سیستم‌های فازی و کاربردها
شاپای چاپی: ۲۷۱۷-۴۴۰۹
سال چهارم، شماره دوم (پائیز و زمستان ۱۴۰۰)
شاپای الکترونیکی: ۲۷۱۷-۳۹۸۴
صص. ۱۸۷ تا ۲۰۴
DOI:10.22034/JFSA.2022.310120.1093

رویکرد بوت‌استرپ در آزمون فرضیه برای واریانس متغیر تصادفی فازی

اعظم شهرکی، محمدحسین دهقان*، حامد احمدزاده

چکیده

هدف این مقاله بیان و محاسبه فاصله اطمینان واریانس با روش چاچی (تبدیل داده‌ها) و بدون تبدیل (داده‌های اولیه) با روش بوت‌استرپ است. به دلیل اینکه تبدیل داده‌ها به مقدار واریانس در مخرج عبارت تبدیل بستگی دارد، دچار نوسانات شده و در نتیجه برآورد فاصله‌های اطمینان و آزمون فرضیه به چالش کشیده می‌شود. نتایج هردو روش را مقایسه و مزایا و معایب آنها بررسی می‌شوند. برای این منظور، در ابتدا با مفاهیمی مثل متغیر تصادفی فازی و فاصله بین داده‌های فازی بر اساس مفهوم α-شک و نیز α-برش پرداخته می‌شوند. سپس فرآیند آزمون فرضیه را برای واریانس داده‌های فازی یک نمونه‌ای انجام می‌شود. در انجام آزمون فرضیه از روش بوت‌استرپ استفاده شده است و فاصله‌های اطمینان مذکور را با فاصله‌های اطمینان حاصل از روش α-برش مقایسه شده است.

۱ مقدمه

پس ازمعرفی نظریه مجموعه‌های فازی در سال ۱۹۶۵ ، علوم مختلف به بهره‌گیری از این مفهوم روی آورده و از کاربردهای گوناگون آن استفاده نموده‌اند. در علم آمار و احتمال نیز اولین بار مفهوم متغیر تصادفی فازی توسط واکرناک در سال

عبارات و کلمات کلیدی: α-برش، α- شک، متغیر تصادفی فازی، بوت‌استرپ، آزمون فرضیه.
Email(s): .
Mathematics Subject Classification: 62J86; 90B50

۱۹۷۸ [۲] و ۱۹۷۹ [۳] معرفی شده‌است و از آن به بعد پژوهشگران زیادی کم و بیش به این موضوع پرداخته‌اند. سرانجام پوری و رالسکو [۴] در سال ۱۹۸۶ رهیافتی متفاوت از شیوه واکرناک ارائه دادند که مبنای بسیاری از مطالعات بعدی در این زمینه بود. در سال ۱۹۸۷ این مفهوم توسط کروز و میر [۵] به‌صورت ساختاری فرمولبندی گردید. در این رویکرد، متغیر تصادفی فازی به‌عنوان یک درک فازی روی یک متغیر تصادفی با مقادیر حقیقی کلاسیک تلقی می‌شود و متغیر فازی تصادفی برای اولین بار توسط فرون [۱۱]معرفی شد و پس از آن توسط واکرناک [۱۲]، پوری و رالسکو [۱۳] و تعدادی دیگر از پژوهشگران تکمیل و تصحیح شد. به عنوان مثال، حسامیان و چاچی [۱۴] با به کارگیری مفهوم شک و اندازه اعتبار، متغیر تصادفی فازی را معرفی و آن را در آزمون کولموگروف-اسمیرنوف برای داده های فازی به کار گرفتند. همچنین پرچمی [۱۵] نوعی جدید از این مفهوم به نام"متغیر تصادفی قطعه قطعه شده خطی " مبتنی بر تعدادی متناهی از برش تصادفی پیشنهاد داد. در استنباط آماری کلاسیک فرض بر دقیق بودن داده‌ها و مشاهدات، فرضیات و پارامترها می‌باشد، در عمل معمولاً یک یا چند فرضیه از فرضیات مذکور برقرار نمی‌باشند. استفاده از نظریه فازی امکان انجام استنباط در شرایط نادقیق را فراهم می‌سازد. نخستین بار آرنولد در سال ۱۹۹۶، [۶، ۷] به بررسی آزمون فرضیه‌های فازی البته بر پایه مشاهدات دقیق پرداخت. در میان رویکردهای تقریبی دقت برآورد به روش بوت‌استرپ از تمام روش‌های موجود بیشتر می‌باشد که در مطالعات پیشنهادی مونتنگرو و همکاران [۸] ،گونزالز-رودریگوئز و همکاران [۹] و گیل و همکاران [۱۰] بیان شده است. برآورد فاصله اطمینان برای میانگین متغیرهای تصادفی فازی نرمال توسط [۱۷] ارائه شده است. اخیراً آزمون فرض واریانس روی یک و دو نمونه‌ای فازی انجام شده است[۱] .

در این تحقیق فاصله های اطمینان با استفاده از تبدیل داده‌ها و داده‌های اولیه برآورد گردیده و سپس آزمون فرضیه انجام شده است. از آنجا که در روش چاچی، تبدیل داده‌ها به مقدار واریانس تحت فرضیه صفر بستگی دارد، اگر مقدار واریانس کوچک باشد تبدیل داده‌ها دارای نوسانات بیشتری خواهند شد. در

نتیجه طول فاصله اطمینان با تاثیر از آن بزرگتر و نامتقارن شده و آزمون فرضیه به درستی تشخیص نخواهد داد. از طرف دیگر در این روش برای هر مقدار واریانس مفروض در فرضیه صفر باید شبیه‌سازی بوت‌استرپی انجام پذیرد تا پذیرش یا رد فرضیه معلوم گردد. درصورتی که با محاسبه فاصله اطمینان بوت‌استرپی برای واریانس با داده‌های بدون تبدیل براحتی می‌توان آزمون فرضیه را برای هر مقدار مفروض واریانس انجام داد. در بخش (۵) این موضوع به تفصیل بررسی و با ذکر دلایلی استفاده از فاصله‌های اطمینان بدون تبدیل داده‌ها را پیشنهاد می‌کنیم.

۲ تعاریف و مفاهیم اولیه

تعریف ۱.۲. زیر مجموعه‌ی فازی $\tilde{A}$ از مجموعه‌ی مرجع X به‌صورت زیر نمایش داده می‌شود:

$$\{(x, \tilde{A}(x)) \mid x \in X\}$$

که در آن $\tilde{A} : X \to [0, 1]$ تابع عضویت x در $\tilde{A}$ است.

تعریف ۲.۲. فرض کنید X یک مجموعه مرجع و $\tilde{A}$ یک مجموعه فازی از آن باشد، تکیه‌گاه $\tilde{A}$ به‌صورت زیر تعریف می‌شود:

$$supp(\tilde{A}) = \{x \in X \mid \tilde{A}(x) > 0\}.$$

مجموعه $\tilde{A}[\alpha]$ یک α-برش از مجموعه فازی $\tilde{A}$ در $\mathbf{R}$ و به ازای هر $\alpha \in (0, 1]$ می‌گوییم و عبارت است از مجموعه عناصری از $\mathbf{R}$ که درجه عضویت آن‌ها در مجموعه فازی $\tilde{A}$ دست کم به بزرگی α باشد.
در این مقاله $\mathbf{R}$ (مجموعه‌ی اعداد حقیقی) به عنوان مجموعه مرجع در نظر گرفته می‌شود.

تعریف ۳.۲. یک مجموعه فازی $\tilde{A}$ از $\mathbf{R}$ یک عدد فازی نامیده می‌شود اگر برای هر $\alpha \in [0, 1]$ مجموعه $\tilde{A}[\alpha]$ یک بازه فشرده غیر تهی باشد.

مجموعه تمام اعداد فازی از $\mathbf{R}$ را با نماد $F(\mathbf{R})$ نشان می‌دهیم.
عدم دقت یا ابهام می‌تواند با استفاده از یک نمایش خاص از اعداد فازی مورد بررسی قرار گیرد. این نمایش خاص، نمایش LR اعداد فازی نامیده می‌شود؛ و در عمل بسیار مفید است، زیرا می‌توانند با سه عدد حقیقی مشخص شوند. این سه عدد شامل مقدار نما (مرکز یا میانه)، پهنای راست و پهنای چپ می‌باشد. و داریم $\tilde{N} = (n, l, r)_{LR}$ با متغیر مرکزی n که متعلق به اعداد حقیقی است و $l \in \mathbf{R}^+, r \in \mathbf{R}^+$، مشخص می‌شود، L و R توابعی غیرصعودی هستند و داریم: $L : \mathbf{R}^+ \to [\circ, ١]$، $R : \mathbf{R}^+ \to [\circ, ١]$ و $L(\circ) = R(\circ) = ١$

تعریف ۴.۲. فرض کنید $\tilde{A} \in F(\mathbf{R})$ و $x \in \mathbf{R}$. شاخص $C : F(\mathbf{R})\times\mathbf{R} \to [\circ, ١]$، که به شکل زیر تعریف می‌شود

$$C\{\tilde{A} \leq x\} = \frac{sup_{y\leq x}\tilde{A}(y) + ١ - sup_{y>x}\tilde{A}(y)}{٢}$$

درجه اعتبار [۱] را زمانی که $\tilde{A}$ کمتر یا مساوی x است نشان می‌دهد.

تعریف ۵.۲. فرض کنید $\tilde{A} \in F(\mathbf{R})$ و $\alpha \in [\circ, ١]$ عبارت زیر مقدار α شک[۲] $\tilde{A}$ نامیده می‌شود، که نشانگر میزان کوچکی عدد فازی $\tilde{A}$ نسبت به مقدار حقیقی x است.

$$\tilde{A}_\alpha = inf\{x \in \tilde{A}[\circ] : C\{\tilde{A} \leq x\} \geq \alpha\}$$

یکی از مهمترین جنبه‌های تحلیل داده‌های فازی، استفاده از یک فاصله مناسب روی خانواده‌ای از اعداد فازی است. مترها و فاصله‌های گوناگونی توسط نویسندگان مختلف شرح داده شده، و مورد استفاده قرار گرفته است. متر زیر بر اساس مفهوم α- شک و توسط حسامیان و چاچی [۱۴] معرفی شده است.

[1] Credibility Degree
[2] $\alpha - Pessimistic$

تعریف ۶.۲. مقیاس فاصله به‌عنوان یک تابع تعریف شده است $D : F(R) \otimes F(R) \to [0, \infty)$ به‌طوری که آن را با دو عدد فازی مرتبط می‌کند. یعنی اگر $\tilde{A}, \tilde{B} \in F(R)$ فاصله بین A و B به شکل زیر تعریف می‌شود:

$$D(\tilde{A}, \tilde{B}) = \int_0^1 (\tilde{A}_\alpha - \tilde{B}_\alpha)^2 d\alpha$$

اگر $\tilde{A} = (a, l_1, r_1)_{LR}$ و $\tilde{B} = (b, l_2, r_2)_{LR}$ را بدست آوریم که برابر است با

$$D(\tilde{A}, \tilde{B}) = (a-b)^2 + \frac{(l_1 - l_2)^2}{2} \int_0^1 (L^{-1}(\alpha))^2 d\alpha +$$

$$\frac{(r_1 - r_2)^2}{2} \int_0^1 (R^{-1}(\alpha))^2 d\alpha - (a-b)(l_1 - l_2) \int_0^1 L^{-1}(\alpha) d\alpha +$$

$$(a-b)(r_1 - r_2) \int_0^1 R^{-1}(\alpha) d\alpha$$

تعریف ۷.۲. تابع $\tilde{X} : \Omega \to F(R)$ را یک متغیر تصادفی فازی گوییم هر گاه

$$\{(\omega, x) : \omega \in \Omega, x \in \tilde{X}[\alpha](\omega)\} \in F \times B \quad \forall \alpha \in [0, 1]$$

که در آن $F(R)$ مجموعه‌ی همه‌ی اعداد فازی روی R بوده و تابع $\tilde{X}[\alpha](\omega)$ به‌صورت زیر تعریف می‌شود:

$$\tilde{X}[\alpha](\omega) = \{x \in X : \mu_{\tilde{X}(\omega)}(x) \geqslant \alpha\}$$

۳ بوت‌استرپ

بوت‌استرپ یک روش باز نمونه‌گیری برای برآورد پارامترهای مورد نظر است. در شرایطی استفاده از روش بوت‌استرپ تنها راه تحلیل داده‌ها است. بوت‌استرپ یک روش شبیه‌سازی مونت کارلویی[3] است و زمانی که هیچ اطلاعی در مورد توزیع

[3] Monte Carlo Simulation

جامعه‌ی هدف نداریم و فقط یک نمونه تصادفی اولیه از آن استخراج شده است، از آن استفاده می‌شود. در این روش نمونه اولیه را به عنوان جامعه در نظر گرفته و از آن به تعداد مورد نیاز و معمولاً با جایگذاری نمونه استخراج و سپس آماره‌های مورد بحث محاسبه می‌گردند. باید توجه داشت که نمونه مرجع نقش مهمی را در تحلیل درست جامعه هدف بازی می‌کند. اگر این نمونه معرف جامعه نباشد، به تحلیل درستی از جامعه دست نخواهیم یافت. پس از پایان بازنمونه‌گیری به تعداد لازم و بر اساس مقادیر برآورد آماره‌ها، فواصل اطمینان بوت‌استرپی، آزمون فرضیه‌ها محاسبه و تحلیل داده‌ها انجام می‌شود.

۴ آزمون فرضیه یک نمونه‌ای روش چاچی

در این بخش نحوه انجام آزمون فرضیه به روش بوت‌استرپ برای واریانس متغیرهای تصادفی فازی تک نمونه‌ای مورد بررسی قرار می‌گیرد.
فرض کنید $\tilde{X} = (\tilde{X}_۱, \tilde{X}_۲, ..., \tilde{X}_n)$ یک نمونه تصادفی فازی باشد. فرضیه‌های زیر را آزمون می‌کنیم:

$$\begin{aligned} H_\circ &: \quad \sigma^۲ = \sigma_\circ^۲ \\ \mathrm{H}_۱ &: \quad \sigma^۲ \neq \sigma_\circ^۲ \end{aligned}$$

چون اطلاعی از توزیع جامعه‌ی هدف نداریم از توزیع تجربی به عنوان برآورد آن یعنی F_n به‌صورت زیر استفاده می‌کنیم؛ که در آن هر کدام از $\tilde{X}_i$ ها $(i : ۱, ۲, ..., n)$ دارای جرم احتمال $\frac{۱}{n}$ هستند. بنابراین احتمال انتخاب شدن، هر $\tilde{X}_i$ $(i : ۱, ۲, ..., n)$ با احتمال انتخاب شدن در یک نمونه جدیدی که از F_n گرفته می‌شود، یکسان است. هنگامی که از F_n به‌عنوان توزیع شبه‌جامعه استفاده می‌کنیم و نمونه‌گیری مجدد با جایگذاری از نمونه اصلی $\tilde{X} = (\tilde{X}_۱, \tilde{X}_۲, ..., \tilde{X}_n)$ انجام می‌دهیم. و نمونه جدید بدست آمده از این روش را بدین شکل نشان می‌دهیم $\tilde{X}^{*b} = (\tilde{X}_۱^{*b}, \tilde{X}_۲^{*b}, ..., \tilde{X}_n^{*b})$ به طوری که $\tilde{X}^{*b}(b = ۱, ۲, ..., B)$ به معنی b امین مجموعه داده بوت‌استرپ می‌باشد. این B بار تکرار بوت‌استرپ در برآورد توزیع آماره آزمون مورد استفاده قرار می‌گیرد که بعدا معرفی خواهد شد. اما برای انجام

آزمون پارامتری فرضیه به توزیعی نیاز داریم که رفتار جامعه تحت فرضیه $H_\circ$ از آن پیروی می‌کند. در ابتدا توجه کنید که توزیع تجربی (یعنی دادن احتمال $\frac{1}{n}$ به هر عضو $\tilde{X}$) ممکن است یک برآورد مناسب برای توزیع جامعه نباشد و از فرضیه $H_\circ$ تبعیت نکند. به هر حال برای برآوردی از توزیع جامعه که واریانس آن برابر $\sigma_\circ^2$ باشد، یک راه ساده این است که توزیع تجربی را به گونه‌ای تغییر بدهیم که واریانس دلخواه بدست آید. به‌عبارت دیگر، از توزیع تجربی روی مقادیر به‌عنوان توزیع صفر (توزیع احتمالی آماره آزمون، زمانی‌که فرضیه صفر درست باشد) استفاده می‌شود. بنابراین به شیوه زیر متغیر $\tilde{X}_c$ معرفی می‌شود:

$$\tilde{X}_c = (\frac{\sigma_\circ \tilde{X}_1}{S_n(\tilde{X})}, \frac{\sigma_\circ \tilde{X}_2}{S_n(\tilde{X})}, ..., \frac{\sigma_\circ \tilde{X}_n}{S_n(\tilde{X})})$$

در نتیجه واریانس نمونه $\tilde{X}_c$ تحت فرضیه $H_\circ$ برابر $\sigma_\circ^2$ است (یعنی $S_n^2(\tilde{X}_c) = \sigma_\circ^2$). در آزمون فرضیه $H_\circ$، در سطح معنی داری $\gamma \in [0, 1]$ آماره آزمون زیر پیشنهاد می‌شود:

$$T(\tilde{X}) = \frac{(n-1)S_n^2(\tilde{X})}{\sigma_\circ^2}$$

برای برآورد توزیع آماره آزمون $T(\tilde{X})$ تحت فرضیه صفر، B نمونه بوت‌استرپ با جایگذاری از نمونه اصلی $\tilde{X}_c = (\tilde{X}_{c_1}, \tilde{X}_{c_2}, ..., \tilde{X}_{c_n})$ بدست آورده و نمونه‌های جدید بدست آمده را با $\tilde{X}_c^{*1}, \tilde{X}_c^{*2}, ..., \tilde{X}_c^{*B}$ نشان می‌دهیم. برای هر نمونه بوت‌استرپ $\tilde{X}_c^{*b} (b = 1, 2, ..., B)$، آماره یکسان را به‌صورت زیر در نظر می‌گیریم:

$$T(\tilde{X}_c^{*b}) = \frac{(n-1)S_n^2(\tilde{X}_c^{*b})}{\sigma_\circ^2}, \qquad b = 1, 2, ..., B$$

که در آن واریانس بر اساس تعریف (۶.۲) و تفاضل α -شک‌ها محاسبه شده است.

هنگامی که این B مقدار $T(\tilde{X}_c^{*b})$ را بدست آوردیم برای برآورد توزیع $T(\tilde{X})$، از آن استفاده خواهد شد. و سرانجام با استفاده از این مقادیر توزیع جامعه، چندک‌ها و فواصل اطمینان را می‌توان برآورد کرد. برای مثال چندک‌ها را به‌صورت زیر

می‌شوند:

$$\frac{\gamma}{۲} = \frac{\#(T(\tilde{X_c}^{*b}) \leq Q_{\frac{\gamma}{۲}})}{B}$$

در نتیجه آزمون تحت فرضیه $H_۰$ رد می‌شود اگر $T(\tilde{X}) < Q_{\frac{\gamma}{۲}}$ یا $T(\tilde{X}) > Q_{۱-\frac{\gamma}{۲}}$ و در سایر نقاط فرض $H_۰$ پذیرفته می‌شود. برای انجام آزمون فرضیه واریانس یک متغیر تصادفی فازی، الگوریتم زیر پیشنهاد می‌شود:

۱. با توجه به نمونه تصادفی $\tilde{X} = (\tilde{X_۱}, \tilde{X_۲}, ..., \tilde{X_n})$، $\tilde{X_c} = (\tilde{X_{c_۱}}, \tilde{X_{c_۲}}, ..., \tilde{X_{c_n}})$ و $T(\tilde{X}) = \dfrac{(n-۱)S_n^۲(\tilde{X})}{\sigma_۰^۲}$ را محاسبه می‌نماییم.
۲. به کمک نمونه‌گیری با جایگذاری از $\tilde{X_c}$ ، B نمونه بوت‌استرپ $\tilde{X_c}^{*۱}, \tilde{X_c}^{*۲}, ..., \tilde{X_c}^{*B}$ را بدست می‌آوریم.
۳. آماره مرحله ۱ را این بار برای نمونه‌های بوت‌استرپ شده مرحله ۲ به کار می بریم یعنی داریم:

$$T(\tilde{X_c}^{*b}) = \frac{(n-۱)S_n^۲(\tilde{X_c}^{*b})}{\sigma_۰^۲}, \qquad b = ۱, ۲, ..., B.$$

۴. از تکرار مرحله قبل برای برآورد توزیع $T(\tilde{X})$ استفاده می‌کنیم.
۵. چندک‌های بوت‌استرپی $Q_{\frac{\gamma}{۲}}$ و $Q_{۱-\frac{\gamma}{۲}}$ را برآورد می‌نماییم.
۶. اگر $T(\tilde{X}) < Q_{\frac{\gamma}{۲}}$ یا $T(\tilde{X}) > Q_{۱-\frac{\gamma}{۲}}$ فرض $H_۰$ رد می‌شود.

مثال ۱.۴. فرض کنید نمونه تصادفی فازی در جدول (۱) از یک جامعه گرفته شده باشد:

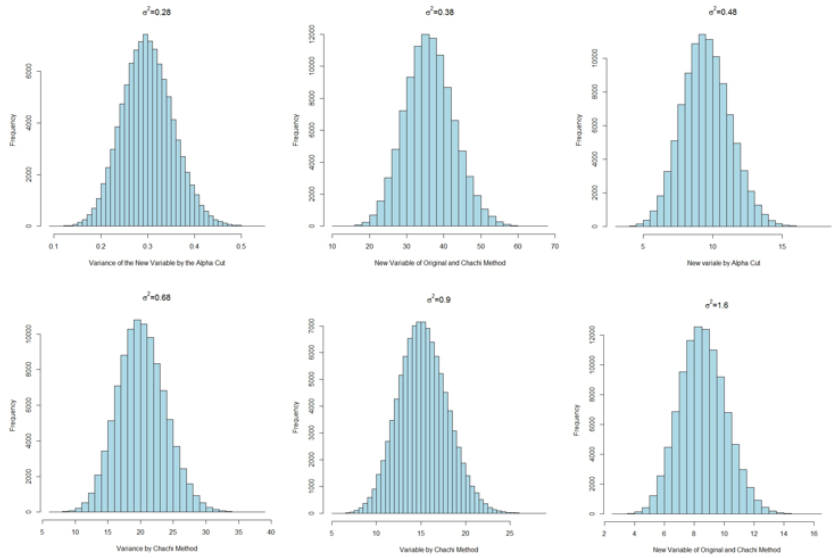
جدول ۱: *جدول داده‌های مثال* ۱.۴

$\tilde{x}_{۱} = (۰/۲۳, ۰/۰۴, ۰/۰۷)_T$	$\tilde{x}_{۱۶} = (۱/۷۸, ۰/۰۴, ۰/۰۶)_T$
$\tilde{x}_{۲} = (۰/۷۶, ۰/۰۵, ۰/۰۲)_T$	$\tilde{x}_{۱۷} = (۱/۹۹, ۰/۰۸, ۰/۰۹)_T$
$\tilde{x}_{۳} = (۰/۹۸, ۰/۱۲, ۰/۰۹)_T$	$\tilde{x}_{۱۸} = (۲/۲۵, ۰/۰۴, ۰/۰۴)_T$
$\tilde{x}_{۴} = (۱/۱۴, ۰/۰۶, ۰/۰۹)_T$	$\tilde{x}_{۱۹} = (۲/۴۵, ۰/۰۱, ۰/۰۸)_T$
$\tilde{x}_{۵} = (۱/۴۶, ۰/۱, ۰/۰۷)_T$	$\tilde{x}_{۲۰} = (۲/۵۷, ۰/۰۷, ۰/۰۲)_T$
$\tilde{x}_{۶} = (۱/۶۹, ۰/۰۵, ۰/۱۲)_T$	$\tilde{x}_{۲۱} = (۰/۶۴, ۰/۱۱, ۰/۰۷)_T$
$\tilde{x}_{۷} = (۱/۹۵, ۰/۰۵, ۰/۱۱)_T$	$\tilde{x}_{۲۲} = (۰/۹۴, ۰/۰۹, ۰/۰۴)_T$
$\tilde{x}_{۸} = (۲/۱۷, ۰/۰۳, ۰/۰۵)_T$	$\tilde{x}_{۲۳} = (۱/۰۸, ۰/۱, ۰/۰۶)_T$
$\tilde{x}_{۹} = (۲/۴, ۰/۰۸, ۰/۱۲)_T$	$\tilde{x}_{۲۴} = (۱/۳۷, ۰/۰۸, ۰/۰۶)_T$
$\tilde{x}_{۱۰} = (۲/۵۱, ۰/۱, ۰/۱۴)_T$	$\tilde{x}_{۲۵} = (۱/۶۴, ۰/۰۲, ۰/۰۸)_T$
$\tilde{x}_{۱۱} = (۰/۴۱, ۰/۰۳, ۰/۰۸)_T$	$\tilde{x}_{۲۶} = (۱/۸۳, ۰/۰۹, ۰/۰۵)_T$
$\tilde{x}_{۱۲} = (۰/۸۶, ۰/۰۸, ۰/۰۴)_T$	$\tilde{x}_{۲۷} = (۲/۰۴, ۰/۱۱, ۰/۰۶)_T$
$\tilde{x}_{۱۳} = (۱/۰۲, ۰/۰۳, ۰/۱)_T$	$\tilde{x}_{۲۸} = (۲/۳۶, ۰/۰۵, ۰/۰۹)_T$
$\tilde{x}_{۱۴} = (۱/۲۳, ۰/۰۳, ۰/۱۴)_T$	$\tilde{x}_{۲۹} = (۲/۴۹, ۰/۱۳, ۰/۰۵)_T$
$\tilde{x}_{۱۵} = (۱/۵۳, ۰/۱۳, ۰/۱۵)_T$	$\tilde{x}_{۳۰} = (۲/۶۱, ۰/۰۸, ۰/۰۶)_T$

بر اساس این مجموعه داده، می‌خواهیم فرضیه‌های زیر را آزمون کنیم.

$$H_{۰}: \quad \text{واریانس جامعه} = \sigma_{۰}^{۲}$$
$$H_{۱}: \quad \text{واریانس جامعه} \neq \sigma_{۰}^{۲}$$

مقادیر با پیروی از مراحل الگوریتم ذکر شده در بالا و با ۲۰۰۰۰۰ بار تکرار بوت‌استرپ $\tilde{X}_c^{*۱}, \tilde{X}_c^{*۲}, ..., \tilde{X}_c^{*B}$ محاسبه شده است. با مرتب سازی این مقادیر به آسانی می‌توانیم چندک γ را برآورد کنیم و همچنین با توجه به سطح معنی‌داری $\gamma = ۰/۰۵$ انتظار داریم برآوردهای کران پایین $(Q_{۰/۰۲۵})$ و کران بالا $(Q_{۰/۹۷۵})$ بترتیب برابر ۵۰۰۰امین و ۱۹۵۰۰۰امین مقدار مرتب شده $T(\tilde{X}_c^{*b})$ باشند. (چون $B \times \frac{\gamma}{۲} = ۲۰۰۰۰۰ \times ۰/۰۲۵ = ۵۰۰۰$) در عمل $Q_{۰/۰۲۵}$ و $Q_{۰/۹۷۵}$ را طوری بدست می‌آوریم که بترتیب در رابطه‌های $\frac{\#(T(\tilde{X}_c^{*b}) \leq Q_{۰/۰۲۵})}{B} = ۰/۰۲۵$ و $\frac{\#(T(\tilde{X}_c^{*b}) \leq Q_{۰/۹۷۵})}{B} = ۰/۹۷۵$ صدق کنند. برآورد $\sigma_{۰}^{۲}$ را در جدول (۲) ملاحظه می‌کنید. در شکل (۱) هیستوگرام توزیع‌های برآورد شده از آماره نمونه $T(\tilde{X})$ برای مقادیر مختلف $\sigma_{۰}^{۲}$ جدول (۲) ارایه شده است.



شکل ۱: *هیستوگرام توزیع‌های برآورد آماره $T(\tilde{X})$ برای مقادیر مختلف* σ_0^2 *در جدول*

۵ تحلیل آزمون فرضیه

از آنجا که در تعریف متغیر تصادفی خی‌دو، متغیرهای اولیه باید دارای توزیع نرمال باشند. از طرفی الزاماً متغیرهای تصادفی فازی دارای توزیع نرمال نیستند و برای این منظور از تقریب استفاده می کنیم که با توجه به اندازه حجم نمونه بدون خطا نخواهد بود. در نتیجه برای تحلیل روش‌های بالا، با روش بوت‌استرپ واریانس داده‌های بدون تبدیل را محاسبه و فاصله اطمینان بوت‌استرپی برای واریانس برآورد می کنیم. سپس این دو روش را مقایسه نموده و به بیان برخی از ویژگی‌ها، مزایا و معایب آنها می‌پردازیم. با در نظر گرفتن حالت یک نمونه‌ای روش چاچی، فرض کنید $error = ۰٫۰۵$ و $N = ۲۰۰۰۰۰$ باشد، واریانس نمونه اصلی برابر $S^۲ = ۰٫۴۸۹۵۲۵۵$ است. با توجه به برآورد فاصله اطمینان ۹۵ درصدی برای واریانس داده‌های جدول (۱) به روش چاچی برابر $(۰٫۳۱۲, ۰٫۶۴۸)$ است. لازم به ذکر است روش تغییر متغیر چاچی روشی بسیار جالب و در واقع مستقل از

واحد اندازه‌گیری است و همیشه دارای میانگین $(n-۱)$ است. در نتیجه فاصله اطمینان بدست آمده در یک نمونه تصادفی به مرکزیت $(n-۱)$ است. باید توجه داشت که برای تعریف متغیر تصادفی خی‌دو، متغیر اولیه باید دارای توزیع نرمال باشد که در اینجا از توزیع واقعی آن اطلاع نداریم. از طرفی توزیع

$$T(\tilde{X})=\frac{(n-۱)S_n^۲(\tilde{X})}{\sigma_۰^۲}$$

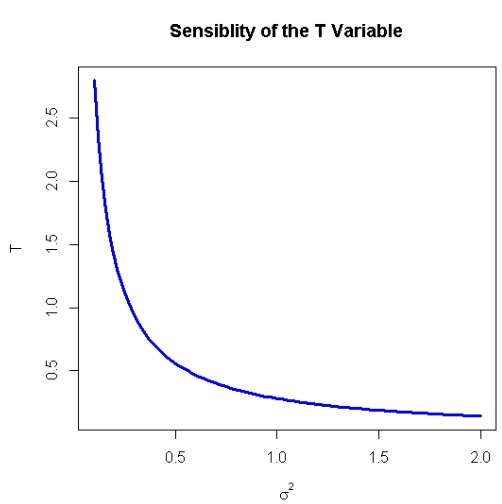
با توجه به مقدار n تقریباً دارای توزیع خی‌دو می‌باشد. با توجه به شکل (۲)، مقادیر برآورد شده به روش چاچی برحسب واریانس مفروض (تحت فرضیه صفر) ملاحظه می‌شود که در این تغییر متغیر، برآوردگر برای مقادیر کوچک واریانس مفروض حساسیت بیشتری خواهد داشت و مقدارش با شدت بیشتری بزرگ خواهد شد در نتیجه دچار اشتباه شده و فرضیه صفر را رد می‌کند. از طرف دیگر برای مقادیر بزرگ واریانس مفروض (تحت فرضیه صفر) به دلیل حساسیت کم دچار اشتباه شده و فرضیه صفر را به نادرستی می‌پذیرد. برای پی بردن به این موضوع، اگر نرخ تغییرات متغیر چاچی را در نظر بگیریم که همان مشتق متغیر می‌باشد داریم:

$$\frac{\partial T(\tilde{X})}{\partial\sigma_۰^۲}=-\frac{(n-۱)S_n^۲(\tilde{X})}{\sigma_۰^۴}$$

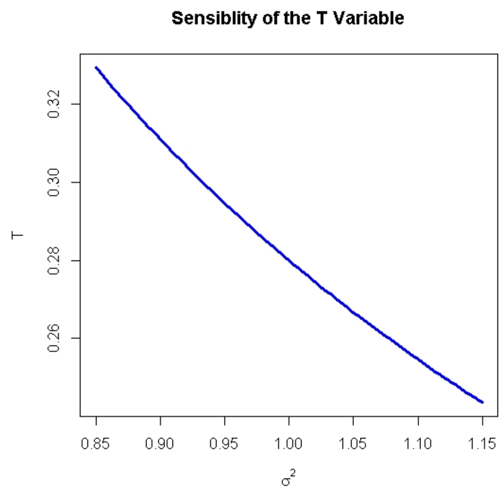
ملاحظه می‌شود که تغییرات منفی و متناسب با معکوس توان دوم واریانس می‌باشند. در نتیجه روش چاچی حساس به مقادیر کوچک واریانس مفروض است، شاید حول مرکز عدد یک، حساسیت آن متقارن گردد (شکل ۳) و با روش بدون تبدیل جوابی تقریبا یکسان داشته باشد. به نظر می‌رسد که این روش مادامی که واریانس داده‌ها نزدیک به عدد یک باشد بسیار خوب عمل می‌کند. لذا برای رهایی از این حساسیت پیشنهاد می‌گردد چون توزیع داده‌ها را نمی‌دانیم، برای آزمون نمودن واریانس از روش بوت‌استرپ بدون تبدیل استفاده گردد. همچنین در این روش یک فاصله اطمینان بدست آورده و سپس هر مقدار مفروضی برای واریانس را می‌توان آزمون کرد و نیازی به شبیه‌سازی با مقادیر جدید واریانس‌های مفروض نیست.

جدول ۲: *نتایج آزمون به دو روش برای مقادیر مختلف* σ_0^2 *در مثال* ۱.۴

σ_0^2	S_{ch}^2	$CI_{0.95}(\sigma_{ch}^2)$		$Result_{ch}$	$CI_{0.95}(\sigma^2)$		$Result$
۰٫۲۸	۵۰٫۷۰	۱۸٫۵۳	۳۸٫۴۴	$reject H_0$	۰٫۳۱۲	۰٫۶۴۸	$reject H_0$
۰٫۳۱	۴۵٫۷۹	۱۸٫۵۳	۳۸٫۴۴	$reject H_0$			$reject H_0$
۰٫۳۲	۴۳٫۸۲	۱۸٫۵۰	۳۸٫۴۰	$reject H_0$			$accept H_0$
۰٫۳۸	۳۷٫۳۶	۱۸٫۵۰	۳۸٫۴۰	$accept H_0$			$accept H_0$
۰٫۴۸	۲۹٫۵۸	۱۸٫۵۲	۳۸٫۴۱	$accept H_0$			$accept H_0$
۰٫۶۸	۲۰٫۸۸	۱۸٫۵۲	۳۸٫۴۰	$accept H_0$			$reject H_0$
۰٫۷	۲۰٫۲۸	۱۹٫۶۱	۲۸٫۰۴	$accept H_0$			$reject H_0$
۰٫۹۰	۱۵٫۷۷	۱۸٫۵۲	۳۸٫۴۰	$reject H_0$			$reject H_0$
۱٫۶۰	۸٫۸۷	۱۸٫۵۱	۳۸٫۴۴	$reject H_0$			$reject H_0$



شکل ۲: *نمودار تغییرات تبدیل چاچی نسبت به مقادیر مفروض واریانس*



شکل ۳: *نمودار حساسیت برآوردگر واریانس*

همان طور که در جدول (۴) ملاحظه می‌شود با تغییر مقادیر آلفا نه تنها کرانهای اطمینان در تبدیل چاچی تغییر نمی‌کند بلکه مقادیر واریانس اولیه و واریانس چاچی نیز تغییر نمی‌کند. به‌عنوان مثال برای مقادیر $\alpha = \{۰/۵, ۱\}$، واریانس داده‌های اولیه برابر با ۰/۱۴۹۳۴۹۱ و فاصله اطمینان ۹۵ درصد برابر $(۰/۰۸۷۴۵۰۱, ۰/۲۲۰۷۳۷)$ است.

جدول ۳: *خلاصه نتایج روش چاچی تک نمونه*

$\sigma^۲$	α_{cut}	$CI_{۰/۹۵}(\sigma^۲_{Tr})$		$\hat{\sigma^۲_{Tr}}$	$CI_{۰/۹۵}(\sigma^۲_{ch})$		$\hat{\sigma^۲_{ch}}$	$\hat{\sigma^۲_{or}}$
۰/۲۸	۰/۱	۱۸/۵۳	۳۸/۴۴	۵۰/۷۰				
۰/۳۱	۰/۳	۱۸/۵۳	۳۸/۴۴	۴۵/۷۹				
۰/۳۲	۰/۵	۱۸/۵۳	۳۸/۴۴	۴۳/۸۲	۰/۳۱۲	۰/۶۴۸	۰/۴۸۹	۰/۱۴۹
۰/۳۸	۰/۷	۱۸/۵۰	۳۸/۴۰	۳۷/۳۶				
۰/۴۸	۰/۹	۱۸/۵۲	۳۸/۴۱	۲۹/۵۷				

۶ آزمون فرضیه یک نمونه‌ای به روش آلفا برش

در این روش به منظور اندازه‌گیری فاصله اجزا در $F(\mathbf{R})$ «خانواده همه اعداد فازی» شکل ساده‌ای از فاصله که متر برتولوزا [۴] نامیده می‌شود، معرفی می‌گردد، که با $d_\theta(\theta > ۰)$ نمایش داده می‌شود و به شکل زیر تعریف شده است.

$$d^{۲}_\theta(A,B) = \int_{[۰,۱]} (mid(A_\alpha) - mid(B_\alpha))^{۲} d\alpha +$$

$$\theta \int_{[۰,۱]} (spr(A_\alpha) - spr(B_\alpha))^{۲} d\alpha$$

که در آن:

$$mid(A_\alpha) = \frac{۱}{۲}(\underline{a}_\alpha + \overline{a}_\alpha) \quad , \quad spr(A\alpha) = \frac{۱}{۲}(\overline{a}_\alpha - \underline{a}_\alpha)$$

برای هر $\alpha \in [۰,۱]$، نقطه میانی A_α و $spr(A_\alpha)$ گستردگی (یا شعاع) A_α نامیده می‌شود. توجه داشته باشید که اندازه پذیری بورل d_θ بستگی به مقدار θ زمانی که $\theta > ۰$ است ندارد، زیرا تمام فواصل متری d_θ توپولوژی‌های یکسانی را در $F(\mathbf{R})$ القاء می‌نمایند.

فرض کنید $X = (X_۱, X_۲, ...X_n)$ نمونه‌ای از یک متغیر تصادفی فازی باشد، میانگین و واریانس آن به شکل زیر تعریف می‌شود.

$$\overline{X}_n = \frac{۱}{n}(X_۱ \oplus X_۲ \oplus ... \oplus X_n)$$
$$\sigma^{۲} = \frac{۱}{n}\Sigma_i^n d^{۲}_\theta(X_i, \overline{X}_n)$$

اکنون آزمون فرضیه مطرح شده در مثال (۱.۴) را به روش آلفا برش بکار می‌بریم. در جدول (۴) نتایج آزمون برای مقادیر مختلف $\sigma^{۲}_۰$ و در شکل (۴) هیستوگرام توزیع برآورد شده از آماره $T(\tilde{X})$ نمایش داده شده است. لازم به ذکر است برای بررسی موضوع از الگوریتم ذکر شده در بخش (۴) استفاده می کنیم و سپس چندکها و فواصل اطمینان مانند بخش (۴) محاسبه می‌شوند.

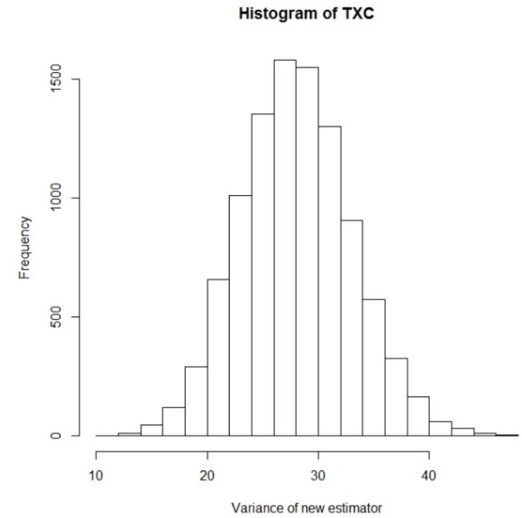
[۴] Bertoluzza

۷ تحلیل آزمون برای روش آلفا برش

نتایج حالت تک نمونه‌ای روش آلفا برش در جدول (۴) داده شده است. همان طور که قبلا ملاحظه شد، آماره آزمون در روش چاپچی وابسته به مقدار $\sigma_{\circ}^{۲}$ می‌باشد اما در آلفا برش حساسیت کمتری نشان می‌دهد. با توجه به محاسبه فاصله اطمینان به روش بوت‌استرپ آلفا برشی، نیاز به شبیه‌سازی و محاسبه مجدد تبدیل داده‌ها به ازای مقادیر واریانس نمی‌باشد.

جدول ۴: *خلاصه نتایج روش آلفا برش تک نمونه‌ای*

$\sigma_{\circ}^{۲}$	$S_{\alpha}^{۲}$	$CI_{\circ/۹۵}(\sigma_{\alpha}^{۲})$		$Result_{\alpha}$	$CI_{\circ/۹۵}(\sigma^{۲})$		$Result.Boot_{\alpha}$
۰/۲۸	۱۶/۳۷	۱۶/۹۳	۲۹	$rejectH_{\circ}$			$rejectH_{\circ}$
۰/۳۱	۱۴/۷۸	۱۶/۹۳	۲۸/۰۳	$rejectH_{\circ}$			$rejectH_{\circ}$
۰/۳۸	۱۲/۰۶	۱۶/۹۳	۲۸/۰۴	$rejectH_{\circ}$			$rejectH_{\circ}$
۰/۴۸	۹/۵۴	۱۶/۹۳	۲۸/۰۱	$rejectH_{\circ}$	۰/۱۱۷	۰/۲۴۹	$rejectH_{\circ}$
۰/۶۸	۶/۷۳	۱۶/۹۳	۲۸/۰۱	$rejectH_{\circ}$			$rejectH_{\circ}$
۰/۷	۶/۵۵	۱۶/۹۳	۲۸/۰۴	$rejectH_{\circ}$			$rejectH_{\circ}$
۰/۹۰	۵/۰۹	۱۶/۹۳	۲۸/۰۱	$rejectH_{\circ}$			$rejectH_{\circ}$
۱/۶۰	۲/۸۷	۱۶/۹۳	۲۸/۰۲	$rejectH_{\circ}$			$rejectH_{\circ}$



شکل ۴: *هیستوگرام توزیع برآورد شده از آماره نمونه* $T(\tilde{X})$

۸ نتیجه‌گیری

وقتی توزیع جامعه مشخص نیست، روش بوت‌استرپ یکی از روش‌های مناسب تحلیل داده‌ها است. با توجه به اینکه حجم نمونه کوچک است (n=۱۵) در مقاله چاچی با فرض نرمال بودن داده‌ها متغیر تصادفی خی‌دو تعریف شده است. اگر چه چاچی با این تغییر متغیر به پارامترهای از پیش تعیین شده‌ای رسیده است، ولی از بستگی تبدیل به مقدار واریانس تحت فرضیه صفر و حساسیت آن به مقادیر کوچک واریانس نباید غافل شد. از طرفی در این روش برای هر مقدار از واریانس باید تبدیل مذکور، روش بوت‌استرپ و فرایند آزمون فرضیه را مجدداً انجام داد که کاری غیر ضروری و زمانبر است. روش جایگزین، استفاده از بوت‌استرپ برای داده‌های اولیه و محاسبه یک فاصله اطمینان را پیشنهاد می‌کند و پس از آن برای هر مقدار از واریانس، آزمون فرضیه را می‌توان بررسی کرد. در روش آلفا برش جدول (۳) ملاحظه می‌شود که تغییر آلفا در برآورد فاصله اطمینان اثری ندارد و محاسبه یک فاصله اطمینان و آزمون فرضیه برای واریانس مجموعه داده‌ها کفایت می‌کند.

مراجع

[۱] پرچمی عباس، متغیرهای تصادفی فازیLR، سیستم های فازی و کاربردها، دوره ۲ ،شماره ۱ - شماره پیاپی ۳ ،بهار و تابستان ۱۳۹۸ صفحه ۸۹ -۱۰۷.

[2] Zadeh, L., A (1965), Fuzzy Set, Iformation and Ccontrol, 8, 338–353

[3] Kwakernaak H. (1978) Fuzzy random variables, part I: Definitions and theorems. Inform Sci 19:1-15

[4] Kwakernaak H (1979) Fuzzy random variables, part II: Algorithms and examples for the discrete case. Inform Sci 17:253-278

[5] Puri ML, Ralescu DA (1986) Fuzzy random variables. Journal of Mathematical Analysis and Applications 114:409-422..

[6] R. Kruse, and K. D. Meyer (1987) Statistics with Vague Data, Reidel Publishing Company, Dordrecht, .

[7] Arnold, B. F. (1996) An apporach to fuzzy hypothesis testing, Metrika 44:119-126.

[8] Arnold, B. F. (1998) Testing fuzzy hypothesis with crisp data, Fuzzy Sets and Systems 94:323-333.

[9] M. Montenegro, A. Colubi, M. R. Casals, and M. A. Gil (2004) Asymptotic and bootstrap techniques for testing the expected value of a fuzzy random variable, Metrika 59:31–49.

[10] G. Gonz alez-Rodr ıguez, M. Montenegro, A. Colubi, and M. A. Gil (2006) Bootstrap techniques and fuzzy random variables: Synergy in hypothesis testing with fuzzy data, Fuzzy Sets and Systems 157:2608–2613.

[11] M. A. Gil, M. Montenegro, G. Gonz alez-Rodr ıguez, A. Colubi, and M. R. Casals (2006) Bootstrap approach to the multi-sample test of means with imprecise data, Computational Statistics and Data Analysis 51:148–162.

[12] Féron R. (1986), Ensembles Aléatoires Flous, C.R. Academic Science Paris Ser. A 575, 1986, 227 -226.

[13] Kwakernaak H. (2002), Fuzzy random variables - I. definitions and theorems, Information Sciences, 0 - 52 , 0237

[14] Puri M.L. and Ralescu D.A. (2002), Fuzzy random variables, Journal of Mathematical Analysis and Application, 222 –255 , 0276.

[15] Hesamian G., and Chachi J. (2013), Two-sample Kolmogorov-Smirnov fuzzy test for fuzzy random variables , Statistical Papers, vol. 56, pp.61–82, 2013

[16] Chachi, J., Taheri, S.M. (2011), Fuzzy confidence intervals for mean of Gaussian fuzzy random variables, Expert Systems with Applications, 38: 5240-5244.

[17] Chachi J. (2017), Bootstrap Approach to the One-Sample and Two-Sample Test of Variances of a Fuzzy Random Variable, Stat., Optim. Inf. Comput., Vol. 5, pp 188–199.