. \tag{۹}$$

برای حالت $j=1$:

$$T_{0\alpha}=\sqrt{d_c^2\left(\frac{\hat{\tilde{\beta}}_{0[\alpha]} *_G \hat{\tilde{\beta}}_{0[\alpha]}}{s_{\hat{\beta}_{1\alpha}}},\{0\}\right)}, \tag{۱۰}$$

$$s^2_{\hat{\beta}_{0\alpha}}=\frac{\sum_{i=1}^n d_c^2(\tilde{y}_{i[\alpha]},\hat{\tilde{y}}_{i[\alpha]})}{(n-2)}\left(\frac{\sum_{i=1}^n x_i^2}{n\sum_{i=1}^n (x_i-\bar{x})^2}\right),$$

که در آن $*_G$ تفاضل هاکوهارا است. همان‌طور که مشاهده می‌کنیم، آماره‌های آزمون (۹) و (۱۰) تعمیمی از حالت کلاسیک بوده و فقط بجای عبارات $e_i=(y_i-\hat{y}_i)^2$ و $\left|\hat{\beta}_j-\hat{\beta}_{j.0}\right|$ در حالت کلاسیک، به ترتیب از عبارات $d_c^2(\tilde{y}_{i[\alpha]},\hat{\tilde{y}}_{i[\alpha]})$ و $\sqrt{d_c^2(\hat{\tilde{\beta}}_{j[\alpha]} *_G \hat{\tilde{\beta}}_{j0[\alpha]},\{0\})}$ $j=0,1$ در حالت فازی استفاده می‌شود. همچنین توجه کنید که در روش بوت‌استرپ در تخمین یک پارامتر، از روش معمول زیر استفاده می‌شود. فرض کنید $\theta=\theta(F)$ پارامتر موردنظر باشد. آنگاه برآوردی که از جایگزین کردن تابع توزیع تجربی به‌جای توزیع جامعه حاصل می‌شود، برآوردگری است مانند $\hat{\theta}=t(\hat{F})$ که برآورد جایگزین نامیده می‌شود. برای مثال، امید ریاضی $(E(X))$ تابعی از تابع توزیع تجمعی $(F(x))$ به‌صورت $\theta=E(X)=\int xdF(x)$ است و $\int xd\hat{F}(x)=\frac{1}{n}\sum_{i=1}^n x_i=\bar{x}$. پس $\hat{\theta}=\bar{X}$ برآوردگر جایگزین $\theta=E(X)$ است. برای توضیح بیشتر در مورد بوت استرپ ناپارامتری، فرض کنید $x=(x_1,x_2,...,x_n)$ نمونه‌ای تصادفی از توزیع نامعلوم F باشد. از طرفی فرض کنید $\theta=\theta(x)$ آماره موردنظر بوده که علاقه‌مند به آگاهی از رفتارهای آن هستیم. ساختار کلی روش بوت‌استرپ ناپارامتری برای بررسی رفتار آماره به‌صورت زیر است:

۱. انجام B بار نمونه گیری با جایگذاری به حجم n از داده‌های اصلی $x=(x_1,x_2,...,x_n)$. این نمونه‌ها را به‌صورت زیر نشان می‌دهیم:

$$x^{*b}=(x_1^{*b},x_1^{*b},...,x_n^{*b}),\qquad b=1,\dots,B,$$

به نمونه‌های فوق، نمونه‌های بوت‌استرپ می‌گویند.

۲. محاسبه آماره موردنظر در هر بار نمونه‌گیری انجام شده در بند ۱ یعنی:

$$\theta^{*b} = \theta(x^{*b}), \qquad b = ۱,\ldots,B,$$

۳. به‌دست آوردن خاصیت موردنظر برای θ بر اساس تابع توزیع تجربی $\hat{\theta}^*$ها.

به‌عنوان مثال، اگر بخواهیم $\mathrm{var}[\theta(x)]$ را به روش بوت‌استرپ برآورد کنیم، می‌توان نوشت:

$$\mathrm{var}(\theta(x)) = \frac{۱}{B}\sum_{b=۱}^{B}(\theta^{*b} - \bar{\theta}^*)^۲,$$

که در آن $\bar{\theta}^* = \frac{۱}{B}\sum_{b=۱}^{B}\theta^{*b}$. برای درک بیشتر و دقیق‌تر مفاهیم در روش بوت‌استرپ، افرن دیاگرام زیر را رسم نمود که در آن فضای بوت‌استرپ، فضای موجود و ارتباط آن‌ها به نمایش گذاشته شده است.

فضای بوت‌استرپ

$$\left[\begin{array}{c} \hat{P} = \hat{F} \to \mathsf{x}^{*b} = (x_۱^{*b}, x_۲^{*b}, \ldots, x_n^{*b}) \\ \downarrow \\ \hat{\theta}^{*b} = \theta\left(\mathsf{x}^{*b}\right) \end{array}\right]$$

مقدار آماره موردنظر با داده‌های بوت‌استرپ

فضای حقیقی

$$\left[\begin{array}{c} P = F \to \mathsf{x} = (x_۱, x_۲, \ldots, x_n) \\ \downarrow \\ \hat{\theta} = \theta(x) \end{array}\right]$$

مقدار آماره موردنظر با داده‌های اصلی

همان‌طور که دیده می‌شود در فضای موجود، داده‌ها توسط توزیع نامعلوم F تولید شده‌اند. لذا آماره موردنظر نیز از داده‌های همین توزیع حاصل می‌گردند. از طرفی در روش بوت استرپ، توزیع داده‌ها توسط تابع توزیع تجربی برآورد می شود. سپس نمونه‌های بوت‌استرپ متناسب با این توزیع از دل داده‌ها تولید می‌گردند. درنهایت پارامتر متناظر توسط جایگذاری داده‌های بوت‌استرپ در تابع $\theta = \theta(x)$ برای هر نمونه حاصل می‌شود. حال با در نظر گرفتن تابع توزیع تجربی و اصل جایگزینی می‌توان رفتار دلخواه آماره را بررسی کرد. برای مثال در قسمت قبل که هدف برآورد واریانس آماره بود، با در نظر گرفتن تابع توزیع تجربی و عمل جایگزینی به نتیجه مطلوب زیر می‌رسیم.

$$\text{var}(\theta) = E_F(\theta - E_F(\theta))^۲ \xrightarrow{F \to \hat{F}} = \text{v}\hat{\text{a}}\text{r}(\theta) = \frac{۱}{n}\sum_{b=۱}^{B}(\theta^{*b} - \overline{\theta}^{*})^۲.$$

به عبارت دیگر، با انتخاب یک الگو یا یک توزیع احتمال، دقت برآورد توسط داده‌های موجود افزایش داده می‌شود.

با توجه به مطالب فوق، برای انجام آزمون فرضیه‌های موردنیاز، به‌صورت زیر عمل می‌کنیم:

گام اول:

ابتدا ضرایب مدل رگرسیون خطی (۴) را با استفاده از روش کمترین مربعات که در قبل بیان شد، محاسبه می‌کنیم. سپس مقدار باقی‌مانده‌ها به‌صورت زیر محاسبه می‌شوند:

$$\widetilde{\varepsilon}_{i[\alpha]} = [\hat{\varepsilon}^-_{i[\alpha]}, \hat{\varepsilon}^+_{i[\alpha]}], \quad i = ۱,\ldots,n,$$

که

$$\hat{\varepsilon}^-_{i[\alpha]} = \min\{y^L_{i\alpha} - \hat{\beta}^-_{۰\alpha} - \hat{\beta}^-_{۱\alpha}x_i, y^U_{i\alpha} - \hat{\beta}^+_{۰\alpha} - \hat{\beta}^+_{۱\alpha}x_i\},$$
$$\hat{\varepsilon}^+_{i[\alpha]} = \max\{y^L_{i\alpha} - \hat{\beta}^-_{۰\alpha} - \hat{\beta}^-_{۱\alpha}x_i, y^U_{i\alpha} - \hat{\beta}^+_{۰\alpha} - \hat{\beta}^+_{۱\alpha}x_i\}.$$

گام دوم:

مانده‌های برآورد شده را به‌صورت زیر به دست می‌آوریم:

$$\widetilde{e}_{i[\alpha]} = [e^-_{i\alpha}, e^+_{i\alpha}], \qquad i = ۱,\ldots,n,$$

$$e^-_{i\alpha} = \min\{\hat{\varepsilon}^-_{i\alpha} - \bar{\hat{\varepsilon}}^-_{i\alpha}, \hat{\varepsilon}^+_{i\alpha} - \bar{\hat{\varepsilon}}^+_{i\alpha}\}, \quad e^+_{i\alpha} = \max\{\hat{\varepsilon}^-_{i\alpha} - \bar{\hat{\varepsilon}}^-_{i\alpha}, \hat{\varepsilon}^+_{i\alpha} - \bar{\hat{\varepsilon}}^+_{i\alpha}\}.$$

گام سوم:

نمونه $(\widetilde{e}^{(۱)}_{j[\alpha]}, j = ۱,\ldots,n)$ با جایگذاری از مقادیر $\widetilde{e}_{j[\alpha]}$ را انتخاب می‌کنیم؛ بنابراین، مقادیر نمونه بوت استرپی $\widetilde{y}^{(۱)}_{j[\alpha]}$ به‌صورت زیر است:

$$\widetilde{y}^{(۱)}_{j[\alpha]} = \widetilde{y}_{j[\alpha]} + \widetilde{e}^{(۱)}_{j[\alpha],} \qquad j = ۱,\ldots,n.$$

گام چهارم:

محاسبه‌ی برآورد آلفابرش ضرایب فازی مدل به روش کمترین مربعات با استفاده از نمونه‌ی بوت استرپی در گام قبل:

$$\tilde{\beta}^{(\circ)}_{[\alpha]} = \left[\hat{\beta}^{(\circ)-}_{\circ[\alpha]}\ ,\ \hat{\beta}^{(\circ)+}_{\circ[\alpha]}\right],\quad \hat{\beta}^{(1)}_{1[\alpha]} = \left[\hat{\beta}^{(1)-}_{1\alpha}\ ,\ \hat{\beta}^{(1)+}_{1\alpha}\right],$$
$$\hat{\beta}^{(1)}_{k[\alpha]} = \min\{\hat{\beta}^{(1)L}_{k\alpha}\ ,\ \hat{\beta}^{(1)U}_{k\alpha}\},\quad \hat{\beta}^{(1)+}_{k[\alpha]} = \max\{\hat{\beta}^{(1)L}_{k\alpha}\ ,\ \hat{\beta}^{(1)U}_{k\alpha}\},\quad k=\circ,1.$$

گام پنجم:

تکرار گام سوم و چهارم برای B مرتبه و به دست آوردن نمونه بوت استرپی برای برآورد آلفا‌ـبرش‌های ضرایب فازی مدل رگرسیون خطی به‌صورت:

$$\hat{\hat{\beta}}^{(1)}_{j[\alpha]}, \hat{\hat{\beta}}^{(2)}_{j[\alpha]},\ldots,\hat{\hat{\beta}}^{(B)}_{j[\alpha]},\qquad j=\circ,1,\quad \alpha\in[\circ,1].$$

سپس با توجه به روابط (۹) و (۱۰)، آماره‌ی بوت استرپی متناسب با آزمون فرضیه را همانند $(T^{(1)}_{j\alpha}, T^{(2)}_{j\alpha},\ldots,T^{(B)}_{j\alpha}),\quad j=\circ,1.$ به دست می‌آوریم:

درنهایت با استفاده از نمونه بوت استرپی آماره آزمون موردنظر، مقدار-احتمال را به‌صورت

$$p-value = \frac{\sum_{b=1}^{B} I(T^{(b)}_{j\alpha}\geq T_{j\alpha})}{B},\quad j=\circ,1\ b=1,\ldots,B,$$

محاسبه می‌کنیم، که

$$I(T^{(b)}_{j\alpha}\geq T_{j\alpha}) = \begin{cases} 1 & T^{(b)}_{j\alpha}\geq T_{j\alpha}, \\ \circ & T^{(b)}_{j\alpha}\geq T_{j\alpha}. \end{cases}$$

توجه کنید که مقدار $T_{j\alpha}$ بر اساس آلفا‌ـبرش متغیر خروجی فازی مشاهده شده $(\tilde{y}_{i[\alpha]})$ و مقادیر برازش داده شده $\hat{\tilde{y}}_{i[\alpha]}$ به دست می‌آید. اگر مقدار-احتمال به دست آمده از λ کمتر باشد $(p-value<\lambda)$ آنگاه فرضیه صفر برای آزمون موردنظر را در سطح λ رد می‌کنیم و در غیر این صورت، فرضیه صفر را می‌پذیریم.

در ادامه با استفاده از توزیع تجربی نمونه بوت استرپی آماره آزمون $T^{(b)}_{j\alpha}$ برای ضرایب مدل رگرسیونی (۴)، روش لی و همکاران در تعیین فاصله اطمینان برای ضرایب مدل در سطح اطمینان $1-\lambda$ را مرور می‌کنیم.

در گام اول با اســـتفاده از روابط (۹) و (۱۰)، نمونه بوت اســـترپی آماره آزمون فرضـــیه‌های $H_\circ : \tilde{\beta}_{j[\alpha]} = [\circ,\circ]$ در مقابل $H_1 : \tilde{\beta}_{j[\alpha]} \neq [\circ,\circ]$ را به‌صورت زیر محاسبه می‌کنیم:

$$(T_{j\alpha}^{(۱)}, T_{j\alpha}^{(۲)}, ..., T_{j\alpha}^{(B)}), \quad j = ۰,۱.$$

در گام دوم با استفاده از توزیع تجربی آماره آزمون، چندک $t_{j\lambda}$ را طوری محاسبه می‌کنیم که در رابطه‌ی $p(T_{j\alpha} < t_{j\lambda}) = ۱-\lambda.$ صدق کند.

برای به دست آوردن چندک بالای نمونه‌ی بوت استرپی آماره آزمون، لازم است تا $(T_{j\alpha}^{(۱)}, T_{j\alpha}^{(۲)}, ..., T_{j\alpha}^{(B)})$ را به‌صورت صعودی مرتب کنیم. در این‌صورت مقدار چندک $t_{j\lambda}$ برابر با عضو $[(۱-\lambda)B]+۱$ام نمونه مرتب شده آماره آزمون است. (منظور از علامت [ ] در اینجا جزء صحیح است.) مقدار چندکی که از نمونه بوت استرپی برای $t_{j\lambda}$ به دست می‌آید را با $t_{j[(۱-\lambda)B]}$ نشان می‌دهیم. سرانجام، فاصله اطمینان برای ضریب رگرسیونی $\tilde{\beta}_{۱[\alpha]}$ به‌صورت زیر به دست می‌آید:

$$p(T_{۱\alpha} < t_{۱\lambda}) = ۱-\lambda \Rightarrow T_{۱\alpha} < t_{۱[(۱-\lambda)B]} \Rightarrow$$

$$(\beta_{۱\alpha}^{C} - \hat{\beta}_{۱\alpha}^{C})^{۲} + (\beta_{۱\alpha}^{W} - \hat{\beta}_{۱\alpha}^{W})^{۲} < t_{۱[(۱-\lambda)B]}^{۲} \cdot \frac{\sum_{i=۱}^{n} d_c^{۲}(\tilde{y}_{i[\alpha]}, \hat{\tilde{y}}_{i[\alpha]})}{(n-۲)\sum_{i=۱}^{n} (x_i - \overline{x})^{۲}}. \qquad (۱۱)$$

به‌طور مشابه، فاصله اطمینان برای ضریب رگرسیون $\tilde{\beta}_{۰[\alpha]}$ به‌صورت زیر حاصل می‌شود:

جدول (۱): داده‌های مثال ۲

ردیف	x_i	$\tilde{y}_i = (y_i; l_{y_i}, r_{y_i})_T$	ردیف	x_i	$\tilde{y}_i = (y_i; l_{y_i}, r_{y_i})_T$
۱	۱	(۶/۲;۰,۰/۶)	۹	۹	(۶;۰/۳,۱/۱)
۲	۲	(۶/۱;۰/۳,۰/۶)	۱۰	۱۰	(۶/۴;۰/۴,۰/۶)
۳	۳	(۵/۷;۰/۱,۰/۶)	۱۱	۱۱	(۶/۶;۰/۳,۰/۶)
۴	۴	(۶/۷;۰/۱,۰/۳)	۱۲	۱۲	(۶/۷;۰/۳,۰/۶)
۵	۵	(۵/۵;۰/۴,۰/۷)	۱۳	۱۳	(۶/۳;۰/۲,۰/۶)
۶	۶	(۵/۹;۰/۳,۰/۶)	۱۴	۱۴	(۶/۵;۰/۳,۰/۸)
۷	۷	(۶/۱;۰/۳,۰/۶)	۱۵	۱۵	(۶/۱;۰/۲,۱/۳)
۸	۸	(۶۰;۰/۶,۰/۹)			

مثال ۲: برای داده‌های جدول ۱، آزمون فرضیه‌های زیر را برای ضرایب مدل رگرسیون خطی (۴) در نظر بگیرید:

$$H_{\circ}:\tilde{\beta}_{j[\alpha]}=I\{\circ\},\quad H_{\backslash}:\tilde{\beta}_{j[\alpha]}\neq I\{\circ\},\qquad j=\circ,\backslash\quad \alpha\in[\circ,\backslash].$$

جدول (۲): $T_{\circ\alpha}$، نقطه بحرانی (CP) و p-value به دست آمده برای آزمون فرض در مورد عرض از مبدأ مدل رگرسیون خطی فازی در مثال ۲

α	۰	۰/۱	۰/۲	۰/۳	۰/۴	۰/۵
$T_{\circ\alpha}$	۳۸/۱۴۸	۳۸/۶۱۱	۳۸/۹۱۲	۳۹/۰۳۷	۳۸/۹۸۱	۳۸/۷۴۵
CP	۳۳/۸۲۳	۳۴/۰۳۷	۳۳/۹۴۶	۳۳/۸۴۶	۳۳/۶۹۴	۳۳/۱۱۸
p-value	۰/۰۱۵۴	۰/۰۱۵۶	۰/۰۱۳۸	۰/۰۱۳۰	۰/۰۱۱۳	۰/۰۰۸۰
α	۰/۶	۰/۷	۰/۸	۰/۹	۱	
$T_{\circ\alpha}$	۳۸/۳۳۸	۳۷/۷۷۵	۳۷/۰۷۶	۳۶/۲۶۴	۳۵/۳۶۴	
CP	۳۲/۴۹۰	۳۱/۷۶۵	۳۰/۷۰۲	۴۱۲/۲۹	۲۸/۰۰۸	
p-value	۰/۰۰۶۹	۰/۰۰۷۰	۰/۰۰۴۴	۰/۰۰۲۷	۰/۰۰۲۷	

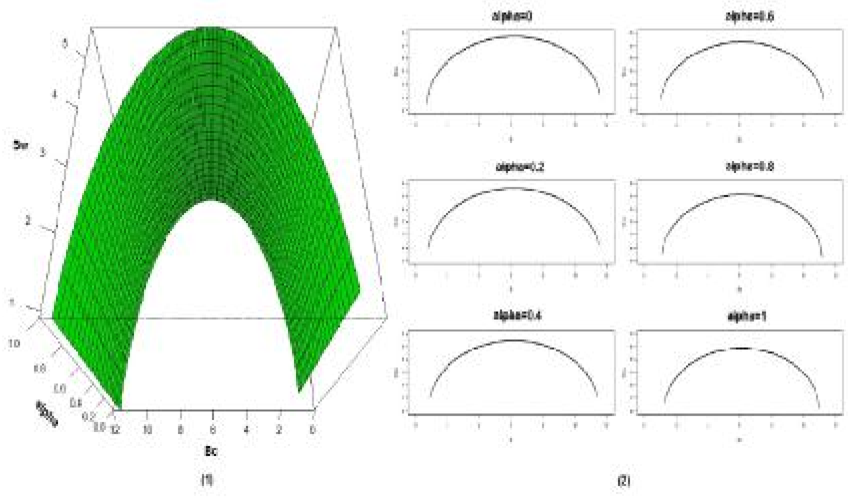
جدول (۳): $T_{\circ\alpha}$، CP و p-value برای آزمون فرض در مورد شیب مدل رگرسیون خطی فازی در مثال ۲

α	۰	۰/۱	۰/۲	۰/۳	۰/۴	۰/۵
$T_{\backslash\alpha}$	۲/۶۴۴۷	۲/۶۳۵۹	۲/۶۱۶۸	۲/۵۸۷۲	۲/۵۴۷۴	۲/۰۴۹۸
CP	۲/۷۷۲۱	۲/۷۵۶۴	۲/۰۷۱۸	۲/۶۹۹۵	۲/۶۳۵۵	۲/۵۶۸۸
p-value	۰/۱۴۲۱	۰/۱۳۷۹	۰/۱۳۲۴	۰/۱۳۲۷	۰/۱۲۸۰	۰/۱۲۰۸
α	۰/۶	۰/۷	۰/۸	۰/۹	۱	
$T_{\backslash\alpha}$	۲/۴۴۰۱	۲/۳۷۵۰	۲/۳۰۴۵	۲/۲۳۰۳	۲/۱۵۳۸	
CP	۲/۵۰۱۸	۲/۴۱۴۷	۲/۳۳۱۴	۲/۲۲۳۱	۲/۰۹۴۶	
p-value	۰/۱۱۸۶	۰/۱۱۱۵	۰/۱۰۸۳	۰/۰۹۷۸	۰/۰۸۲۰	

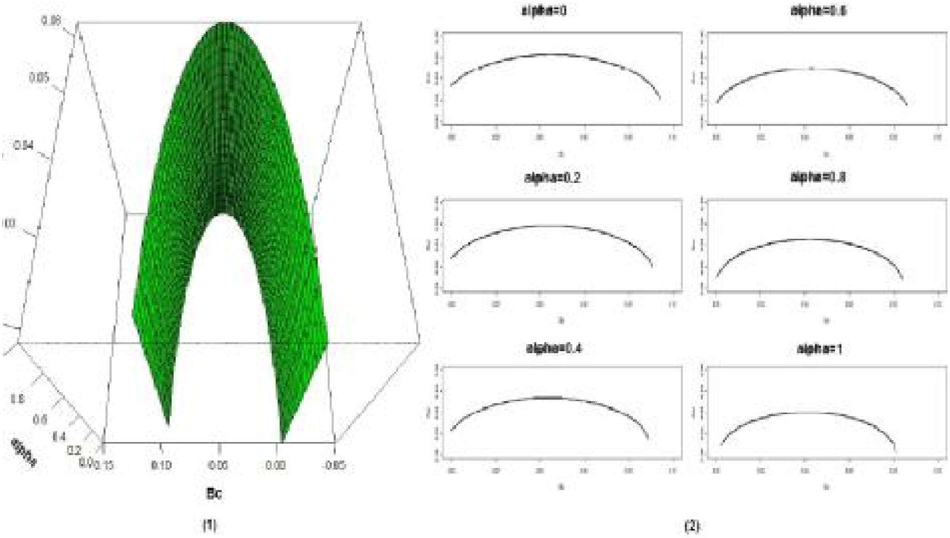
ابتدا برای هر یک از آزمون‌های بالا، با در نظر گرفتن $\lambda=\circ/\backslash$ و حجم نمونه بوت استرپی $B=\backslash\circ\circ\circ\circ$ مقدار $T_{j\alpha}$، نقطه بحرانی (CP) و p-value را برای ۱۱ آلفای مشخص به دست می‌آوریم. مقادیر به دست آمده را در جداول ۲ و ۳ مشاهده می‌کنید. با مقایسه‌ی این مقادیر، به ازای هر آلفا می‌توان برای رد یا پذیرش فرضیه صفر تصمیم گرفت. مقایسه‌ی مقادیر موجود در

جدول ۲ نتیجه می‌دهد که به ازای تمام مقادیر آلفا، فرضیه صفر کاملاً رد می‌شود. آنچه از جدول ۳ نتیجه می‌گیریم این است که به جزء برای آلفاهای ۰/۹ و ۱، دلیلی برای رد فرضیه صفر در سطح λ وجود ندارد؛ یعنی با کم شدن ابهام متغیر خروجی فازی تا آلفا ۰/۹ و ۱، فرض اینکه شیب مدل رگرسیون خطی فازی صفر است، توسط روش لی و همکاران رد نمی‌شود. همچنین فواصل اطمینان برای عرض از مبدأ فازی و شیب فازی خط رگرسیونی (۴) در سطح ۹۰%به ترتیب در نمودارهای ۲ و ۳ ترسیم شدند.

تذکر ۱: در آزمون فرضیه‌ی که در مثال ۱ مورد بررسی قرار گرفت، ازآنجاکه مدل (۴) به سطح آلفا–برش بستگی دارد، بهتر است فرضیه صفر معادل با برابری با آلفا–برش‌های دو عدد فازی باشد. در این صورت، برای تحلیل آزمون فرضیه‌ی دوطرفه برای ضرایب مدل (۳) می‌توان از نتایج به دست آمده استفاده کرد. همانند روش کلاسیک، این نواحی نشان می‌دهند که اگر در آزمون فرضیه‌ی دوطرفه برای آلفا–برش، ضرایب فازی موردنظر فرض صفر طوری اختیار شود که در این ناحیه قرار نگیرد، آنگاه در سطح ۰/۹ فرضیه صفر را رد می‌کنیم و در غیر این صورت دلیلی برای رد فرض صفر نداریم.



شکل (۲): (۱) نمودار سه‌بعدی فاصله اطمینان برای عرض از مبدأ $(\tilde{\beta}_{\circ[\alpha]})$ مدل رگرسیونی (۴) برحسب α و نقطه میانی $(\beta^{C}_{\circ\alpha})$ و پهنای $(\beta^{W}_{\circ\alpha})$ این ضریب. (۲) نمودارهای فاصله اطمینان برای عرض از مبدأ $(\tilde{\beta}_{\circ[\alpha]})$ مدل رگرسیونی (۴) برای برخی از αهای مشخص در مثال ۲



شکل (۳): (۱) نمودار سه‌بعدی فاصله اطمینان برای شیب $(\tilde{\beta}_{1[\alpha]})$ مدل رگرسیونی (۴) برحسب α و نقطه میانی $(\beta_{1\alpha}^{C})$ و پهنای $(\beta_{1\alpha}^{W})$ این ضریب. (۲) نمودارهای فاصله اطمینان برای شیب $(\tilde{\beta}_{1[\alpha]})$ مدل رگرسیونی (۴) برای برخی از αهای مشخص در مثال ۲

تذکر ۲: توجه داشته باشید که نواحی اطمینان، بر اساس روش بوت استرپی به دست آمده‌اند؛ بنابراین هنگامی‌که آلفا ـ برش ضریب فازی با آلفابرش فازی مشخصی برابر است، مطابق با تعبیر فاصله اطمینان کلاسیک، اگر ۱۰۰۰ مرتبه نمونه‌گیری کنیم در ۹۰۰ مرتبه آلفابرش ضریب فازی مدل برابر با این مقدار مشخص می‌شود و در ۱۰۰ مرتبه، برابر آلفابرش ضریب فازی مدل برابر آن نمی‌شود. ملاحظه می‌شود که برای چنین آلفابرش‌هایی، روش لی و همکاران در این مثال به‌خوبی عمل‌نکرده است؛ زیرا محدوده‌ای به‌عنوان فاصله اطمینان به دست آورده که دور از واقعیت است.

جدول (۴): داده‌های مثال ۳

ردیف	x_i	$\tilde{y}_i = (y_i; l_{y_i}, r_{y_i})_T$	ردیف	x_i	$\tilde{y}_i = (y_i; l_{y_i}, r_{y_i})_T$
۱	۵	(۱۱;۷,۸)	۵	۱۷	(۲۵;۲,۲)
۲	۸	(۱۶;۵,۴)	۶	۱۹	(۳۰;۴,۴)
۳	۱۱	(۱۸;۳,۳)	۷	۲۲	(۳۱;۴,۸)
۴	۱۴	(۲۴;۳,۲)	۸	۲۴	(۳۷;۹,۱۱)

مثال ۳: برای داده‌های جدول ۴ ضرایب مدل رگرسیون خطی (۳) را بر اساس روش ارائه شده توسط لی و همکارانش به دست می‌آوریم و مانند مثال ۲ آزمون فرضیه‌ها را برای ضرایب مدل رگرسیونی (۴) انجام می‌دهیم. با استفاده از مقادیر جدول ۵ و رابطه‌ی (۸)، برآورد ضرایب فازی مدل رگرسیونی (۳) به‌صورت زیر حاصل می‌شوند:

$$\hat{\tilde{\beta}}_{۱}=(۱/۲۷۲;۱/۲۴۴,۱/۴۴۰)_{T},\quad \hat{\tilde{\beta}}_{۰}=(۴/۹۱۸;۰/۷۲۰,۷/۶۵۲)_{T}.$$

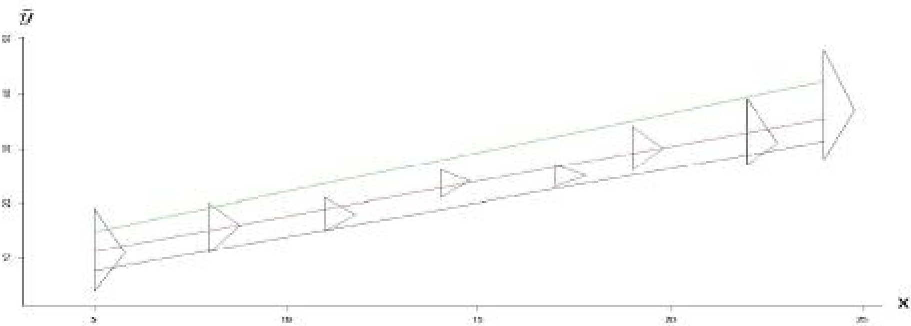
جدول (۵): برآورد آلفا ـ برش ضرایب رگرسیونی به دست آمده برای مثال ۳

α	۰	۰/۱	۰/۲	۰/۳	۰/۴	۰/۵
$\hat{\beta}^{-}_{۰\alpha}$	۰/۷۲۰	۱/۱۴۰	۱/۵۵۹	۱/۹۷۹	۲/۳۹۹	۲/۸۱۹
$\hat{\beta}^{+}_{۰\alpha}$	۷/۶۵۲	۷/۳۷۸	۷/۱۰۵	۶/۸۳۲	۶/۵۵۸	۶/۲۸۵
$\hat{\beta}^{-}_{۱\alpha}$	۱/۲۴۴	۱/۲۴۷	۱/۲۴۹	۱/۲۵۲	۱/۲۵۵	۱/۲۵۸
$\hat{\beta}^{+}_{۱\alpha}$	۱/۴۴۰	۱/۴۲۳	۱/۴۰۶	۱/۳۹۰	۱/۳۷۳	۱/۳۵۶
α	۰/۶	۰/۷	۰/۸	۰/۹	۱	
$\hat{\beta}^{-}_{۰\alpha}$	۳/۲۳۹	۳/۶۵۸	۴/۰۷۸	۴/۴۹۸	۴/۹۱۸	
$\hat{\beta}^{+}_{۰\alpha}$	۶/۰۱۱	۵/۷۳۸	۵/۴۶۸	۵/۱۹۱	۴/۹۱۸	
$\hat{\beta}^{-}_{۱\alpha}$	۱/۲۶۱	۱/۲۶۴	۱/۲۶۶	۱/۲۶۹	۱/۲۷۲	
$\hat{\beta}^{+}_{۱\alpha}$	۱/۳۳۹	۱/۳۲۲	۱/۳۰۶	۱/۲۸۹	۱/۲۷۲	

جدول (۶): نقطه بحرانی به دست آمده برای آزمون فرضیه‌ی عرض از مبدأ مدل رگرسیون خطی فازی برای داده‌های مثال ۳

α	۰	۰/۱	۰/۲	۰/۳	۰/۴	۰/۵
$T_{۰\alpha}$	۱/۸۹۹	۲/۰۱۲	۲/۱۵۰	۲/۳۱۸	۲/۵۲۰	۲/۷۵۶
CP	۲/۵۹۱	۲/۶۸۰	۲/۸۰۸	۲/۹۳۵	۳/۰۷۵	۳/۲۲۷
p-value	۰/۶۷۷۲	۰/۶۱۰۲	۰/۵۵۳۸	۰/۴۷۹۲	۰/۳۸۳۲	۰/۳۰۶۱
α	۰/۶	۰/۷	۰/۸	۰/۹	۱	
$T_{۰\alpha}$	۳/۰۲۲	۳/۲۹۵	۳/۵۳۶	۳/۶۸۹	۳/۷۱۰	
CP	۳/۴۰۴	۳/۵۴۲	۳/۶۰۸	۳/۵۴۶	۳/۲۳۰	
p-value	۰/۲۲۹۱	۰/۱۶۲۴	۰/۱۱۳۶	۰/۰۷۷۰	۰/۰۳۹۳	

همچنین با توجه به مقادیر جداول ۶ و ۷ می‌توان نتیجه گرفت که فرض صفر بودن عرض از مبدأ مدل رگرسیونی خطی فازی (۴)، برای $\alpha =$ ۰/۹،۱ در سطح $\lambda =$ ۰/۱ رد می‌شود اما برای سایر مقادیر آلفا، دلیلی برای رد فرضیه صفر در سطح $\lambda =$ ۰/۱ وجود ندارد. همچنین فرضیه صفر، به عبارت دیگر صفر بودن شیب مدل رگرسیونی خطی (۴)، برای $\alpha =$ ۰/۷،۰/۸،۰/۹،۱ در سطح $\lambda =$ ۰/۱ رد می‌شود و برای سایر مقادیر آلفا دلیلی برای رد فرضیه صفر وجود ندارد. این نتیجه با آنچه از پراکنش داده‌ها در شکل ۴ مشاهده می‌شود، هم خوانی ندارد. لذا این روش عملکرد ضعیفی در انجام آزمون فرضیه در این مثال دارد.



شکل (۴): نمودار برازش مدل رگرسیون خطی (۳) بر اساس روش لی و همکاران به داده‌های مثال ۳

جدول (۷): $T_{1\alpha}$ ، نقطه بحرانی (CP) و p-value به دست آمده برای آزمون فرضیه شیب مدل رگرسیون خطی فازی برای داده‌های مثال ۳

α	۰	۰/۱	۰/۲	۰/۳	۰/۴	۰/۵
$T_{1\alpha}$	۷/۶۴۳۸	۸/۲۹۲۵	۹/۰۴۴۶	۹/۹۱۵۳	۱۰/۹۱۴۷	۱۲/۰۳۶۴
CP	۸/۴۰۱۸	۹/۰۹۱۰	۹/۸۳۵۶	۱۰/۷۲۰۹	۱۱/۷۱۲۲	۱۲/۷۲۶۶
p-value	۰/۲۸۳۵	۰/۲۷۱۷	۰/۲۵۵۵	۰/۲۳۰۸	۰/۱۹۶۷	
α	۰/۶	۰/۷	۰/۸	۰/۹	۱	
$T_{1\alpha}$	۱۳/۲۳۷۱	۱۴/۴۰۵۹	۱۵/۳۴۳۰	۱۵/۷۹۹۰	۱۵/۶۰۸۵	
CP	۱۳/۷۳۶۰	۱۴/۳۰۹۷	۱۴/۶۸۷۹	۱۴/۱۲۶۳	۱۲/۶۷۰۷	
p-value	۰/۱۳۵۹	۰/۰۹۴۷	۰/۰۶۷۹	۰/۰۳۹۶	۰/۰۱۷۲	

درنهایت برخی از نقاط ضعف روش لی و همکاران عبارت‌اند از:

۱) لی و همکاران از تفاضل هاکوها استفاده کردند که این تفاضل دارای یک شرط وجودی بوده و در صورت عدم وجود آن، وجود نخواهد داشت.
۲) از دیگر نقاط ضعف این روش، تحلیل گنگ روش لی و همکاران در بررسی آزمون فرضیه‌ها برای ضرایب فازی مدل رگرسیون خطی فازی است؛ زیرا این روش به‌طور مشخص تعیین نمی‌کند که فرضیه صفر رد یا قبول می‌شود. در این روش رد یا قبول فرضیه H_0 به مقدار آلفا بستگی دارد.
۳) فواصل اطمینان نیز منوط به مقدار آلفا است.
۴) از محاسبات و مراحل زیاد برای رسیدن به برآورد ضرایب فازی استفاده شده است.
۵) برآورد ضرایب رگرسیونی بستگی به سطوح مختلف آلفا دارد. اما لی و همکاران به این نکته اشاره نکرده‌اند که این دنباله از اعداد، چگونه تشکیل یک عدد فازی می‌دهند.

۳- بهبود روش لی و همکاران در برآورد، مسئله آزمون فرضیه و فاصله اطمینان برای ضرایب رگرسیون خطی فازی

مدل رگرسیون خطی فازی به‌صورت زیر را در نظر بگیرید:

$$\tilde{Y}_i = \tilde{\beta}_0 \oplus \tilde{\beta}_1 x_i \oplus \tilde{\varepsilon}_i,$$

که

$$\tilde{Y}_i = (Y_i; l_{Y_i}, r_{Y_i})_T, \quad \tilde{\beta}_j = (\beta_j; l_{\beta_j}, r_{\beta_j})_T, \quad j = 0,1, \quad i = 1,\dots,n.$$

برای برآورد ضرایب فازی مدل رگرسیون فوق، از روش کمترین میانگین مربعات خطا به‌صورت زیر استفاده می‌کنیم:

$$MSE = \frac{1}{n}\sum_{i=1}^{n} d^2(\tilde{Y}_i, \tilde{Y}_i),$$

که d^2 مربع فاصله بین دو عدد فازی $\tilde{A}$ و $\tilde{B}$ است که در قضیه‌ی ۲ معرفی کردیم. توجه شود که معیار MSE را می‌توان به‌صورت مدل خطی زیر نوشت:

$$MSE = (Y - X\beta)'(Y - X\beta) + \frac{۱}{۲}(Y - X\beta)$$
$$\Big\{(R_Y - L_Y) - (X_s R_\beta - X_{۱-s} L_\beta - X_s l_\beta + X_{۱-s} R_\beta)$$
$$+\frac{۱}{۶}\Big[(L_Y - X_s(R_\beta + X_{۱-s}(L_\beta)'(L_Y - X_s(R_\beta + X_{۱-s}(L_\beta)\Big]$$
$$+\frac{۱}{۶}((R_Y - X_s(R_\beta + X_{۱-s}(L_\beta)'(R_Y - X_s(R_\beta + X_{۱-s}(L_\beta)\Big\},$$

که $\tilde{\beta} = (\beta; L_\beta . R_\beta)_T$ و

$$Y = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}, L_Y = \begin{pmatrix} l_{Y_1} \\ \vdots \\ l_{Y_n} \end{pmatrix}, R_Y = \begin{pmatrix} r_{Y_1} \\ \vdots \\ r_{Y_n} \end{pmatrix}, \beta = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_k \end{pmatrix}, L_\beta = \begin{pmatrix} l_{\beta_1} \\ \vdots \\ l_{\beta_k} \end{pmatrix}, R_\beta = \begin{pmatrix} r_{\beta_1} \\ \vdots \\ r_{\beta_k} \end{pmatrix},$$

$$X = \begin{pmatrix} x_{۱۱} & \cdots & x_{k۱} \\ \vdots & \ddots & \vdots \\ x_{۱n} & \cdots & x_{kn} \end{pmatrix}. \quad X_s = \begin{pmatrix} s_{۱۱}x_{۱۱} & \cdots & s_{k۱}x_{k۱} \\ \vdots & \ddots & \vdots \\ s_{۱n}x_{۱n} & \cdots & s_{kn}x_{kn} \end{pmatrix}. \quad s_{li} = I(x_{li} \geq ۰).$$

بنابراین پارامتر مجهول $\tilde{\beta} = (\beta; L_\beta, R_\beta)_T$ از حل دستگاه معادلات زیر حاصل می‌شود:

$$۰ = \frac{\partial}{\partial \beta} MSE = ۲X'X\beta - ۲X'Y + \frac{۱}{۲}X'[(R_Y - L_Y)$$
$$-(X_s R_\beta - X_{۱-s} L_\beta - X_s l_\beta + X_{۱-s} R_\beta)].$$
$$X_{۱-s}(R_\beta) + ۲X's(R_Y - X_s(R_\beta + X_{۱-s}(L_\beta),$$

$$۰ = \frac{\partial}{\partial R_\beta} MSE = \frac{۱}{۲}(I - X'_s + X'_{۱-s})(Y - X\beta)' - ۲X'_{۱-s}W_h(L_Y - X_s(L_\beta +$$
$$X_s R_\beta) + ۲X's(R_Y - X_s(R_\beta + X_{۱-s}(L_\beta).$$

درنهایت دستگاه معادلات فوق منجر به یک جواب یکتا با پهناهای مثبت به‌صورت $\tilde{\hat{\beta}} = (\hat{\beta}; \max\{۰, L_{\hat{\beta}}\}. \max\{۰, R_{\hat{\beta}}\})_T$ است که:

$$\begin{aligned}
\widehat{\beta} &= A^{-۱}(B-CD+CE(R_{\widehat{\beta}}-L_{\widehat{\beta}})),\\
L_{\widehat{\beta}} &= A_{۲}^{-۱}A_{۳}-A_{۲}^{-۱}A_{۱}(A_{۱}-A_{۲}(B_{۲})^{-۱}B_{۱})^{-۱}(A_{۳}-A_{۲}(B_{۲})^{-۱}B_{۳}),\\
R_{\widehat{\beta}} &= (A_{۱}-A_{۲}(B_{۲})^{-۱}B_{۱})^{-۱}(A_{۳}-A_{۲}(B_{۲})^{-۱}B_{۳}),
\end{aligned}$$

و

$$A=۲X'X,\ B=۲X'Y,\quad K_{۱}=\frac{۱}{۲}(E-I)Y,\ K_{۲}=\frac{۱}{۲}(E-I)X,$$

$$C=\frac{۱}{۲}X,\ D=R_Y-R_Y,\ E=X_s-X_{۱-s},$$

$$K_{۱}=\frac{۱}{۲}(E-I)Y,\ K_{۲}=\frac{۱}{۲}(E-I)X,$$

$$K_{۳}=-۴C_sL_Y+۴C_{۱-s}R_Y,\ K_{۴}=A_{ss}+A_{۱-s۱-s},$$

$$K_{۵}=A_{۱-ss}+A_{s۱-s},\quad A_{\alpha\beta}=۲X'_{\alpha}X_{\beta},$$

$$A_{۱}=K_{۲}A^{-۱}CE-K_{۵},\quad A_{۲}=-(K_{۲}A^{-۱}CE-K_{۴}),$$

$$A_{۳}=-(K_{۱}+K_{۲}A^{-۱}(B-CD)+K_{۳}),$$

$$B_{۱}=K_{۲}A^{-۱}CE-K_{۴},\quad B_{۲}=-(K_{۲}A^{-۱}CE-K_{۵}),$$

$$B_{۳}=-(K_{۱}+K_{۲}A^{-۱}(B-CD)+K_{۴}).$$

حال پس از تخمین ضرایب رگرسیونی فازی به روش فوق، آزمون فرضیه‌های فازی برای مدل رگرسیون خطی با ورودی‌های دقیق و خروجی فازی را به‌صورت زیر اصلاح می‌کنیم:

$$H_{\circ}:\tilde{\beta}_j=\tilde{\beta}_{\circ j}\qquad H_{\circ}:\tilde{\beta}_j\neq\tilde{\beta}_{\circ j}\qquad j=\circ,۱.$$

ملاحظه کنید که برخلاف روش لی و همکاران، فرضیه‌ها فوق به مقدار آلفا بستگی ندارند. برای انجام آزمون فرضیه‌ی فازی فوق، آماره آزمون و مقدار–احتمال، به روش بوت استرپ کلاسیک و طبق گام‌های زیر به محیط فازی تعمیم داده می‌شوند:

گام اول

نمونه $x_i^{(1)}, \tilde{y}_i^{(1)}$ به حجم n از بین متغیرهای خروجی فازی $\tilde{Y}_i$ و متغیرهای ورودی غیر فازی X_i را با جایگذاری انتخاب کرده و برآورد $\hat{\tilde{\beta}}_j^{(1)}$ را با استفاده از روش کمترین مربعات به دست می‌آوریم. در این صورت برآورد متغیر خروجی فازی به‌صورت $\hat{\tilde{y}}_i^{(1)} = \hat{\tilde{\beta}}_0^{(1)} \oplus \hat{\tilde{\beta}}_1^{(1)} \otimes x_i^{(1)}$ محاسبه می‌شود.

گام دوم

مقدار آماره‌های آزمون را به کمک مقادیر و نمونه‌هایی که از گام قبلی در اختیار داریم به‌صورت زیر به دست می‌آوریم:

$$\begin{cases} T_j^{(1)} = \dfrac{d(\hat{\tilde{\beta}}_j^{(1)}, \tilde{\beta}_{0j})}{S_{\tilde{\beta}_j^{(1)}}}, & j = 0,1, \\ S^2_{\tilde{\beta}_1^{(1)}} = \dfrac{\sum_{i=1}^{n} d^2(\hat{\tilde{y}}_i^{(1)}, \tilde{y}_i)}{(n-1)\sum_{i=1}^{n}(x_i^{(1)} - \overline{x}^{(1)})^2}, \end{cases}$$

که

$$S^2_{\tilde{\beta}_0^{(1)}} = \sum_{i=1}^{n} d^2(\hat{\tilde{y}}_i^{(1)}, \tilde{y}_i)\left(\frac{\sum_{i=1}^{n}(x_i^{(1)})^2}{n(n-2)\sum_{n}^{n}(x_i^{(1)} - \overline{x}^{(1)})^2}\right).$$

همان‌طور که مشاهده می‌شود، آماره‌های فوق در واقع تعمیم آماره‌های آزمون در حالت کلاسیک برای آزمون ضرایب رگرسیونی بر اساس مشاهدات و ضرایب فازی هستند.

گام سوم

گام‌های ۱ تا ۲ را B بار تکرار کرده و یک نمونه‌ی B تایی از آماره‌های آزمون معرفی شده در گام ۲ را به‌صورت $T_j^{(1)}, T_j^{(2)}, \ldots, T_j^{(B)}$ به دست می‌آوریم.

گام چهارم

قرار می‌دهیم:

$$p-value = \frac{\sum_{b=1}^{B} I(T_j^{(b)} > t_j)}{B}, \; j = 0,1. \tag{۱۳}$$

که در آن

$$I(T_j^{(b)} > t_j) = \begin{cases} ۱ & T_j^{(b)} > t_j \\ ۰ & T_j^{(b)} \le t_j \end{cases},\ t_j = \frac{d(\hat{\tilde{\beta}}_j, \tilde{\beta}_{۰j})}{S_{\tilde{\beta}_j}},$$

و

$$S_{\beta_۱}^۲ = \frac{\sum_{i=۱}^n d^۲(\hat{\tilde{y}}_i^{(۱)}, \tilde{y}_i)}{(n-۲)\sum_{i=۱}^n (x_i - \overline{x})^۲},\ S_{\tilde{\beta}_۰}^۲ = \sum_{i=۱}^n d(\hat{\tilde{\beta}}_j^{(۱)}, \tilde{\beta}_{۰j})(\frac{\sum_{i=۱}^n x_۱^۲}{n(n-۲)\sum_n^n (x_i - \overline{x})^۲}.$$

در روابط فوق، $\hat{\tilde{y}}_i$ و $\hat{\tilde{\beta}}_j$ برآوردهایی است که از گام (۱) حاصل شده‌اند. در این صورت، مقدار-احتمال را با استفاده از رابطه (۱۳) محاسبه می‌کنیم. درنهایت اگر $p-value<\lambda$ آنگاه فرضیه صفر را رد می‌کنیم و در غیر این صورت دلیلی برای رد فرضیه صفر در سطح λ نداریم. روش دیگری که می‌توان از آن برای انجام این آزمون‌ها استفاده کرد مقایسه $T_{j\alpha}$ با نقطه بحرانی (CP) است. نقطه‌ی بحرانی (CP) برابر با $[B(۱-\lambda)]+۱$ امین عضو نمونه بوت استرپی مرتب شده‌ی $T_{j\alpha}^{(b)}$ ها است. اگر $T_{j\alpha}$ بیشتر از نقطه بحرانی باشد، فرضیه صفر را در سطح λ رد می‌کنیم و در غیر این صورت آنرا می‌پذیریم.

مثال ۴: برای داده‌های جدول ۱ در مثال ۳ با استفاده از روش پیشنهادی، آزمون فرضیه‌ی دوطرفه را برای ضرایب فازی مدل رگرسیون خطی فازی انجام داده‌ایم. از جداول ۸ و ۹ می‌توان حدود تغییرات این ضرایب فازی را حدس بزنیم.

در ادامه این بخش، با استفاده از آماره آزمون بوت استرپی $T_j^{(b)}$ برای ضرایب مدل رگرسیونی (۱۲)، یک ناحیه اطمینانی در سطح $۱-\lambda$ برای این ضرایب پیشنهاد می‌شود. ابتدا گام ۱ تا ۳ معرفی شده را برای انجام آزمون فرضیه زیر انجام داده و نمونه بوت استرپی آماره آزمون موردنظر $(T_j^{(۱)}, T_j^{(۲)}, ..., T_j^{(B)})$ را به دست می‌آوریم.

$$H_۰ : \tilde{\beta}_{۰j} = (۰;۰,۰)_T, \quad H_۱ : \tilde{\beta}_{۰j} \ne (۰;۰,۰)_T, \qquad j = ۰,۱.$$

سپس با استفاده از توزیع تجربی آماره آزمون، مقدار چندک $t_{j\lambda}$ این آماره آزمون را به‌صورت

$$p(T_j < t_{j\lambda}) = ۱ - \lambda, \qquad j = ۰,۱.$$

تعیین می‌کنیم. برای تعیین چندک بالا، ابتدا نمونه بوت استرپی آماره آزمون را به‌صورت صعودی مرتب می‌کنیم. در این صورت، چندک $t_{j\lambda}$ برابر با عضو $[(۱-\lambda)B]+۱$ ام نمونه مرتب شده

آماره آزمون است. مقدار چندکی که از نمونه بوت استرپی برای $t_{j\lambda}$ به دست می‌آید را با $t_{j[(۱-\lambda)B]}$ نشان می‌دهیم. با توجه به رابطه‌ی بالا و برآورد چندک $t_{j\lambda}$، فاصله اطمینان برای ضرایب رگرسیونی $\tilde{\beta}_j$ را می‌توان از رابطه‌ی زیر به دست آورد:

$$p(T_j < t_{j\lambda}) = ۱-\lambda \Rightarrow T_j < t_{j[(۱-\lambda)B]} \Rightarrow d_c^{۲}(\hat{\tilde{\beta}}_j, \tilde{\beta}_j) < t^{۲}_{j[(۱-\lambda)B]} S^{۲}_{\tilde{\beta}_j} \Rightarrow$$

$$(\beta_j - \hat{\beta}_j)^{۲} + \frac{۱}{۲}(\beta_j - \hat{\beta}_j)((r_{\beta_j} - l_{\beta_j}) - (r_{\hat{\beta}_j} - -l_{\hat{\beta}_j})) +$$

$$\frac{۱}{۶}((r_{\beta_j} - r_{\hat{\beta}_j})^{۲} + (l_{\beta_j} - l_{\hat{\beta}_j})^{۲}) < t^{۲}_{j[(۱-\lambda)B]} S^{۲}_{\tilde{\beta}_j}, \qquad j = ۰,۱.$$

توجه کنید که این ناحیه برحسب میانه (β_j)، پهنای چپ (l_{β_j})، پهنای راست (r_{β_j}) و ضرایب فازی مدل رگرسیون خطی فازی (۱۲) است که در آن $\hat{\tilde{\beta}}_j$ برآورد ضرایب رگرسیونی با استفاده از روش کمترین مربعات خطا با متر d است و $S^{۲}_{\tilde{\beta}_j}$ به‌صورت زیر تعریف می‌شود:

$$S^{۲}_{\beta_۱} = \frac{\sum_{i=۱}^{n} d^{۲}(\hat{\tilde{y}}_i^{(۱)}, \tilde{y}_i)}{(n-۲)\sum_{i=۱}^{n}(x_i - \overline{x})^{۲}}, \;\; S^{۲}_{\beta_۰} = \sum_{i=۱}^{n} s^{۲}(\hat{\tilde{y}}_i^{(۱)}, y_i)(\frac{\sum_{i=۱}^{n} x_۱^{۲}}{n(n-۲)\sum_{n}^{n}(x_i - \overline{x})^{۲}}.$$

جدول (۸): آزمون فرضیه برای ضرایب رگرسیون خطی فازی برای داده‌های مثال ۴

$\beta_{۰۰}$	$(۰,۰,۰)_T$	$(۵/۵;۰/۲,۰/۳)_T$	$(۵/۵;۰/۲,۰/۳)_T$	$(۶;۰,۰)T$	$(۶;۰/۲,۰/۳)_T$
p − value	۰/۰۲۴۳	۰/۱۲۵۲	۰/۱۳۵۵	۰/۲۹۹۶	۰/۴۱۰۱
نتیجه	رد	قبول	قبول	قبول	قبول
$\beta_{۰۰}$	$(۶/۵;۰,۰)_T$	$(۶/۵;۰/۲,۰/۳)_T$	$(۷;۰,۰)_T$	$(۷:۰/۲,۰/۳)_T$	
p − value	۰/۱۱۹۵	۰/۱۲۵۶	۰/۰۲۱۰	۰/۰۱۸۵	
نتیجه	قبول	قبول	رد	رد	
$\beta_{۰۰}$	$(۱;۰,۰)_T$	$(۰;۰,۰)_T$	$(۰/۰,۰/۰۱,۰/۰۲)_T$	$(۰/۰۵;۰,۰)_T$	$(۰/۰۵;۰/۰,۰/۰۲)_T$
p − value	۰/۰۰۸۸	۰/۰۶۲۳	۰/۰۸۲۳	۰/۵۲۲۶	۰/۵۲۹۶
نتیجه	رد	رد	رد	قبول	قبول
$\beta_{۰۰}$	$(۰/۰۷;۰,۰)_T$	$(۰/۰۷;۰/۰۱,۰/۰۲)_T$	$(۰/۱۵;۰,۰)_T$	$(۰/۱۵;۰/۱,۰/۲)_T$	
p − value	۰/۲۶۸۹	۰/۲۶۸۴	۰/۰۷۰۳	۰/۰۶۸۳	
نتیجه	قبول	قبول	رد	رد	

جدول (۹): آزمون فرضیه برای ضریب رگرسیون خطی فازی برای داده‌های مثال ۴

$\beta_{۰۰}$	$(۰;۰,۰)_T$	$(۴;۰,۰)_T$	$(۴;۴,۳)_T$	$(۵;۰,۰)T$	$(۵;۴,۳)_T$
p – value	۰	۰	۰/۰۰۵۲	۰	۰/۵۴۷۹
نتیجه	رد	رد	رد	رد	قبول
$\beta_{۰۰}$	$(۵/۵;۰,۰)_T$	$(۵/۵;۴,۳)_T$	$(۷;۰,۰)_T$	$(۷;۴,۳)_T$	
p – value	۰/۰۰۰۴	۰/۰۱۶۸	۰	۰/۰۰۲۷	
نتیجه	رد	رد	رد	رد	
$\beta_{۰۰}$	$(۰;۰,۰)_T$	$(۱/۲۵;۰,۰)_T$	$(۱/۲۵;۰/۰۵,۰/۲)_T$	$(۱/۲۷;۰,۰)T$	$(۱/۲۷;۰/۰۵,۰/۲)_T$
p – value	۰	۰/۰۰۰۷	۰/۲۳۳۷	۰/۰۰۱۱	۰/۴۹۵۷
نتیجه	رد	رد	قبول	رد	قبول
$\beta_{۰۱}$	$(۱/۳;۰,۰)_T$	$(۱/۳;۰/۰۵;۰/۲)_T$	$(۱/۵;۰,۰)_T$	$(۱/۵;۰/۰۵,۰/۲)_T$	
p – value	۰/۰۰۳۸	۰/۰۸۸۷	۰	۰	
نتیجه	رد	رد	رد	رد	

تذکر ۳: در حالت خاص که ضرایب و متغیر خروجی مدل رگرسیونی (۱۲) اعداد فازی مثلثی متقارن باشند، فاصله اطمینان برای ضرایب فازی به‌صورت زیر تبدیل می‌شود:

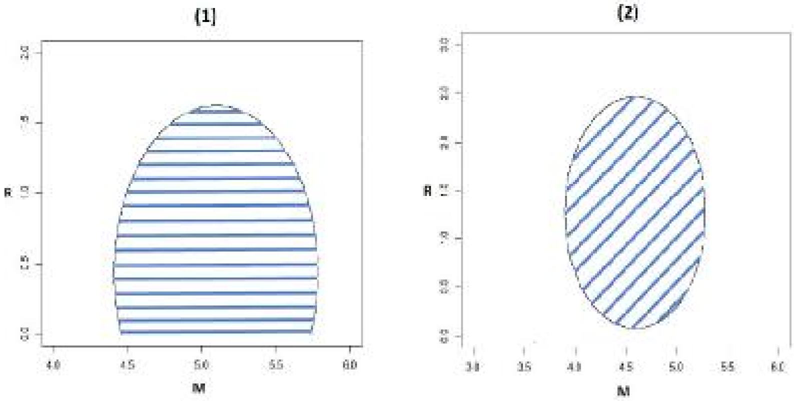
$$(\beta_j - \hat{\beta}_j)^۲ + \frac{۱}{۳}(r_{\beta_j} - r_{\hat{\beta}_j})^۲ < t^۲_{j[(۱-\lambda)B]} S^۲_{\beta_j}, \quad j = ۰,۱. \tag{۱۴}$$

جدول (۱۰): داده‌های مثال ۵

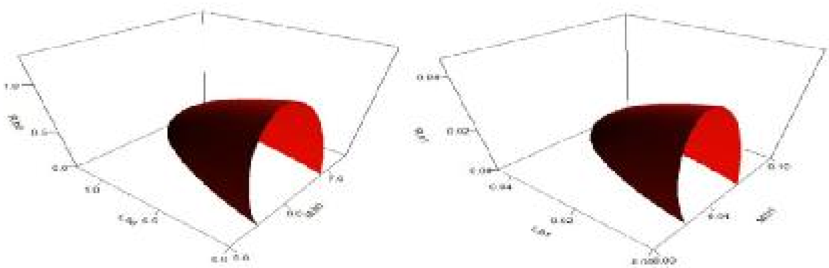
ردیف	x_i	$\tilde{y}_i$	ردیف	x_i	$\tilde{y}_i$
۱	۱/۴۹	$(۷۸/۱;۴۴/۱۱)_T$	۷	۳/۶۲	$(۸۱/۲;۵/۲۳)_T$
۲	۱/۶۷	$(۳۸/۲;/۱۴)_T$	۹	۳/۶۲	$(۸۴/۲;۳۴/۲۳)_T$
۳	۱/۹۹	$(/۲;۲۸/۱۵)_T$	۹	۳/۷۲	$(۸۷/۲;/۲۳)_T$
۴	۲/۰۳	$(۱۳/۰۷;۲/۰۱)_T$	۱۰	۴/۱۳	$(۲۵/۱۷;۳/۰۷)_T$
۵	۲/۸۹	$(۵۴/۲;۷۸/۱۹)_T$	۱۱	۴/۱۶	$(۱۲/۳;۶۷/۲۶)_T$
۶	۳/۰۱	$(۲۰/۷۸;۲/۵۱)_T$	۱۲	۴/۹۲	$(۴۷/۳;۸۳/۲۸)_T$

مثال ۵: برای داده‌های جدول ۱۰ فاصله اطمینان برای ضرایب فازی مثلثی متقارن مدل رگرسیونی (۱۲) در شکل ۵ رسم شده است. نواحی هاشور زده نشان‌دهنده‌ی فاصله اطمینان

برای ضرایب فازی مدل رگرسیون خطی موردنظر است. این نواحی نشان می‌دهند که اگر در آزمون فرضیه دوطرفه برای ضرایب فازی موردنظر، فرضیه صفر دارای میانه و پهنایی باشد که در این ناحیه قرار نگیرد، آنگاه در سطح معنی‌داری ۰/۱ فرضیه صفر را رد می‌کنیم و در غیر این صورت دلیلی برای رد آن نداریم.



شکل (۵): (۱) فاصله اطمینان برای شیب خط فازی (۲) فاصله اطمینان برای عرض از مبدأ فازی در سطح $1-\lambda = ۰/۹$ بر اساس داده‌های مثال ۵



شکل (۶): نمایش سه‌بعدی فاصله اطمینان در سطح $1-\lambda = ۰/۹$ برای عرض از مبدأ و شیب مدل رگرسیون خطی فازی بر اساس میانه، پهنای چپ و راست برای داده‌های مثال ۶

مثال ۶: برای داده‌های جدول ۱ با استفاده از روش پیشنهادی، فاصله اطمینان برای ضرایب فازی در سطح اطمینان $1-\lambda = ۰/۹$ به دست آوردیم. شکل ۶ نیز نمایش سه‌بعدی برای فاصله اطمینان در سطح ۰/۹ برای ضرایب فازی مدل رگرسیون خطی فازی (۱۲) بر اساس میانه، پهنای چپ و راست این ضرایب ارائه می‌دهد. همان‌طور که مشاهده می‌شود، این نمایش سه‌بعدی

با نتایج به دست آمده در جدول ۸ به‌خوبی هماهنگ است. برای مثال بر اساس نتایج جدول ۸، دلیلی برای رد فرضیه صفر $H_{\circ}:\tilde{\beta}_{\circ}=(۶;\circ,\circ)_T$ در سطح معنی‌داری $\lambda=\circ/۱$ وجود ندارد. از شکل ۶ همچنین نتیجه می‌شود که مقدار در نظر گرفته برای فرضیه صفر در محدوده این ناحیه سه‌بعدی قرار می‌گیرد؛ بنابراین، نتایج حاکی از عدم رد این فرضیه در سطح $\lambda=\circ/۱$ است.

بحث و نتیجه‌گیری

آزمون فرضیه و فاصله اطمینان دو موضوع با اهمیت در استنباط آماری هستند. در این مقاله به مقوله‌ی برآورد، آزمون فرضیه و فاصله اطمینان در یک مدل متداول رگرسیون خطی فازی با ورودی‌های دقیق و خروجی‌های فازی پرداخته شد. برای این منظور، روش لی و همکاران در محیط فازی مورد بررسی و کنکاش قرار گرفت. لی و همکاران، با تعریف فرضیه‌های فازی در سطوح مختلف آلفا-برش، روش بوت استرپ کلاسیک را برای انجام آزمون فرضیه و فاصله اطمینان در یک رگرسیون خطی فازی به کاربردند؛ اما در روش آن‌ها، نتیجه‌گیری در مورد رد یا پذیرش فرض صفر منوط به انتخاب یک سطح آلفا ویژه است. مشکل این روش این است که در بسیاری موارد برای کاربر مشخص نیست که فرض صفر درنهایت پذیرفته یا رد می‌شود؛ زیرا روش تصمیم‌گیری ایشان بستگی به این دارد که چه سطحی از آلفا-برش مدنظر کاربر باشد. در این راستا، با تکیه‌بر یک رویکرد اصلاح شده در تعیین ضرایب فازی و فرضیه‌های فازی، این مقاله تلاش کرد تا بر معایب روش لی و همکاران چیره شود. برای این منظور با اصلاح متری که لی و همکاران پیشنهاد کردند، آماره‌های آزمون همانند روش کلاسیک محاسبه شدند. در محاسبه آماره‌های آزمون، روش‌های بوت استرپ با استفاده از حساب اعداد فازی به کار گرفته شد. درنهایت همانند روش‌های کلاسیک، با مقایسه‌ی مقدار-احتمال و یک سطح معنی‌داری مشخص، به‌سادگی می‌توان در مورد رد یا قبول فرض صفر ابراز نظر نمود. همچنین سادگی در محاسبات، مزیت دیگر روش پیشنهادی ما در مقایسه با روش لی و همکاران است. در کل، مزیت‌های رویکرد جایگزین را می‌توان به‌صورت زیر بیان کرد:

۱) عدم وابستگی آماره‌های آزمون به مقدار آلفا-برش‌ها.
۲) عدم وابستگی نواحی اطمینان به مقدار آلفا-برش‌ها.
۳) سرعت بالا در محاسبات.
۴) تحلیل و نتیجه‌گیری آسان در تصمیم.

خاطرنشان می‌شود که روش پیشنهادی در این مقاله یک روش کلی در تحلیل رگرسیونی در محیط فازی است و می‌توان آن را در سایر مدل‌های رگرسیونی مانند رگرسیون نیمه پارامتری یا رگرسیون ریج به کاربرد.

منابع

[1] Diamond, P. and Korner, R. (1997). Extended fuzzy linear models and least squares estimates. *Computers and Mathematics with Applications*, **33**, 15-32.

[2] Mohammadi, J. and Taheri, S. M. (2004). Pedomodels ffitting with fuzzy least squares regression. *Iranian Journal of Fuzzy Systems*, **1** ,45-62.

[3] Lee, W. J., Jung, H. Y., Yoon, O. J. H. and Choi, S. H. (2014). The statistical inferences of fuzzy regression based on bootstrap techniques. *Soft Computing*, **19**, 883-890.

[4] Kao, C. and Chyu, C. L. (2003). Least-squares estimates in fuzzy regression analysis. *European Journal of Operational Research*, **148**, 426-435.

[5] Choi, S. H. and Buckley, J. J. (2008). Fuzzy regression using least absolute deviation estimators. *Soft Computing*, **12**, 257-263.

[6] Ferraro, M. B. and Coppi, R., Gonazlez-Rodriguez, G. and Colubi, A. (2010). A linear regression model for imprecise response. *International Journal of Approximate Reasoning*, **51**, 759-770.

[7] Chachi, J. and Taheri, S. M. (2013). A least-absolutes regression Model for imprecise response based on the generalized hausdorff-metric. *Journal of Uncertain Systems*, **7**, 265-276.

[8] Lee, W. J., Jung, H. Y., Yoon, O. J. H. and Choi, S. H. (2014). The statistical inferences of fuzzy regression based on bootstrap techniques. *Soft Computing*, **19,** 883-890.

[9] Zadeh, L. A. (1956). Fuzzy sets. *Information and Control*, **8**, 338-353.

[10] Stefanini, L. (2010). A generalization of hukuhara difference for interval and fuzzy arithmetic. *Fuzzy Sets and Systems*, **161**, 1564-1576.

[11] Heilpern, S. (1997). Representation and application of fuzzy numbers, *Fuzzy sets and Systems*, **91**, 259-268.

[12] Hardy, G. H., Littlewood, J. E. and Pólya, G. (1952). Inequalities. Cambridge Mathematical Library (second Ed.). Cambridge: Cambridge University Press.

[13] Efron, B. (1982). The jackknife, the bootstrap and other resampling plans. Society of Industrial and Applied Mathematics CBMS-NSF Monographs, Philadelphia. ISBN 0898711797.

An Improved Method Estimating and Performing Confidence Interval and Hypothesis Test in a Fuzzy Simple Linear Regression Model With Non-Fuzzy Inputs and Fuzzy Outputs

Mohammad Ghasem Akbari [*] , Gholamreza Hesamian [**]

[*] Department of Statistics, University of Birjand, Birjand, Iran

[**] Department of Statistics, Payame Noor University, Tehran, Iran.

Abstract

In this paper, an existing method to estimate coefficients of a fuzzy regression model with non-fuzzy inputs and fuzzy out-puts as well as its method to perform a fuzzy hypothesis test and a fuzzy confidence interval are recalled. Then, the disadvantages and shortcoming of this method are examined and criticized by analyzing several numerical examples. By introducing an appropriate alternative approach to estimating coefficients and applying conventional fuzzy hypotheses in a fuzzy environment, we try to improve the method in the structure and decision making to accept or reject fuzzy hypotheses. For this purpose, employing a common distance measure and Bootstrap technique, the required test statistics are defined as non-fuzzy criteria. Then, by comparing the p-value obtained from the test statistics and a given significance, unlike the existing one, one can easily decide to accept or reject the null hypothesis. Also, using some applied examples, the possible advantageous of the proposed approach in hypothesis testing and confidence interval for fuzzy coefficients are compared and discussed.

Keywords: Fuzzy observation, Fuzzy hypothesis, Confidence interval, Bootstrap, Distance measure between two fuzzy numbers.

Mathematics Subjects Classifications (2010): 62G08, 94D05.