سومین کنفرانس انجمن رمز ایران
دانشگاه صنعتی اصفهان
۱۶و۱۷ شهریورماه ۱۳۸۴

انجمن رمز ایران

ارزیابی فازی مولدهای شبه‌تصادفی

حمید ملا
دانشگاه صنعتی اصفهان
mala@ec.iut.ac.ir

محمد دخیل‌علیان
دانشگاه صنعتی اصفهان
mdalian@cc.iut.ac.ir

چکیده

مولدهای شبه‌تصادفی از اهمیت ویژه‌ای در رمزنگاری به خصوص جهت تولید دنبالهٔ کلید اجرایی در سیستمهای رمز پی‌درپی برخوردارند. در دهه‌های اخیر آزمونهای آماری متعددی به منظور ارزیابی این‌گونه مولدها ارائه شده است. در روش متداول ارزیابی یک مولد شبه تصادفی، ابتدا دنباله‌های متعدد تولید شده توسط مولد تحت آزمونهای آماری نظیر آزمون فرکانس، آزمون رنها[1]، آزمون پوکر[2] و... قرار می‌گیرند. سپس نتایج این آزمونها خود تحت آزمونهای کلی‌تر همچون آزمون[3] KS، آزمون یکنواختی مقادیر احتمال ویا آزمون سازگاری رفتار مولد با مدل فرض شده، قرار می‌گیرند و در نهایت نتیجه چنین آزمونی به صورت رد یا قبول مولد به عنوان یک مولد شبه‌تصادفی می‌باشد. در این مقاله با در اختیار داشتن نتایج حاصل از اعمال آزمون آماری T بر روی دنباله‌های تولید شده توسط مولد، فرضیهٔ شبه‌تصادفی بودن به صورت یک فرضیهٔ فازی مدل می‌شود و سپس آزمون سازگاری جهت بررسی این فرضیهٔ فازی مورد استفاده قرار می‌گیرد. نتیجهٔ این روش بیان میزان شبه‌تصادفی بودن مولد به صورت یک تابع عضویت می‌باشد.

واژه‌های کلیدی: مولد شبه‌تصادفی، آزمون آماری، آزمون فرض فازی، آزمون سازگاری فازی

۱- مقدمه

امروزه اهمیت مسأله ارزیابی مولدهای شبه‌تصادفی به ویژه در زمینه امنیت تبادل اطلاعات، مدام در حال افزایش است. در واقع بسیاری از روشهای امن‌سازی اطلاعات و ارتباطات همچون رمزنگاری کلید عمومی، امضای دیجیتال، پروتکلهای تبادل کلید و احراز اصالت و به خصوص سیستمهای رمز پی‌درپی مبتنی بر روشهای تولید امن دنباله‌های باینری شبه‌تصادفی هستند[۱]. تاکنون آزمونهای آماری متعددی جهت ارزیابی مولدهای گوناگون ارائه شده است[۱, ۲, ۳]. در واقع در آزمونهای آماری ویژگیهایی چون تعداد یکها و صفرها، میزان پیچیدگی، قابلیت فشرده‌شدن، میزان خودهمبستگی و... در مورد یک دنبالهٔ ایده‌آل که

[1] *Runs test*
[2] *Poker test*
[3] *Kolmogrov-Smirnov test*

بیتهای آن مستقل از هم و دارای توزیع یکنواخت باشند، مورد بررسی قرار می‌گیرد و مدل احتمالی هر یک از این ویژگیها بدست می‌آید. با بدست آوردن مدل احتمال یک ویژگی خاص برای انجام آزمون T ،آماره[1]ای از روی دنبالهٔ تحت آزمون ساخته می‌شود. با انتخاب سطح اهمیت[2] α دنباله مورد نظر تحت آزمون فرض آماری قرار می‌گیرد. در آزمون فرض، فرضیهٔ صفرِ H_0 که بیان می‌کند دنباله از توزیع یکسان و یکنواخت تولید شده، در برابر فرض مقابلِ H_1 که بیان می‌کند دنباله طبق یک قانون معین وقطعی بدست‌آمده‌است، مورد بررسی قرارمی‌گیرد. اگر کمیت مقدار احتمال[3] که از روی آمارهٔ آزمون ساخته می‌شود بزرگتر از α باشد، دنباله مورد قبول واقع ودر غیر اینصورت رد می‌شود. برای اظهار نظر دربارهٔ مولد بایستی دنباله‌های متعدد تولید شده توسط آن تحت آزمون آماریِ T قرارگیرند و سپس مجموعهٔ نتایج بدست آمده تحت یکی از آزمونهای کلی قرارگیرد. برخی از این آزمونهای کلی عبارتند از: آزمون KS که در مورد تطابق تابع توزیع تجمعی آماره‌های حاصل از تکرارآزمون T بر روی دنباله‌های تولید شده توسط مولد با مدل احتمالی بدست‌آمده، بحث می‌کند، آزمون سازگاری رفتار مولد با مدل فرض شده در آزمون T که در بخش بعد بیشتر به آن خواهیم پرداخت و نیز آزمون یکنواختی مقادیر احتمال که تطابق مقادیر احتمال بدست‌آمده از چندین بار اجرای آزمون T بر روی دنباله‌های مختلف مولد را با توزیع یکنواخت مورد بررسی قرارمی‌دهد.

به هرحال، صرف‌نظر از نوع آزمون T و نوع آزمون نهایی بکاررفته جهت ارزیابی مولد، نتیجه این ارزیابی به دو حالت رد یا قبول خلاصه می‌شود. اما در بسیاری مواقع، ابهام در ارزیابی آماری دنباله‌ها و منابع تولید دنباله‌های شبه‌تصادفی با قبول و یا رد دنباله در آزمون برطرف نمی‌گردد؛ چراکه ممکن است یک دنباله با توجه به مقدار α وضعیت کاملاً نامطلوبی داشته باشد و دنبالهٔ دیگر وضعیت خوبی داشته باشد اما به صورت مرزی در آزمون رد شود، ولی از دیدگاه آزمون آماری هر دو یکسان باشند. به همین خاطر رفتن به سمت برخورد فازی با دنباله‌ها ومنابع تولید آنها به منظور ارزیابی دقیقتر کیفیت مولدهای شبه‌تصادفی، برخوردی منطقی و صحیح به نظر می‌رسد. بر همین اساس در این مقاله پس از بیان آزمون سازگاری رفتار مولد با مدل فرض شده در آزمون T، با استفاده از ایدهٔ آزمون فرض فازی برای فرضیه‌های فازی، رویکرد جدیدی جهت ارزیابی مولدهای شبه‌تصادفی ارائه می‌کنیم. نتیجه استفاده از این روش بدست آوردن تابع عضویت[4] میزان تصادفی بودن بر حسب مقدار احتمال حاصل از آزمون سازگاری است، یعنی به هر مولد، عددی بین صفر تا یک که مشخص کنندهٔ میزان تصادفی بودن مولد است نسبت داده خواهد شد.

۲- آزمون سازگاری رفتار مولد با مدل فرض شده در آزمون *T*

می‌دانیم که در آزمون فرض، α برابر با احتمال رد شدن فرض صفر به شرط صحیح بودن آن است. حال فرض کنید آزمون T که همان آزمون فرکانس، آزمون پوکر، آزمون رنها یا... است، بر روی N دنباله مجزا از یک مولد اعمال شود. اگر نتیجهٔ قبول شدن یک دنباله را با یک ونتیجهٔ رد شدن آن را با صفر نشان دهیم، در نتیجه می‌توان نتایج حاصل از آزمون T بر روی N دنبالهٔ مجزا به طول n از مولد را به صورت زیر نمایش داد:

$$T^N = t_1, t_2, \ldots, t_N \qquad , t_i \in \{0,1\} \qquad (۱)$$

اگر فرضیهٔ H_0 صحیح باشد، با توجه به تعریف α وفرض استقلال t_i ها داریم:

---

[1] *Statistic*
[2] *Level of significance*
[3] *P-value*
[4] *Membership function*

$$P\{t_i=0\}=\alpha \quad , \quad P\{t_i=1\}=1-\alpha \quad , \ i=1,2,...,N \qquad (۲)$$

و با توجه به قضیهٔ حد مرکزی، S_N یعنی تعداد دنباله‌هایی که از بین N دنباله از آزمون T عبور کرده‌اند، دارای توزیع نرمال با میانگین $N(1-\alpha)$ و واریانس $N\alpha(1-\alpha)$ خواهد بود.

$$S_N = t_1 + t_2 + \cdots + t_N \qquad (۳)$$

اکنون می‌توان آزمون زیر را با انتخاب α جدید که آن را با α_{new} نمایش می‌دهیم، به انجام رسانید.آماره آزمون که انتظار می‌رود برای یک مولد شبه‌تصادفی دارای توزیع نرمال استاندارد باشد، به صورت زیر خواهد بود:

$$T_N = \frac{S_N - N(1-\alpha)}{\sqrt{N\alpha(1-\alpha)}} \qquad (۴)$$

مقدار احتمال متناظر با آمارهٔ مشاهده شده نیز که آن را با PV نشان می‌دهیم به صورت زیر بدست می‌آید.

$$PV = 2\{1-\varphi(T_N)\} \qquad (۵)$$

که در آن $\varphi(\cdot)$ تابع توزیع تجمعی متغیر تصادفی نرمال استاندارد می‌باشد. نتیجهٔ آزمون سازگاری به این نحو خواهد بود که اگر مقدار احتمال بدست آمده بزرگتر از α_{new} باشد، مولد از آزمون عبور می‌کند و به عنوان یک مولد شبه تصادفی شناخته می‌شود که این به معنای سازگار بودن رفتار مولد با ویژگی مدل شده در آزمون T می‌باشد. اما چنانچه مقدار احتمال از α_{new}کوچکتر باشد، مولد رد می‌شود [۱ ,۴]. در ادامه پس از بیان مقدمات لازم، ایدهٔ فازی کردن آزمون سازگاری را بیان می‌کنیم.

۳- آزمون سازگاری فازی

همانطور که بیان شد، در آزمون سازگاری معمولی نیز همچون سایر آزمونهای آماری مشابه، نتیجه آزمون تنها به صورت رد یا قبول مولد تحت تست اظهار می‌شود. از سوی دیگر پارامترهایی نظیر طول وتعداد دنباله‌ها و یا مقدار اولیه[1] مولد نیز در میزان دقت آزمون مؤثرند اما تأثیر آنها به طور کامل در روند ذکر شده لحاظ نمی‌شود. به عنوان مثال با افزایش تعداد دنباله‌ها، تقریب زدن آماره آزمون در رابطه (۴) با توزیع نرمال استاندارد دقیقتر خواهد شد. این نکته که به خصوص در مواردی که تعداد دنباله‌ها کم باشد اهمیت بیشتری خواهد داشت، در روند آزمون سازگاری لحاظ نشده است. از طرف دیگر نتایج آزمون T که به عنوان ورودی آزمون سازگاری هستند، از دقت کافی برخوردار نیستند، زیرا هریک از مؤلفه‌های دنباله‌ی T^N در حقیقت کوانتیزه شده‌ی مقدار احتمال حاصل از انجام آزمون Tبر روی یکی از دنباله‌های خروجی مولد هستند. بنابراین می‌توان آنها را به عنوان داده‌های نادقیق برای آزمون سازگاری تلقی نمود. وجود این قبیل نایقینی[2]ها و نیاز به تصمیم‌گیری دقیقتر به جای تصمیم‌گیریِ دو ارزشی، در مورد مولدها ما را به سوی استفاده از روشهای تصمیم گیریِ فازی سوق می‌دهد.

اساس هر آزمون آماری یک آزمون فرض است که فرضیهٔ تصادفی بودن را در مقابل فرضیهٔ قطعی بودن توزیع دنباله مورد بررسی قرار می‌دهد. ایدهٔ ما برای حصول نتیجهٔ فازی درباره کیفیت یک مولد این است که تمام نایقینی‌های فوق‌الذکر را به صورت یکجا در فرضیهٔ تحت آزمون جمع کنیم. به عبارت دیگر پیشنهاد می‌کنیم که فرضیهٔ مورد نظر را به صورت فازی بیان کنیم اما داده‌ها را که همان دنبالهٔ T^N می‌باشد به صورت دقیق لحاظ کنیم. سپس با استفاده از آزمون فرض فازی[3] به یک بیان فازی دربارهٔ کیفیت مولد خواهیم رسید. آزمون فرض فازی از مباحث نسبتاً جدید در آمار فازی است اما تاکنون به شاخه‌ای از آن که مشاهدات (داده‌ها)

---

[1] *Seed*
[2] *Uncertainty*
[3] *Fuzzy Hypothesis Testing*

دقیق و غیر فازی اما فرضیه‌ها فازی باشند کمتر توجه شده است. مسألهٔ آزمون فرضها هنگامی که مشاهدات دقیق باشند اما خود فرضها مبهم و نادقیق باشند، نخستین بار توسط آرنولد[۵ ،۶] مطالعه شد. واتانابه و ایمایزومی نیز با شیوه‌ای متفاوت این مسأله را مورد بررسی قرار دادند[۷]. در روش آنها نتیجهٔ آزمون نیز به صورت فازی بیان می‌شود. در ادامه نخست، فازی کردن فرضیهٔ شبه تصادفی بودن مولد را مطرح می‌کنیم و سپس در بخش بعد آزمون سازگاری با فرضیهٔ فازی را بیان مینماییم.

۳-۱ فازی کردن فرضیه‌ها

در آزمونهای آماری متداول، فرضیه‌ها به صورت قطعی و غیرفازی هستند ونتیجه آزمون نیز به صورت رد یا قبول، یعنی دوارزشی و غیرفازی بیان می‌شود. پیشنهاد ما برای رسیدن به یک ارزیابی واقعی‌تر درباره‌ی مولد، فازی کردن فرضیه‌هاست. در آزمون سازگاری، فرضیه صفر به صورت زیر بیان می‌شود:

H_0 : تعداد دنباله‌های عبور کرده از آزمون T برابر با $N(1-\alpha)$ است.

برای فازی کردن این فرضیهٔ دقیق، نخست آن را به صورت زیر بیان می‌کنیم:

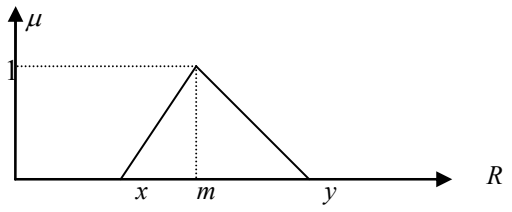
$H_{f,0}$: تعداد دنباله‌های عبور کرده از آزمون T تقریباً برابر با $N(1-\alpha)$ است.

سپس برای مشخص کردن مفهوم "تقریباً $N(1-\alpha)$"آن را با یک عدد فازی که به صورت زیر تعریف می‌شود، نمایش می‌دهیم.

تعریف عدد فازی: مجموعه فازی $A\in IF(IR)$ یک عدد (حقیقیِ) فازی نامیده می‌شود، اگر وتنها اگر A محدب باشد و دقیقاً یک عدد حقیقی a وجود داشته باشد که $\mu_A(a)=1$ باشد [۸].

در ادامه برای سادگی از اعداد فازی مثلثی استفاده می‌کنیم و عدد فازی مثلثی تقریباً m یا به عبارت دیگر مجموعه‌ی اعداد نزدیک به m را به صورت زیرنمایش می‌دهیم:

$$\tilde{m}=(x,m,y)_T \quad , \quad x\le m\le y \qquad (۶)$$



شکل۱: عدد فازی مثلثی $\tilde{m}$

که در آن درجهٔ تعلق m به مجموعهٔ اعداد نزدیک به عدد m برابر واحد است و اعداد کوچکتر از x یا بزرگتر از y به این مجموعه تعلق ندارند و درجهٔ عضویت آنها برابر صفر است. هر کدام از اعداد کوچکتر از y و بزرگتر از x نیز دارای درجهٔ عضویتی بین صفر تا یک هستند.

۳-۲ بیان آزمون سازگاری فازی

گفتیم که در آزمون سازگاری فاری، هدف بررسی فرضیهٔ فازیِ برابری تعداد دنباله‌های عبور کرده از آزمون T با عدد فازی "تقریباً $N(1-\alpha)$" می‌باشد. برای انجام این آزمون، نخست آزمون معمولیِ زیر با سطح اهمیت مشخصِ α_{new} وپارامتر ψ انجام می‌شود:

$$H_0(\psi): p=\psi \qquad , \qquad H_1(\psi): p\neq\psi$$

که در آن p احتمال عبور یک دنباله تولید شده توسط مولد از آزمون T است. در این حالت آماره آزمون تابعی از ψ خواهد بود که آن را با $T_N(\psi)$ نمایش می‌دهیم و طبق رابطهٔ (۴) برابر است با :

$$T_N(\psi)=\frac{S_N-N\psi}{\sqrt{N\psi(1-\psi)}} \qquad (۷)$$

واضح است که ناحیهٔ قبول برای این آماره، همانند آزمون سازگاری معمولی به ψ بستگی ندارد و تابعی از فقط α_{new} می‌باشد. به علاوه از آنجا که آماره دارای توزیع نرمال استاندارد است و آزمون فرض، دوطرفه می‌باشد، ناحیهٔ قبول آماره، متقارن و به صورت $[-k_\alpha, k_\alpha]$ خواهد بود. در آزمون معمولی ψ برابر با $1-\alpha$ بود و چنانچه آماره در ناحیهٔ $[-k_\alpha, k_\alpha]$ قرار می‌گرفت، مولد از آزمون عبورمی‌نمود. حال با مشخص بودن ناحیهٔ قبول آمارهٔ آزمون، ناحیهٔ قبول پارامتر ψ را بدست می‌آوریم.

$$-k_\alpha \le T_N(\psi)=\frac{S_N-N\psi}{\sqrt{N\psi(1-\psi)}} \le k_\alpha \qquad (۸)$$

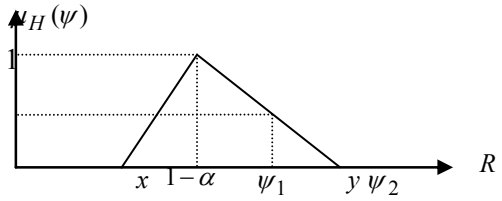
پس از انجام چند عملیات جبری ساده بر روی ناحیهٔ مشخص شده در رابطهٔ (۸)، ناحیهٔ قبول پارامتر ψ به صورت $[\psi_1, \psi_2]$ بدست می‌آید که در آن ψ_1 و ψ_2 از رابطهٔ (۹) بدست می‌آیند.

$$\psi_{1,2}=\frac{2NS_N+Nk_\alpha{}^2 \pm k_\alpha\sqrt{N^2k_\alpha{}^2+4NS_N(N-S_N)}}{2(N^2+Nk_\alpha{}^2)} \qquad (۹)$$

در مرجع [۶] درجهٔ قبول فرض صفر که ما آن را با $\mu_{H_{f,0}}(T_N(\psi))$ نمایش می‌دهیم، با استفاده از عملگر کمکی مثلثی[1] *sup* به صورت رابطهٔ (۱۰) پیشنهاد شده‌است.

$$\mu_{H_{f,0}}(T_N)=\sup_{\psi:T_N(\psi)\in[-k_\alpha,k_\alpha]}\mu_H(\psi) \qquad (۱۰)$$

که در آن $\mu_H(\psi)$ تابع عضویت عدد فازی تقریباً $1-\alpha$ می‌باشد. به بیان ساده‌تر درجهٔ قبول فرضیهٔ فازیِ $H_{f,0}$ (احتمال عبور یک دنبالهٔ تولید شده توسط مولد از آزمون T تقریباً برابر با $1-\alpha$)، برابر با ماکزیمم تابع عضویت عدد فازی $1-\alpha$ در ناحیهٔ قبول پارامتر ψ است. شکل (۲) روند بدست آوردن درجهٔ قبول فرضیهٔ فازیِ صفر را نشان می‌دهد.



شکل۲ : تصمیم گیری فازی درباره درجه قبول فرض فازی $H_{f,0}$

در شکل (۲) با ψ_1 و ψ_2 نشان داده شده، درجهٔ قبول فرض صفر برابر $\mu_H(\psi_1)$ می‌باشد. به طور کلی با توجه به اینکه تابع عضویت عدد فازی $1-\alpha$ قبل از $1-\alpha$ صعودی و بعد از آن نزولی است، اگر ψ_1 و ψ_2 بزرگتر از $1-\alpha$ باشند، درجهٔ قبول فرض صفر برابر $\mu_H(\psi_1)$ خواهد بود و اگر هر دو کوچکتر از $1-\alpha$ باشند، درجهٔ قبول برابر $\mu_H(\psi_2)$ خواهد بود.

۳-۲ نحوهٔ انتخاب عدد فازی

به طور کلی همانطور که α بنا به صلاحدید و بر اساس دقت مورد نیاز آزمونگرانتخاب می‌شود، انتخاب x و y در عدد فازی مثلثی $(x,1-\alpha,y)_T$ جهت انجام آزمون سازگاری فازی نیز تابع قانون مشخصی نیست، اما می‌توان ملاحظاتی را در انتخاب آنها مدنظر قرار داد. یک عامل در این انتخاب، مقدار N یعنی تعداد دفعات انجام آزمون آماری T می‌باشد. به طور کلی انتظار داریم که با افزایش N از میزان ابهام کاسته شود و x و y به $1-\alpha$ نزدیکتر شوند. در حد وقتی که $x=y=1-\alpha$ باشد، عدد

[1] *Triangular co-norm*

فازی به یک عدد معمولی تبدیل شده و آزمون فازی دقیقاً مانند آزمون معمولی عمل خواهد کرد. پیشنهاد ما برای لحاظ این ویژگی در انتخاب عدد فازی این است که فاصله x و y تا $1-\alpha$ متناسب با عکس N انتخاب شوند.

$$(1-\alpha)-x=\frac{k_1}{N} \qquad , \qquad y-(1-\alpha)=\frac{k_2}{N}$$

که در آن k_1 و k_2 اعداد ثابت هستند، اما لزوماً با هم برابر نمی‌باشند. به عنوان مثال فرض کنید آزمون T ویژگی خاصی از دنباله تصادفی ایده‌آل را در مورد یک مولد مورد ارزیابی قرار می‌دهد. همچنین فرض کنید که سطح اهمیت برابر با پنج درصد انتخاب شده باشد. بنابراین انتظار می‌رود برای یک مولد ایده‌آل ۹۵درصد از دنباله‌ها از آزمون عبور کنند. اما چنانچه برای یک مولد ۹۹درصد دنباله‌ها از آزمون عبورکنند این بدان مفهوم است که این مولد دنباله‌های ضعیف کمتری تولید می‌کند و نسبت به مولدی که ۹۱درصد از دنباله‌های آن از آزمون عبور کنند ارجح است، در حالیکه قدرمطلق آماره آزمون سازگاری برای هر دو مولد یکسان می‌باشد، اگرچه که رغبت بیشتر آزمونگر نسبت به مولدی که دنباله‌های ضعیف کمتر تولید می‌کند، تا حدی فرض استقلال بیتها را زیر پا می‌گذارد. مثالی که در ادامه مطرح می‌شود، مؤید همین مطلب است.

مثال ۱ : فرض کنید ۱۰۰۰ دنباله ۱۰۰ بیتی از یک مولد، تحت آزمون فرکانس قرار گرفته باشند. اگر در آزمون فرکانس مورد نظر $\alpha=0.05$ انتخاب شده باشد، واضح است که $N\alpha\geq 5$ می‌باشد و شرط استفاده از آزمون سازگاری برقرار است. اگر از بین این ۱۰۰۰ دنباله، ۷۵ دنباله از آزمون فرکانس عبور نکنند و $\alpha_{new}=0.01$ انتخاب شود، یک بار به صورت متداول و یک بار طبق روش بیان شده و با عدد فازی $(0.925,0.95,1)_T$ درباره کیفیت مولد اظهار نظر می‌کنیم.

در روش متداول، آماره آزمون سازگاری طبق رابطهٔ (۴) برابر با $T_N=-3.627$ بدست می‌آید که از مقدار آستانه $-k_\alpha=-1.96$ کوچکتر است. در نتیجه این مولد از آزمون سازگاری معمولی عبور نمی‌کند. اما در روش فازی بیان شده با جایگذاری $S_N=925$ در رابطهٔ (۹) بازهٔ قبول پارامتر ψ به صورت $[\psi_1,\psi_2]=[0.9070,0.9397]$ بدست می‌آید. بنابراین با توجه به اینکه ψ_1 و ψ_2 هر دو کوچکتر از $1-\alpha$ هستند، داریم:

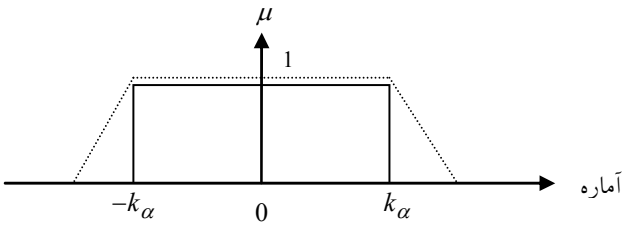
$$\mu_{H_{f,0}}(T_N)=\sup_{\psi:T_N(\psi)\in[-k_\alpha,k_\alpha]}\mu_H(\psi)\ =\mu_H(\psi_2)=0.588$$

از این رو درجهٔ تصادفی بودن مولد مورد نظر برابر با 0.588 بدست می‌آید. حال اگر تحت همان شرایط فقط ۲۵ دنباله از آزمون عبورنکنند، این بار آمارهٔ آزمون برابر با $T_N=3.627$ خواهد بود که از $k_\alpha=1.96$ بزرگتر است و از آزمون متداول عبور نمی‌کند. این بار بازه قبول پارامتر ψ به صورت $[\psi_1,\psi_2]=[0.9634,0.9830]$ بدست می‌آید. بنابراین با توجه به اینکه ψ_1 و ψ_2 هر دو بزرگتر از $1-\alpha$ هستند، داریم:

$$\mu_{H_{f,0}}(T_N)=\sup_{\psi:T_N(\psi)\in[-k_\alpha,k_\alpha]}\mu_H(\psi)\ =\mu_H(\psi_1)=0.732$$

همانطور که ملاحظه شد، دنباله‌ای که در آزمون معمولی رد شده بود از درجه‌ای از قبول در آزمون فازی برخوردار است. در حالت کلی فرض کنید در آزمون سازگاری معمولی، ناحیهٔ قبول آماره به صورت $[-k_\alpha,k_\alpha]$ و ناحیهٔ رد، خارج از این بازه باشد. در این صورت با انجام عملیات سادهٔ جبری می‌توان نشان داد که اگر S_N به گونه‌ای باشد که T_N در ناحیه $[-k_\alpha,k_\alpha]$ قرار گیرد، آنگاه در آزمون فازی، ψ_1 کوچکتر یا مساوی $1-\alpha$ و ψ_2 بزرگتر یا مساوی $1-\alpha$ خواهد بود و این به مفهوم یک بودن درجه قبول در این ناحیه است. به عبارت ساده‌تر در آزمون فازی ناحیهٔ قبول دست نخورده باقی می‌ماند و به ناحیهٔ رد نیز درجه‌ای از قبول نسبت داده می‌شود. شکل (۳) درجهٔ قبول فرض صفر را به صورت یک تابع عضویت از آمارهٔ آزمون نشان می‌دهد. در آزمون معمولی درجهٔ قبول را در ناحیهٔ $[-k_\alpha,k_\alpha]$ با یک و در خارج از این ناحیه با صفر نشان داده‌ایم (خط پر).

همانطور که مشاهده می‌شود در آزمون فازی (نقطه چین) قسمتی از ناحیهٔ رد در آزمون معمولی با همان سطح اهمیت، از درجهٔ قبول غیر صفر برخوردار است. به عبارت دیگر برای مقایسهٔ دو آزمون سازگاری معمولی و فازی با α_{new}های برابر می‌توان گفت که ناحیهٔ قبول آزمون فازی، قدری وسیعتر از ناحیهٔ قبول آزمون معمولی است.



شکل ۳ : درجهٔ قبول فرض صفر در آزمون معمولی و فازی

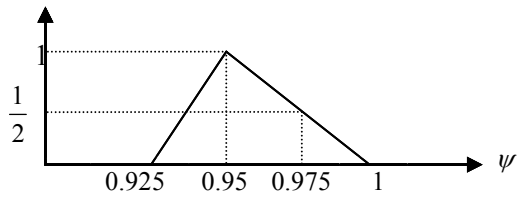
تا این جا آزمونگر با دو مسأله روبروست. یکی انتخاب α_{new} و دیگری تصمیم‌گیری نهایی در مورد استفاده یا عدم استفاده از مولد با توجه به درجهٔ قبول بدست آمده. از طرفی با کاهش α_{new} واضح است که k_α نیز کم می‌شود، به طوری که اگر α_{new} را به سمت صفر میل دهیم، k_α نیز به سمت صفر میل می‌کند. البته با صفر شدن α_{new} مفهوم آزمون آماری متداول را با اشکال مواجه می‌سازد، زیرا در این حالت ناحیهٔ قبول آزمون معمولی برابر با یک نقطه می‌شود و مولد، تنها هنگامی از آزمون عبور می‌کند که آماره برابر با صفر باشد. اما در آزمون فازی چنین اشکالی وجود نخواهد داشت و میتوان α_{new} را به اندازهٔ دلخواه به صفر نزدیک نمود، زیرا در آزمون فازی با صفر شدن α_{new}، تابع عضویت قبول آماره که قبلاً به صورت ذوزنقه بود به مثلث تبدیل می‌شود. بنابراین پیشنهاد می‌کنیم که در آزمون فازی ارائه شده، α_{new} برابر با صفر انتخاب شود تا هم مشکل انتخاب α_{new} برای آزمونگر و هم مسألهٔ بزرگ شدن ناحیهٔ قبول در آزمون فازی حل شود. با میل نمودن k_α به سمت صفر از رابطهٔ (۹) داریم:

$$\psi_1 = \psi_2 = \frac{S_N}{N} \qquad (۱۱)$$

بنابراین رابطهٔ (۱۰) به صورت زیر ساده می‌شود:

$$\mu_{H_{f,0}}(T_N) = \mu_H\left(\frac{S_N}{N}\right) \qquad (۱۲)$$

مثال ۲ : یک بار دیگر به مثال ۱ برمی‌گردیم. این بار اگر α_{new} در حد، برابر صفر باشد و ۷۵ دنباله از آزمون T عبور نکنند، آنگاه $\psi_1 = \psi_2 = 0.925$ و درجهٔ قبول فرض صفر، مطابق شکل (۴) برابر با $\mu_H(0.925) = 0$ خواهد شد. به طور مشابه اگر ۲۵ دنباله از آزمون عبور نکنند، درجهٔ قبول برابر با $\mu_H(0.975) = \frac{1}{2}$ خواهد شد.



شکل ۴ : تصمیم‌گیری فازی با شرط $\alpha_{new} = 0$

باز هم تأکید می‌شود که صفر شدن α_{new} در آزمون آماریِ معمولی، بدون اشکال نیست و موجب تبدیل ناحیهٔ قبول به صرفاً یک نقطه می‌گردد. اما در آزمون فازی، با میل دادن α_{new} به سوی صفر، اشکالی پیش نمی‌آید زیرا این بار با مثلثی شدنِ ناحیهٔ قبول تکیه‌گاهِ آن به صورت یک بازه است نه یک نقطه.

۴- نتیجه‌گیری

در این مقاله، نخست مشکل اصلی آزمونهای آماری متداول که جهت ارزیابی مولدهای شبه تصادفی بکار می‌روند، بیان گردید. این مشکل این است که نتیجهٔ این گونه آزمونها، باینری است و تنها به صورت رد یا قبول مولد اعلام می‌گردد. سپس روش جدیدی مبتنی بر روشهای فازی برای ارزیابی مولدها پیشنهاد دادیم. اساس روش پیشنهادی بر فازی کردن فرضیهٔ شبه تصادفی بودن در آزمون سازگاری متداول استوار است. نتیجهٔ روش جدید نیز به صورت فازی بدست می‌آید و به هر مولد، درجه‌ای از قبول بین صفر تا یک نسبت می‌دهد. سپس در ادامه برای بهبود روش پیشنهادی، با میل دادن سطح اهمیت (خطای نوع اول در آزمون فرض غیر فازی) به سوی صفر، مشکل انتخاب سطح اهمیت که خود، یکی از نکات ابهام آمیز در آزمون معمولی است، مرتفع گردید و تصمیم گیری نیز بسیار ساده‌تر شد. روش ارائه شده، نخستین گام جهت ارزیابی آماری- فازی دنباله‌ها و مولدهای شبه‌تصادفی است. اگرچه این روش در ساده‌ترین حالت، یعنی با استفاده از توابع عضویت مثلثی بیان شده است، دو نتیجه مهم را در بر دارد. نخست اینکه با کنار گذاشتن منطق دو ارزشی و استفاده از روش فازی می‌توان اطلاعات بیشتری دربارهٔ مولدها بیان نمود. دوم اینکه همانطور که نشان داده شد دیگر نیازی به انتخاب سطح اهمیت توسط آزمونگر نیست. به نظر می‌رسد در آینده با مدلسازی دقیقترِ ابهام در آزمونهای آماری و استفاده از توابع عضویت دیگر که نمایندهٔ دقیقتر این ابهام باشند، بتوان روشهای مناسبتر وکاملتری برای ارزیابی مولدها و دنباله‌های شبه‌تصادفی ارائه نمود.

مراجع

[1] A.L. Rukhin, Testing randomness: a suite of statistical procedures. SIAM J. Theory Probability Appl. (45) 2000.

[2] D.E. Knuth, The Art of Computer Programming, 2, third ed., Prentice Hall PTR, 1997.

[3] A.L. Rukhin, et al,"A Statistical Test Suite for Random and Pseudo Random Number Generators for Cryptographic Applications ", NIST Special Publication 800-22, www.nist.gov, 2001.

[٤] محمد دخیل علیان، ارزیابی دنباله‌های شبه تصادفی و طراحی مولدهای آشوبی، رساله دکتری، دانشکده برق و کامپیوتر، دانشگاه صنعتی اصفهان،۱۳۷۷.

[5] B.F. Arnold. An approach to fuzzy hypothesis testing. Metrika, 44:119–126, 1996.

[6] B.F. Arnold. Testing fuzzy hypothesis with crisp data. Fuzzy Sets Syst., 94(2):323–333,1998.

[7] N. Watanabe and T. Imaizumi. A fuzzy statistical test of fuzzy hypotheses. Fuzzy SetsSyst., 53:167–178, 1993.

[8] H. Bandemer and S. Gottwald. Fuzzy Sets Fuzzy Logic Fuzzy Methods with Applications. John-Wiley, 1996.