

# توسعه‌ی الگوریتم خوشه‌بندی برای برآورد نقاط تغییر چندگانه در پایش پروفایل خطی ساده

مهران مهتابی\* (دانشجوی کارشناسی ارشد)

یاسر صمیمی (استادیار)

دانشکده‌ی مهندسی صنایع، دانشگاه صنعتی خواجه نصیرالدین طوسی

مهندسی صنایع و مدیریت شریف، تابستان ۱۳۹۵  
دوری ۱ - ۳۲، شماره‌ی ۱/۲، ص. ۴۷-۵۵

شناسایی تغییر در پارامترهای فرایند از مسائل حائز اهمیت در کنترل فرایند آماری به شمار می‌رود؛ چرا که ارائه‌ی اطلاعات دقیق درخصوص زمان و الگوی تغییر در پارامترهای فرایند، اقدامات اصلاحی مؤثرتر را به دنبال خواهد داشت. به‌طور کلی، در مطالعات انجام شده در زمینه‌ی برآورد نقطه‌ی تغییر، عمدتاً موضوع تغییرات انفرادی پارامترها مورد بررسی قرار گرفته است. این در حالی است که یک هشدار در نمودار کنترل ممکن است متأثر از چندین تغییر باشد؛ تغییراتی که عدم کشف آنها ممکن است باعث تأخیر یا حتی جهت‌گیری اشتباه در روند انجام اقدامات اصلاحی باشد. از سوی دیگر، در کاربردهای کنترل فرایند آماری عملکرد یک فرایند یا محصول معمولاً به‌صورت یک مشخصه‌ی کیفی انفرادی یا برداری از مشخصه‌های کیفی که در طول زمان جمع‌آوری شده مورد ارزیابی قرار می‌گیرد. با این وجود، به‌ویژه طی سال‌های اخیر تحقیقات متعدد به بررسی شرایطی می‌پردازند که عملکرد یک فرایند از طریق تحلیل رابطه‌ی تابعی بین یک متغیر وابسته و یک یا چند متغیر مستقل به شکل مناسب‌تری قابل توصیف است. چنین تابعی تحت عنوان پروفایل شناخته می‌شود. مروری بر تحقیقات انجام شده در زمینه‌ی شناسایی نقطه‌ی تغییر در پایش پروفایل‌ها نشان می‌دهد تا این زمان موضوع تغییرات چندگانه در این زمینه مورد بررسی قرار نگرفته است. در این مقاله، روشی برای کشف تعداد و برآورد زمان وقوع تغییرات چندگانه در زمینه‌ی پایش پروفایل خطی ساده با استفاده از یک الگوریتم خوشه‌بندی ارائه می‌شود. نتایج مطالعه‌ی شبیه‌سازی درخصوص دو و سه تغییر پله‌یی در پارامترهای مدل نشان می‌دهد روش ارائه شده علاوه بر امکان پایش فرایند در سطح معناداری معین، با دقت مناسبی قادر به شناسایی نقاط تغییر چندگانه است.

واژگان کلیدی: کنترل فرایند آماری، نقاط تغییر چندگانه، الگوریتم خوشه‌بندی، پروفایل خطی ساده.

mehran\_mahtabi@ymail.com  
y\_samimi@kntu.ac.ir

## ۱. مقدمه

شناسایی تغییر در پارامتر یک فرایند از مسائل حائز اهمیت در زمینه‌ی کنترل فرایند آماری<sup>۱</sup> محسوب می‌شود؛ چرا که هرچه اطلاعات دقیق‌تری درخصوص زمان و الگوی تغییر در پارامترهای فرایند فراهم شود، اقدامات اصلاحی نیز به‌شکل مؤثرتر صورت گرفته و با کشف علل ریشه‌یی و رفع آنها، از هزینه‌ی تولید محصول معیوب یا ارائه‌ی خدمت معیوب به میزان چشم‌گیری جلوگیری خواهد شد. سیگنال خارج از کنترل در نمودارهای کنترل مرسوم نشان‌دهنده‌ی محل واقعی وقوع انحراف با دلیل نیست. این موضوع از ضرورت به‌کارگیری روش مناسب برای برآورد محل واقعی بروز تغییر حکایت دارد. به‌علاوه، تعدد تغییر در پارامترهای فرایند موضوعی

است که در تحقیقات مربوط به برآورد نقطه‌ی تغییر کم‌تر مورد بررسی قرار گرفته است. در واقع، علامت هشدار در نمودار کنترل می‌تواند متأثر از چندین تغییر باشد که در مقاطع مختلف زمانی رخ داده و عدم شناسایی صحیح تعداد و محل آنها می‌تواند موجب جهت‌گیری نادرست در ریشه‌یابی وضعیت خارج از کنترل و متعاقباً روند انجام اقدامات اصلاحی باشد. انواع تغییر در پارامترهای فرایند را می‌توان در چهار دسته کلی مونوتونیک<sup>۲</sup>، تغییرات پله‌یی<sup>۳</sup>، روند خطی<sup>۴</sup> و نامنظم طبقه‌بندی کرد.<sup>[۱]</sup>

در کاربرد کنترل فرایند آماری، عملکرد یک فرایند یا محصول، معمولاً به‌صورت یک مشخصه‌ی کیفی انفرادی یا برداری از مشخصه‌های کیفی که در طول زمان جمع‌آوری شده‌اند مورد ارزیابی قرار می‌گیرد. اما در برخی موارد عملکرد یک فرایند از طریق رابطه‌ی تابعی بین یک متغیر وابسته با یک یا چند متغیر مستقل به‌نحو

\* نویسنده مسئول

تاریخ: دریافت ۱۳۹۳/۱/۲۷، اصلاحیه ۱۳۹۳/۵/۱۹، پذیرش ۱۳۹۳/۶/۱۵.

مناسب‌تر قابل توصیف است. پایش این نوع از مشخصه‌های کیفی که تحت عنوان پروفایل شناخته می‌شود به‌ویژه طی دهه‌ی اخیر توجه محققین را به خود معطوف داشته است. پروفایل‌های خطی، چندجمله‌یی، لجستیک و ناپارامتریک از مهم‌ترین انواع توابعی محسوب می‌شوند که در مطالعات محققین مورد بررسی قرار گرفته است.<sup>[۴]</sup>

بررسی‌ها نشان می‌دهد که تحقیقات مرتبط با کشف نقطه‌ی تغییر در پایش پروفایل مبتنی بر شناسایی تغییرات انفرادی است. براین اساس، در این مقاله روشی برای برآورد نقاط تغییر چندگانه<sup>۵</sup> در زمان پایش مدل پروفایل خطی ساده بیان خواهد شد. این روش مبتنی بر توسعه‌ی یک روش خوشه‌بندی است که سابقاً برای شناسایی نقاط تغییر چندگانه در نمودارهای کنترل مرسوم طراحی شده است. پیش از بیان مسئله و شرح روش حل، مروری مختصر بر تحقیقات مرتبط با موضوع مقاله ارائه می‌شود.

پایش پروفایل‌های خطی زمینه‌ی نسبتاً جدید از کاربردهای کنترل کیفیت آماری است و بیشتر فعالیت‌های انجام‌شده در این حوزه مربوط به سال‌های اخیر است. در مرور روش‌های پایش پروفایل خطی<sup>[۳]</sup> یک چارچوب کلی برای پایش فرایندها با استفاده از مشخصه‌هایی از نوع پروفایل ارائه شده است. در این زمینه روشی برای پایش پروفایل‌های خطی در فاز II پیشنهاد شد<sup>[۴]</sup> که در آن از دو نمودار کنترل شامل یک نمودار کنترل دومتغیره ( $T^2$ ) برای پایش ضرایب رگرسیون، و یک نمودار EWMA/R برای پایش باقی‌مانده‌ها استفاده می‌شود. محققین براساس کدگذاری متغیر پیش‌بینی، روشی بر مبنای استفاده از سه نمودار کنترل تک‌متغیره مجزا پیشنهاد کردند؛<sup>[۵]</sup> آنها از نمودارهای میانگین متحرک موزون نمایی<sup>۶</sup> برای پایش شیب، عرض از مبدأ و پراکندگی خطا استفاده کردند. همچنین با پیشنهاد نمودارهای کنترلی<sup>[۶]</sup> شیب ارتباط بین دو فرایند اندازه‌گیری را به‌منظور اطمینان از دقت اندازه‌گیری پایش کردند. از سوی دیگر، با استفاده از آزمون نسبت درست‌نمایی و یک نمودار EWMA روشی برای پایش پروفایل‌های خطی ساده در فاز II ارائه شد<sup>[۷]</sup> که طی آن موضوع فواصل نمونه‌گیری متغیر نیز بررسی شد. برای پروفایل‌های خطی ساده در فاز II نیز روشی براساس نمودار کنترل جمع تجمعی<sup>۷</sup> ارائه شده است.<sup>[۸]</sup>

درخصوص شناسایی نقطه‌ی تغییر در نمودارهای کنترل تک‌متغیره و چندمتغیره کاربرد روش‌های خوشه‌بندی<sup>۸</sup> مورد توجه محققین قرار گرفته است. روشی جدید برای شناسایی نقاط تغییر چندگانه در فاز I برای نمودار کنترل مشاهدات انفرادی ارائه شده<sup>[۹]</sup> که در آن الگوریتم خوشه‌بندی قادر به شناسایی تعداد تغییرات و مکان متناظر آنهاست. همچنین با استفاده از خوشه‌بندی فازی، روش تخمین نقطه‌ی تغییر در انواع مختلف نمودارهای کنترل با اندازه نمونه‌های ثابت و متغیر<sup>[۱۰]</sup> ارائه شده است. محققین استفاده از روش خوشه‌بندی را برای تخمین نقطه‌ی تغییر پس از دریافت هشدار از نمودار کنترل  $\bar{X}$  پیشنهاد کردند.<sup>[۱۱]</sup> آنان همچنین برآورد نقاط تغییر چندگانه در بردار میانگین و ماتریس کوواریانس فرایند نرمال چندمتغیره را ارائه کردند.<sup>[۱۲]</sup> مروری جامع بر تحقیقات مرتبط با موضوع شناسایی نقطه‌ی تغییر<sup>[۱]</sup> نیز ارائه شده است.

در تحقیقات متعددی به مسئله‌ی شناسایی نقطه‌ی تغییر در پایش پروفایل‌ها پرداخته شده است. بدین‌منظور روشی مبتنی بر نقطه‌ی تغییر برای پایش پروفایل‌های خطی ساده در فاز II پیشنهاد شد.<sup>[۱۳]</sup> این روش با استفاده از آماره‌ی نسبت درست‌نمایی تعمیم یافته، و شیفت در عرض از مبدأ، شیب و انحراف معیار را با استفاده از یک نمودار کنترل کشف می‌کند. محققین روشی برای شناسایی نقطه‌ی تغییر با استفاده از مدل رگرسیون بخش‌بندی شده<sup>[۱۴]</sup> ارائه کردند. برای این منظور

روشی مبتنی بر آزمون درست‌نمایی در یک مدل رگرسیون خطی ساده بخش‌بندی شده پیشنهاد شده است. از این روش می‌توان برای تشخیص ثبات و شناسایی نقاط تغییر در مجموعه داده پروفایل خطی ساده در فاز I استفاده کرد. پژوهش‌گران در زمینه‌ی پایش پروفایل خطی نیز روشی مبتنی بر تخمین بیشترین درست‌نمایی از نقطه‌ی تغییر برای شناسایی مکان شیفت<sup>[۱۵]</sup> ارائه کردند که در آن از نمودار کنترل میانگین متحرک موزون نمایی چندمتغیره<sup>۹</sup> برای پایش پروفایل خطی استفاده شده است. آنان رویکرد نقطه‌ی تغییر را براساس مدل رگرسیون ناپارامتریک مطرح کردند<sup>[۱۶]</sup> و از آزمون نسبت درست‌نمایی<sup>۱۰</sup> تعمیم یافته استفاده کردند. همچنین با استفاده از تخمین زنده‌ی بیشترین درست‌نمایی یک مدل نقطه‌ی تغییر برای تخمین مکان شیفت در پروفایل‌های خطی چندگانه‌ی چندمتغیره ارائه کردند.<sup>[۱۷]</sup> با استفاده از تخمین بیشترین درست‌نمایی، روشی برای شناسایی زمان تغییرات پله‌یی در پروفایل‌های پواسون<sup>[۱۸]</sup> ارائه شد. کشف زمان تغییر در پروفایل باینری برای تغییرات پله‌یی<sup>[۱۹]</sup> و روند خطی<sup>[۲۰]</sup> صورت پذیرفته است. اخیراً نیز موضوع شناسایی زمان تغییر برای تغییر از نوع روند خطی<sup>[۲۱]</sup> در پروفایل خطی چندمتغیره بررسی شده است.

مرور تحقیقات مرتبط با موضوع نشان می‌دهد مطالعه‌ی برای شناسایی نقطه‌ی تغییر چندگانه در پایش پروفایل‌های خطی صورت نگرفته است. در نوشتار حاضر براساس توسعه‌ی الگوریتم خوشه‌بندی<sup>[۹]</sup> روشی برای شناسایی نقاط تغییر چندگانه در پایش پروفایل خطی ساده ارائه می‌شود.

در بخش بعدی، ضمن بیان مدل رگرسیون خطی به بیان صورت مسئله می‌پردازیم. سپس در بخش سوم نحوه‌ی توسعه‌ی الگوریتم خوشه‌بندی تشریح، و چگونگی تعیین حد تصمیم ارائه می‌شود. در بخش چهارم نیز یک مثال عددی برای درک نحوه‌ی پیاده‌سازی الگوریتم ارائه می‌شود. سپس در بخش پنجم نتایج شبیه‌سازی روش ارائه‌شده بررسی، و در نهایت نتیجه‌گیری در بخش ششم ارائه می‌شود.

## ۲. بیان مسئله

در نوشتار حاضر، با در نظر گرفتن حالتی که در آن ویژگی داده‌ها با استفاده از یک معادله‌ی رگرسیون خطی ساده مشخص می‌شود، روش خوشه‌بندی سالیوان<sup>[۹]</sup> برای شناسایی نقاط تغییر چندگانه در پایش پروفایل خطی ساده توسعه می‌یابد. لازم به ذکر است حوزه‌ی کاری مقاله در فاز ۱ است و شیفت در انحراف معیار نیز در نظر گرفته نشده است. مدل پروفایل خطی ساده با استفاده از رابطه‌ی ۱ نشان داده می‌شود:

$$y_{ij} = a_j + b_j x_{ij} + \varepsilon_{ij}, \quad i = 1, 2, \dots, n_j, \quad j = 1, 2, \dots, m \quad (1)$$

در معادله‌ی رگرسیونی،  $(x_{ij}, y_{ij})$  بیان‌گر مشاهده‌ی  $i$ ام از نمونه‌ی  $j$ ام است. همچنین فرض می‌شود  $\varepsilon_{ij}$ ها متغیرهای تصادفی مستقلی با توزیع نرمال و میانگین صفر و واریانس یکسان  $\sigma^2$  هستند. همانند بیشتر مطالعات انجام شده در زمینه‌ی پروفایل‌های خطی، مقادیر  $X$  ثابت، و در هر نمونه دارای مقادیر یکسان فرض می‌شود. در مدل‌سازی رگرسیون خطی، حالتی بررسی می‌شود که در آن مقادیر  $X$  به شکل تصادفی در نظر گرفته می‌شود.<sup>[۲۲]</sup>

هدف اصلی از تحلیل پروفایل‌ها در فاز I، شناسایی پروفایل‌های دارای شرایط خارج از کنترل و دست‌یابی به برآوردی معتبر از پروفایل مرجع است. در مدل رگرسیونی مورد استفاده، بعد مسئله به سه پارامتر شیب ( $b$ )، عرض از مبدأ ( $a$ ) و واریانس

خطا ( $\sigma^2$ ) کاهش می‌یابد. در نتیجه، برای مجموعه داده‌های پروفایل با حداقل دو مشاهده در هر نمونه ( $n_j \geq 2$ ) باید این پارامترها برآورد شوند. تمرکز روش ارائه شده در این قسمت، بر شناسایی تغییرات در پارامترهای  $a$  و  $b$  از نمونه‌ی به نمونه‌ی دیگر است و فرض می‌شود که تغییری در داخل یک نمونه رخ نمی‌دهد. از آنجا که در بررسی‌های فاز I اطلاعاتی از فرایند در دسترس نیست، با فرض تحت کنترل بودن فرایند باید پارامترهای مدل (پارامترهای رگرسیون) را تخمین زد. چنان که در بخش قبل بیان شد، مقادیر تحت کنترل پارامترهای  $a$ ،  $b$  و  $\sigma^2$  در معادله‌ی فوق ناشناخته‌اند. اگر برای همه نمونه‌ها ( $j = 1, 2, \dots, m$ )، روابط  $a_j = a$  و  $b_j = b$ ،  $\sigma_j^2 = \sigma^2$  اما هدف اصلی در فاز I ارزیابی ثبات فرایند و تخمین پارامترهای مدل در شرایط تحت کنترل است؛ لذا در بخش بعد روش تخمین پارامترها ارائه خواهد شد.

### ۳. برآورد نقاط تغییر چندگانه

در بخش قبل، مدل پروفایل مورد استفاده و همچنین صورت مسئله‌ی نقاط تغییر چندگانه معرفی شد. در ادامه، الگوریتم پیشنهادی برای برآورد نقاط تغییر در پایش پروفایل خطی ساده طرح خواهد شد.

#### ۱.۳. بیان الگوریتم

در این بخش مراحل توسعه‌ی الگوریتم خوشه‌بندی سالیوان<sup>[۹]</sup> و متدولوژی حل آن برای پایش پروفایل خطی ساده مطرح می‌شود. با در نظر گرفتن  $m$  پروفایل در فاز I، الگوریتم خوشه‌بندی با  $m - 1$  حد آغاز می‌شود، یعنی  $m$  پروفایل ابتدا به صورت جداگانه در خوشه‌های تکی قرار می‌گیرند. در نظر بگیرید بردار  $(a_j, b_j)$  معرف پارامترهای پروفایل  $j$ ام است که در یک خوشه‌ی تکی واقع شده است. سپس با استفاده از رابطه‌ی  $d_k$  میزان عدم تشابه در خوشه‌های مجاور بررسی می‌شود. برای این منظور از آماره‌ی ماهالاتوئیس برای محاسبه‌ی فاصله‌ی مراکز هر دو خوشه‌ی مجاور استفاده می‌شود. در اینجا، فاصله‌ی ماهالاتوئیس که برای محاسبه‌ی فواصل در فضای چندمتغیره مورد استفاده قرار می‌گیرد توسط رابطه‌ی ۶ ارائه شده است:<sup>[۲۳]</sup>

$$d_k = (\bar{Z}_k - \bar{Z}_{k+1})' \left[ \left( \frac{1}{n_k} + \frac{1}{n_{k+1}} \right) \times S_{pool} \right]^{-1} (\bar{Z}_k - \bar{Z}_{k+1}) \quad (6)$$

که در آن  $\bar{Z}_k$  میان‌گر متوسط بردار ضرایب پروفایل‌هایی است که در خوشه‌ی  $k$ ام قرار می‌گیرد.  $n_k$  بیان‌گر تعداد بردارها در خوشه‌ی  $k$ ام و  $S_{pool}$  ماتریس واریانس - کوواریانس برآورد ضرایب است که از رابطه‌ی ۷ محاسبه می‌شود:

$$S_{pool} = \begin{bmatrix} S_{\cdot}^t & S_{\cdot 1}^t \\ S_{\cdot 1}^t & S_1^t \end{bmatrix} \quad (7)$$

به طوری که در آن  $S_{\cdot}^t = \sigma^2 \times S_{xx}^{-1}$  و  $S_{\cdot 1}^t = \sigma^2 \times (n^{-1} + \bar{X}^t S_{xx}^{-1})$  و همچنین  $S_{\cdot 1}^t = -\sigma^2 \times \bar{X}^t S_{xx}^{-1}$  شکل ماتریسی رابطه‌ی ۶ عبارت است از:

$$d_k = \begin{bmatrix} a_k - a_{k+1} \\ b_k - b_{k+1} \end{bmatrix}^T * \left[ \left( \frac{1}{n_k} + \frac{1}{n_{k+1}} \right) \times \begin{bmatrix} S_{\cdot}^t & S_{\cdot 1}^t \\ S_{\cdot 1}^t & S_1^t \end{bmatrix} \right]^{-1} * \begin{bmatrix} a_k - a_{k+1} \\ b_k - b_{k+1} \end{bmatrix} \quad (8)$$

خطا ( $\sigma^2$ ) کاهش می‌یابد. در نتیجه، برای مجموعه داده‌های پروفایل با حداقل دو مشاهده در هر نمونه ( $n_j \geq 2$ ) باید این پارامترها برآورد شوند. تمرکز روش ارائه شده در این قسمت، بر شناسایی تغییرات در پارامترهای  $a$  و  $b$  از نمونه‌ی به نمونه‌ی دیگر است و فرض می‌شود که تغییری در داخل یک نمونه رخ نمی‌دهد. از آنجا که در بررسی‌های فاز I اطلاعاتی از فرایند در دسترس نیست، با فرض تحت کنترل بودن فرایند باید پارامترهای مدل (پارامترهای رگرسیون) را تخمین زد.

چنان که در بخش قبل بیان شد، مقادیر تحت کنترل پارامترهای  $a$ ،  $b$  و  $\sigma^2$  در معادله‌ی فوق ناشناخته‌اند. اگر برای همه نمونه‌ها ( $j = 1, 2, \dots, m$ )، روابط  $a_j = a$  و  $b_j = b$ ،  $\sigma_j^2 = \sigma^2$  اما هدف اصلی در فاز I ارزیابی ثبات فرایند و تخمین پارامترهای مدل در شرایط تحت کنترل است؛ لذا در بخش بعد روش تخمین پارامترها ارائه خواهد شد.

#### ۱.۲. تخمین پارامترهای مدل

رایج‌ترین روش برای تخمین پارامترهای مدل رگرسیون خطی ساده روش حداقل مربعات خطاست. این روش تخمینی از پارامترها ارائه می‌دهد که اختلاف بین پروفایل‌های مشاهده شده و پروفایل‌های متناسب را کمینه می‌کند. تخمین حداقل مربعات برای پارامترهای مدل رگرسیونی خطی ساده شامل شیب و عرض از مبدأ با استفاده از روابط ۲ و ۳ انجام می‌شود:<sup>[۱۴]</sup>

$$b_j = S_{xy(j)} / S_{xx(j)} \quad (2)$$

$$a_j = \bar{Y}_j - b_j \bar{X}_j \quad (3)$$

چنان که  $\bar{Y}_j = \sum_{i=1}^{n_j} \frac{y_{ij}}{n_j}$  میانگین متغیر پاسخ در نمونه‌ی  $j$ ام،  $\bar{X}_j = \sum_{i=1}^{n_j} \frac{x_{ij}}{n_j}$  میانگین متغیر مستقل در نمونه‌ی  $j$ ام است. همچنین داریم:

$$S_{xy(j)} = \sum_{i=1}^{n_j} (X_{ij} - \bar{X}_j) Y_{ij}, \quad S_{xx(j)} = \sum_{i=1}^{n_j} (X_{ij} - \bar{X}_j)^2$$

$\sigma_j^2$  به وسیله‌ی میانگین مربعات خطا تخمین زده می‌شود:

$$MSE_j = SSE_j / (n_j - 2) \quad (4)$$

به طوری که  $SSE_j = \sum_{i=1}^{n_j} e_{ij}^2$  برای  $j = 1, 2, \dots, m$  و  $e_{ij} = Y_{ij} - a_j - b_j X_{ij}$  به ازای  $i = 1, 2, \dots, n_j$  باقی‌مانده پیش‌بینی محسوب می‌شود. لازم به ذکر است در روش مورد بحث، واریانس خطا تحت کنترل در نظر گرفته شده و فرض می‌شود انحرافات بادلایل تنها بر بردار ضرایب پروفایل اثر می‌گذارند.

تخمین نقطه‌ی تغییر رویکردی برگرفته از نظریه‌ی کنترل فرایند آماری است که برای شناسایی زمان دقیق تغییر در پارامتر فرایند استفاده می‌شود. از آنجا که زمان هشدار در نمودار کنترل لزوماً بیان‌گر زمان دقیق تغییر در فرایند نیست، برآورد نقطه‌ی تغییر مبتنی است بر شناسایی نقطه‌ی از زمان که در آن پارامترهای فرایند تحت انحرافات بادلایل تغییر می‌کند. مدل رگرسیونی نقطه‌ی تغییر چنین بیان می‌شود:

$$Y_{ij} = \begin{cases} A_0 + B_0 X_{ij} + \varepsilon_{ij}, & i \leq \tau \\ A_1 + B_1 X_{ij} + \varepsilon_{ij}, & i \geq \tau + 1 \end{cases} \quad (5)$$

به طوری که  $\varepsilon_{ij}$  ها متغیرهای تصادفی مستقل با توزیع نرمال و میانگین صفر و واریانس ثابت‌اند. در مسئله‌ی آزمون فرض، فرض صفر بیان می‌دارد که برای تمامی مقادیر

در آغاز مرحله‌ی  $k$  که  $k = 1, 2, \dots, m-1$  تعداد  $m-k$  حد وجود دارد. ابتدا با استفاده از رابطه‌ی ۸ فواصل مراکز خوشه‌های مجاور محاسبه، و سپس کوچک‌ترین فاصله‌ی به دست آمده حذف می‌شود. با این فرض که  $k^* = \arg \min [d_k]$  موقعیت مربوط به کوچک‌ترین فاصله را نشان می‌دهد، فواصل جدید با  $m-k-1$  حد، به روز می‌شوند. مکان حدی که حذف شده یعنی  $l_{m-k}^* = l_k^*$  و فاصله‌ی آن یعنی  $d_{m-k}^* = d_k^*$  ذخیره می‌شود. فرایند تا جایی ادامه می‌یابد که تمامی حدود حذف شده و داده‌ها در یک دسته قرار می‌گیرند. اساس این روش به‌گونه‌ی است که اگر فرایند تحت کنترل باشد، توالی  $d_k^*$  ها به تدریج کاهش می‌یابد اما اگر تعداد  $R$  شیفت در فرایند وجود داشته باشد مقدار  $d_R^* - d_{R+1}^*$  مقدار بزرگی است و فواصل به دست آمده بعد از  $d_{R+1}^*$  به آرامی کاهش می‌یابد.

### ۲.۳. برآورد حد تصمیم

برای تخمین حد تصمیم از آنجا که در فاز I کنترل هستیم، روش متداولی وجود ندارد. سالیوان<sup>[۹]</sup> نشان داد که فواصل  $d_k^*$  در صورت وجود انحرافات با دلیل افزایش می‌یابد و به این نتیجه رسید که دنباله‌ی  $d_k^*$  ها به صورت یک بردار با توزیع‌های وابسته نشان داده می‌شود. او با در نظر گرفتن مقادیر مورد انتظار در شرایط تحت کنترل از اولین مقادیر  $d_k^*$  ها برای تخمین حد تصمیم استفاده کرد. به طور کلی، آماره‌ی  $\max[d_1^*, d_2^*]$  برای تعیین قاعده‌ی تصمیم‌گیری عملکرد مناسبی نشان داده است. با آگاهی از یک الگوی خارج از کنترل خاص، مانند چند شیفت یا چند دورافتادگی، آماره‌ی مذکور حساسیت مناسبی در تحلیل فرایند از خود نشان می‌دهد. مشخصاً، استفاده‌ی تکی از فاصله‌ی اول به‌تنهایی حساسیت چندانی در شناسایی شیفت یا دورافتادگی ندارد. از سوی دیگر در نظر گرفتن تعداد زیاد  $d_k^*$  ها مسئله‌ی مقایسات چندگانه را مطرح کرده و منجر به افزایش احتمال تشخیص نادرست خواهد شد. بنابراین، استفاده از دو فاصله‌ی اول قابل قبول به نظر می‌رسد، چرا که از هر تعداد شیفت یا دورافتادگی به اندازه‌ی قابل قبولی تأثیر می‌پذیرد.

برای تعیین قاعده‌ی تصمیم‌گیری مناسب در فاز I نیاز به یک احتمال مشخص برای خطای نوع ۱ است. احتمال خطای نوع ۱ مورد استفاده در یک نمودار کنترل شوهارت با پارامترهای معلوم و حدود سه انحراف معیار عبارت است از:

$$\alpha = 1 - (1 - 2 \times \phi[-3])^m \quad (9)$$

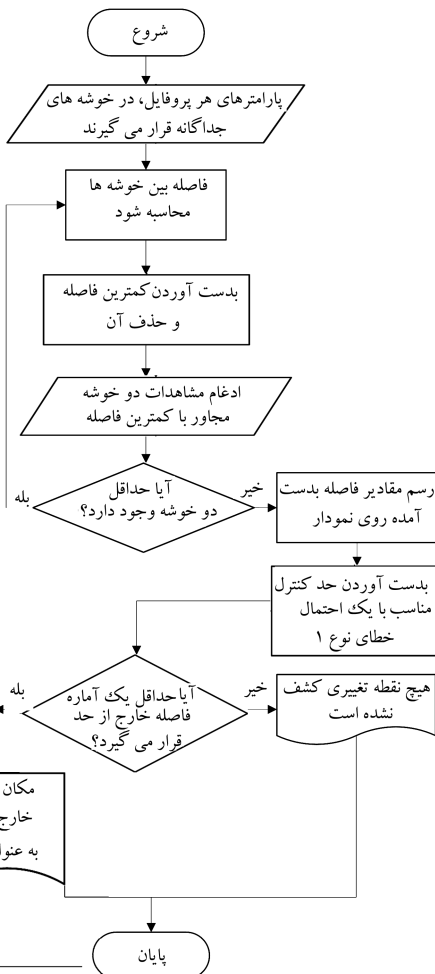
طوری که  $\phi$  تابع توزیع تجمعی نرمال استاندارد است. در این پژوهش حد تصمیم براساس مقدار احتمال خطای نوع ۱ و با استفاده از شبیه‌سازی برآورد می‌شود. در بخش بعد با ارائه‌ی مثالی برای درک الگوریتم، روش تخمین حد کنترل و نتایج شبیه‌سازی برای اثبات اعتبار آن ارائه خواهد شد. در شکل ۱، نمودار جریان الگوریتم خوشه‌بندی نشان داده شده است.

برای به دست آوردن حد کنترل، با فرض این که تغییری در پارامترهای فرایند وجود ندارد و با استفاده از داده‌های مطالعات پیشین<sup>[۴]</sup> با ۱۰۰۰۰ بار شبیه‌سازی، مقادیر  $\max[d_1^*, d_2^*]$  در هر تکرار ذخیره می‌شود. با استفاده از روش به دست آوردن احتمال خطای نوع ۱ در رابطه‌ی ۹، به‌ازای تعداد  $m = 25$  پروفایل در فاز I، مقدار  $\alpha = 0.0654$  حاصل می‌شود. سپس مقدار حد کنترل برابر  $11.10287$  به‌عنوان نتیجه‌ی حاصل از شبیه‌سازی به دست آمد. برای اعتبارسنجی روش شبیه‌سازی، به‌ازای مقادیر مختلف احتمال خطای نوع ۱ و با استفاده از ۱۰۰۰۰ بار شبیه‌سازی نتایج مطلوبی برای حد کنترل حاصل شد (جدول ۱).

### ۴. مثال عددی

در این بخش، نحوه‌ی برآورد نقاط تغییر چندگانه توسط الگوریتم خوشه‌بندی، در قالب یک مثال عددی با دو شیفت پله‌ی چندگانه در عرض از مبدأ پروفایل در زمان‌های ۱۰ و ۲۰ ارائه می‌شود. برای این منظور، از داده‌های مثال کنگ و آلبین<sup>[۴]</sup> استفاده شده است. آنها از مدل پروفایل خطی ساده که در رابطه‌ی ۱۰ ارائه شده بهره‌گرفته‌اند:

$$y_{ij} = 3 + 2x_{ij} + \varepsilon_{ij}, \quad i = 1, \dots, 4, \quad j = 1, 2, \dots, 25 \quad (10)$$



شکل ۱. نمودار جریان الگوریتم خوشه‌بندی.

جدول ۱. شبیه‌سازی احتمال خطای نوع ۱.

مقدار اسمی احتمال خطای نوع ۱	نتیجه شبیه‌سازی
۰٫۰۱	۰٫۰۱۱
۰٫۰۲	۰٫۰۲۰۱
۰٫۰۳	۰٫۰۳۰۲
۰٫۰۵	۰٫۰۵۰۴
۰٫۰۶۵۳	۰٫۰۶۴۲

است. چنان که در نمودار فوق مشخص است آماره‌ی مربوط به  $d_i^*$  خارج از کنترل است و اختلاف بین  $d_i^* - d_{i-1}^*$  مقدار قابل توجهی است. بنابراین، نمودار دو شیفت با مکان‌های تخمینی  $\{l_i^*, i = 1, 2\} = \{20, 9\}$  را شناسایی کرده است.

## ۵. ارزیابی عملکرد

در این قسمت به بررسی عملکرد روش ارائه شده در شناسایی نقاط تغییر چندگانه می‌پردازیم، به همین منظور از شبیه‌سازی در محیط نرم افزار Matlab نسخه ۲۰۰۹a استفاده شده است. مدل پروفایل ارائه شده در رابطه‌ی ۱۰ را در نظر بگیرید؛ در اینجا از مدل مذکور و با فرض  $\varepsilon \sim N(0, 1)$ ،  $m = 25$ ،  $n = 4$  و  $X = 2(2)8$  داده‌ها شبیه‌سازی می‌شوند. برای ارزیابی عملکرد روش خوشه‌بندی در پارامترهای مدل، دو سه تغییر در شیب و عرض از مبدأ ایجاد می‌شود. تغییرات به صورت شیفت پله‌یی چندگانه و مونوتونیک (ترکیبی از صعودی و نزولی) اعمال شده است. به منظور تعیین میزان تغییر در پارامترهای رگرسیون از ضریب  $\lambda$  استفاده شده تا اندازه‌ی تغییر به صورت مضربی از انحراف معیار فرایند بیان شود. برای شیفت در عرض از مبدأ مقدار تحت کنترل  $A$  به  $(\sqrt{\sigma^2 \times (n-1) + \bar{X}^T S_{xx}^{-1}})$  و  $A + \lambda_1$ ، و برای شیفت در شیب مقدار تحت کنترل  $B$  به  $B + \lambda_2 \frac{\sigma}{\sqrt{S_{xx}}}$  تغییر می‌یابد. در این مقاله، انحراف معیار برآورد پارامترهای عرض از مبدأ و شیب به ترتیب برابر با ۱/۲۲۵ و ۰/۲۲۴ است. روند ایجاد تغییر در پارامترها بدین صورت است که در مدل رگرسیون خطی در حالت دو تغییر، تغییرات در نقاط ۱۰ و ۲۰ و در حالت سه تغییر، تغییرات در نقاط ۱۰، ۱۵ و ۲۰ بر پارامترها اعمال می‌شود. نتایج شبیه‌سازی به‌ازای تغییر در عرض از مبدأ و شیب به ترتیب در جداول ۳ و ۴ ارائه شده است. احتمال خطای نوع ۱ و حد کنترل همانند مقادیر به دست آمده در مثال عددی بخش ۴ است.

نحوه‌ی تفسیر نتایج مندرج در جداول ۳ و ۴ بدین صورت است که سه ستون اول از سمت راست با برچسب  $\gamma$  نشان‌دهنده‌ی میزان تغییر مطلق (اندازه‌ی کل تغییر) ایجاد شده در پارامترهای پروفایل است. در این مقاله، بازه‌ی تغییرات در عرض از مبدأ در دامنه‌ی ۰/۲۵ تا ۲/۵ و در شیب در دامنه‌ی ۰/۲۵ تا ۰/۲۵ است. ستون چهارم از سمت راست با برچسب «تعداد نقاط تغییر» بیان‌گر تعداد نقاط بالقوه‌یی است که روش‌های مورد استفاده می‌توانند تخمین بزنند. اگرچه ممکن است تعداد نقاط شناسایی شده بیشتر از ۳ باشد، به منظور جلوگیری از طولانی شدن مطلب تنها برآورد تا ۳ نقطه‌ی تغییر در نظر گرفته شده است. ستون با برچسب  $P$  بیان‌گر درصدی از اوقات است که روش مورد استفاده، تعداد نقاط تغییر بالقوه را برآورد می‌کند. ادامه‌ی جدول شامل برآورد مکان تغییر از ۱ الی ۳ در نظر گرفته شده است. لازم به ذکر است، مقادیری که به‌عنوان برآورد مکان تغییر در هر سلول نمایان است، برابر است با امید ریاضی نقطه‌ی تغییری که در تکرارهای شبیه‌سازی شناسایی شده است. همچنین انحراف استاندارد برآورد برای هر نقطه‌ی تغییر در داخل پرانتز و در سلول زیرین امید ریاضی درج شده است.

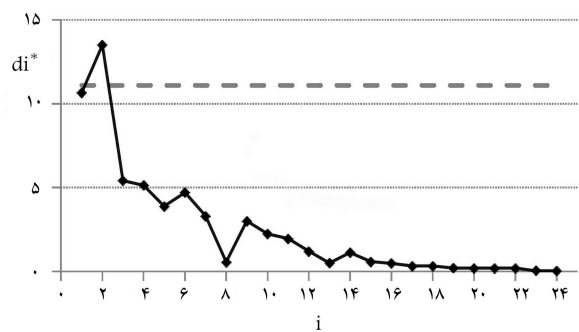
چنان که از جدول ۳ پیداست، برای دو تغییر در عرض از مبدأ به‌ازای تغییرات  $0.75 \leq \gamma_1 < 1$ ، در زمان ۱۰، اگر قدر مطلق اختلاف بین  $\gamma_1$  و  $\gamma_2$  کوچک باشد، آنگاه در اکثر موارد تنها یک تغییر توسط الگوریتم خوشه‌بندی شناسایی می‌شود. به‌عنوان مثال به‌ازای تغییر  $\gamma_1 = 0.5$  در نقطه‌ی ۱۰ و تغییر  $\gamma_2 = 1.5$  در نقطه‌ی ۲۰، الگوریتم در ۸۳/۹۸٪ موارد، یک نقطه‌ی تغییر را شناسایی می‌کند که تقریباً در نیمی از ۸۳/۹۸٪ تغییرات شناسایی شده توسط الگوریتم در پروفایل‌های ۱۹ تا ۲۱ بوده است و دلیل آن تغییر نسبتاً بزرگی است که در پروفایل ۲۰ اعمال شده

مقادیر مفروض پارامترها عبارت‌اند از:  $A = 3$  و  $B = 2$ . بردار  $X_j = (2, 4, 6, 8)$ . در هر پروفایل ثابت و یکسان است. تعداد پروفایل‌ها در فاز  $I$ ،  $m = 25$  و تعداد مشاهدات در هر پروفایل  $n = 4$  است. همچنین فرض می‌شود  $\varepsilon \sim NID(0, 1)$ . با استفاده از مفروضات فوق مقادیر  $y_i$  برای هر پروفایل تولید می‌شود. مطابق مراحل شرح داده شده، در جدول ۲ خلاصه محاسبات ارائه شده است.

مطابق جریان عملیات الگوریتم، ابتدا ۲۴ حد میان ۲۵ خوشه در نظر گرفته می‌شود. در هر خوشه، بردار ضرایب مربوط به هر پروفایل قرار می‌گیرد. معادل فرمول ارائه شده برای محاسبه‌ی فاصله‌ی بین خوشه‌ها در رابطه‌ی ۸، کوچک‌ترین فاصله مربوط به حد ۱۶ است. اطلاعات  $l_i^* = 16$  و  $d_i^* = 0.211$  ثبت می‌شود و حد ۱۶ که فاصله‌ی بین خوشه‌های ۱۶ و ۱۷ است حذف می‌شود و این دو خوشه با یکدیگر ادغام می‌شوند. حال مقداری که به‌عنوان میانگین خوشه در نظر گرفته می‌شود برابر است با متوسط پارامترهای این دو خوشه، یعنی  $(\frac{a_{16}+a_{17}}{2}, \frac{b_{16}+b_{17}}{2})$ . سپس محاسبات فواصل به‌روز می‌شود و مجدداً حدی که کوچک‌ترین مقدار را دارد حذف شده و خوشه‌های مجاور آن ادغام می‌شوند. به همین ترتیب الگوریتم تا جایی ادامه می‌یابد که در نهایت همه‌ی خوشه‌ها ادغام شده و همه‌ی حدود حذف شوند. در شکل ۲ آماره‌های رسم شده روی نمودار و روش پایش آن ارائه شده است. روش تشخیص در این نمودار چنان است که مکان آخرین آماره‌ی خارج از حد کنترل و موقعیت‌های ما قبل آن را خارج از کنترل اعلام می‌کنند. توجه شود که آخرین آماره‌ی خارج از کنترل با آماره‌ی بعد از آن دارای فاصله‌ی قابل توجهی

جدول ۲. خلاصه نتایج مثال عددی.

$d_i^*$	$l_i^*$	$i$	$d_i^*$	$l_i^*$	$i$
۱/۱۱۲۷	۸	۱۴	۱۰/۶۶۶۱	۲۰	۱
۰/۵۶۰۷	۱۷	۱۵	۱۳/۵۰۴۴	۹	۲
۰/۴۶۴۹	۱۰	۱۶	۵/۴۱۳۶	۱	۳
۰/۳۰۶۳	۱۸	۱۷	۵/۱۳۳۳	۲	۴
۰/۳۰۳۹	۵	۱۸	۳/۸۷۰۵	۷	۵
۰/۱۹۳۶	۱۲	۱۹	۴/۷۰۱۸	۶	۶
۰/۱۸۲۱	۱۳	۲۰	۳/۲۸۸۱	۲۱	۷
۰/۱۷۶۶	۱۹	۲۱	۰/۵۴۱۳	۲۲	۸
۰/۱۷۶۱	۱۱	۲۲	۲/۹۹۲۵	۲۴	۹
۰/۰۴۳۴	۲۳	۲۳	۲/۲۲۴۵	۱۴	۱۰
۰/۰۲۱	۱۶	۲۴	۱/۹۴۱۱	۱۵	۱۱
--	--	۲۵	۱/۱۹۰۶	۴	۱۲
			۰/۴۹۱۱	۳	۱۳



شکل ۲. رسم فواصل و حد تصمیم به دست آمده از مثال عددی.

جدول ۳. عملکرد روش خوشه‌بندی در تخمین تعداد و مکان نقاط تغییر به‌ازای ۱۰۰۰۰ بار تکرار شبیه‌سازی (مکان ایجاد دو تغییر در پارامتر عرض از مبدأ در نقاط ۴۱۰ و ۴۰، سه تغییر در نقاط ۱۰، ۱۵ و ۲۰) (انحراف معیار برآوردها در پرانتز نشان داده شده است).

برآورد نقاط تغییر			تعداد نقطه‌ی تغییر		$\gamma_1$	$\gamma_2$	$\gamma_3$
$\hat{\pi}_3$	$\hat{\pi}_2$	$\hat{\pi}_1$	$P(\%)$	تغییر			
			۱,۴۸	۰			
		۱۷,۹۸۹	۸۳,۹۸	۱			
	۲۰,۳۹۶ (۲,۶۶۳)	۹,۴۸۴ (۳,۷۷۱)	۱۱,۲	۲	—	۱,۵	۰,۵
۱۹,۵۶	۱۲,۹۲	۹,۳۵۳	۳,۲۴	۳			
			۰	۰			
		۱۹,۴۴۴	۴۴,۷۲	۱			
	۲۰,۱۵۳ (۰,۸۸۱)	۹,۹۳۱ (۲,۲۳۱)	۵۱,۱۶	۲	—	۲,۲۵	۰,۷۵
۲۰,۱۸۶	۱۳,۶۵۴	۹,۳۸۹	۳,۶۵	۳			
			۰	۰			
		۱۳,۱۲۵	۱۱,۸۹	۱			
	۱۹,۹۸۷ (۱,۱۹۳)	۹,۹۹۷ (۱,۲۱۹)	۸۳,۷۲	۲	—	۲,۵	۱,۲۵
۲۰,۱۵۹	۱۴,۰۲۵	۸,۳۴۹	۳,۶۳	۳			
			۰	۰			
		۲۰,۶۴۲	۰,۱۴	۱			
	۱۹,۹۸۳ (۰,۶۶۲)	۹,۹۹۸ (۰,۸۴۶)	۹۴,۵۲	۲	—	۰,۵	-۱,۵
۲۰,۸۴۳	۱۴,۲۷۶	۷,۹۵۵	۴,۰۸	۳			
			۰,۰۲	۰			
		۱۶,۴۹۱	۶۳,۷۴	۱			
	۲۰,۷۴۳	۱۱,۹۶۹	۳۲,۹۳	۲	۲	۱,۲۵	۰,۵
۱۹,۹۱۹ (۳,۱۸۴)	۱۴,۳۹۹ (۵,۷۳۸)	۹,۶۱۷ (۴,۷۴۱)	۲,۹۸	۳			
			۰	۰			
		۲۰,۰۰۶	۳۱,۳۳	۱			
	۲۰,۰۱۵	۱۵,۱۴۷	۵۰,۸۷	۲	-۲,۵	۰,۷۵	-۰,۵
۲۰,۰۹۳ (۰,۵۵۷)	۱۴,۸۴۶ (۱,۶۱۹)	۱۰,۰۱۱ (۲,۶۶۴)	۱۶,۲۳	۳			
			۱۰,۷۹	۰			
		۱۱,۶۴۵	۲۴,۱۹	۱			
	۱۵,۰۸۵	۱۰,۶۲۹	۱۵,۷۳	۲	-۱	۰	-۱,۲۵
۱۹,۹۴۲ (۱,۳۴۶)	۱۵,۰۶۷ (۱,۰۸۴)	۹,۹۴۲ (۱,۱۷۱)	۴۷,۴۲	۳			

جدول ۴. عملکرد روش خوشه بندی در تخمین تعداد و مکان نقاط تغییر به ازای ۱۰۰۰۰ بار تکرار شبیه سازی (مکان ایجاد دو تغییر در پارامتر شیب در نقاط ۱۰ و ۲۰، سه تغییر در نقاط ۱۰، ۱۵ و ۲۰) (انحراف معیار برآوردها در پرانتز نشان داده شده است).

برآورد نقاط تغییر			تعداد نقطه‌ی تغییر		γ <sub>۳</sub>	γ <sub>۲</sub>	γ <sub>۱</sub>
$\hat{\tau}_3$	$\hat{\tau}_2$	$\hat{\tau}_1$	P(%)	تغییر			
			۳۶,۳۳	۰			
		۱۶,۷۷۱	۵۹,۳۵	۱			
	۲۰,۳۹۴	۱۱,۷۱۷	۳,۰۴	۲	--	۰,۱۵	۰,۰۵
	(۶,۳۵۵)	(۶,۹۲۵)					
۱۸,۷۱	۱۳,۵۴	۹,۸۸۵	۱,۲	۳			
			۱,۲۸	۰			
		۱۴,۲۶	۷۹,۹۲	۱			
	۲۰,۱۱	۱۰,۰۵۶	۱۵,۵۷	۲	--	۰,۲۵	۰,۱۲۵
	(۲,۸۸۵)	(۲,۸۵۸)					
۱۹,۹۲	۱۳,۹۸	۹,۶۷	۳,۱۱	۳			
			۲۶,۳۵	۰			
		۱۸,۳۸۱	۱۷,۸۹	۱			
	۱۹,۴۷۷	۱۰,۳۹۴	۵۲,۹۵	۲	--	۰,۰۵	-۰,۱۲۵
	(۱,۸۶۴)	(۲,۵۶۸)					
۲۰,۷۶۶	۱۳,۱۹۳	۷,۸۶۷	۲,۲۷	۳			
			۰,۱۸	۰			
		۱۰,۰۱۱	۰,۹۳	۱			
	۱۹,۹۹۵	۱۰,۰۲۸	۹۴,۱۱	۲	--	۰	-۰,۲۵
	(۰,۹۰۴)	(۰,۸۴۷)					
۲۰,۷۸۴	۱۴,۵۹۴	۸,۱۳۱	۳,۷۵	۳			
			۲,۳	۰			
		۱۴,۵۹۱	۸۷,۹۵	۱			
	۱۹,۳۷۱	۸,۸۷۸	۷,۵۷	۲	۰,۲۲۵	۰,۱۵	۰,۰۷۵
۱۸,۴۳۹	۱۳,۱۴۱	۹,۲۶۸	۲,۰۵	۳			
	(۳,۵۸۹)	(۵,۷۰۸)	(۴,۲۳۳)				
			۱۶,۹۶	۰			
		۱۹,۴۱۶	۲۵,۲۵	۱			
	۱۹,۴۱۳	۱۴,۹۹۶	۲۴,۱۳	۲	۰,۱۵	-۰,۰۷۵	۰,۱۲۵
۲۰,۰۱۶	۱۴,۸۳۵	۹,۹۱۴	۳۲,۰۴	۳			
	(۰,۹۶۱)	(۱,۱۵۶)	(۱,۹۷۸)				
			۰,۳۲	۰			
		۱۳,۴۶۷	۸,۲۲	۱			
	۱۷,۵۳۹	۱۰,۳۲۷	۳,۸۲	۲	۰,۲۵	۰	۰,۲۵
۱۹,۹۹۷	۱۴,۹۹۶	۹,۹۶۹	۸۴,۵	۳			
	(۰,۷۶۷)	(۰,۷۲)	(۰,۸۰۷)				

هر میزان که باشند به‌خوبی شناسایی خواهند کرد و همچنین انحراف معیار تغییرات شناسایی شده نیز اندک خواهد بود.

جدول ۴ نتایج شبیه‌سازی به‌ازای ۲ و ۳ تغییر در شیب را نشان می‌دهد که نحوه‌ی تفسیر نتایج آن مشابه نتایج جدول ۳ است.

## ۶. نتیجه‌گیری

در این مقاله با استفاده از الگوریتم خوشه‌بندی برآوردی از نقاط تغییر چندگانه در پایش پروفایل خطی ساده ارائه شد. همچنین بررسی‌ها در فاز I پایش پروفایل انجام شد. نتایج ۱۰۰۰۰ بار شبیه‌سازی نشان می‌دهد که الگوریتم مذکور در برآورد نقاط تغییر چندگانه به‌ازای شیف‌های کوچک، هم در عرض از مبدأ و هم در شیب مدل پروفایل خطی ساده عملکرد نسبتاً قابل قبولی دارد ولی مزیت روش فوق عملکرد خوب آن در شناسایی نقاط تغییر چندگانه به‌ازای شیف‌های متوسط تا بزرگ است.

بنابراین، در ادامه‌ی این مطالعه بهبود توان تشخیص تعداد نقاط تغییر در وضعیت خارج از کنترل به‌ازای تغییرات کوچک به‌طور خاص مد نظر نگارندگان است.

با توجه به پتانسیل مناسب تحقیق در حوزه‌ی برآورد نقاط تغییر چندگانه، می‌توان از موارد زیر به‌عنوان فرصت‌هایی برای تحقیقات آتی نام برد:

- برآورد نقاط تغییر چندگانه در زمینه‌ی پایش انواع مختلف پروفایل‌ها (چندمتغیره، چند جمله‌بی و...);
- بررسی تغییرات چندگانه به‌طور همزمان در انحراف معیار مشاهدات;
- در نظر گرفتن تغییرات با روند خطی برای برآورد نقطه‌ی تغییر چندگانه;
- کاربرد روش ارائه شده در شرایط نقض فرض استقلال مشاهدات.

است. در این حال، احتمال شناسایی دو تغییر به‌صورت همزمان برابر ۱۱/۲٪ است. با افزایش قدر مطلق اختلاف بین  $\lambda_1, \lambda_2$  احتمال شناسایی دو نقطه‌ی تغییر توسط الگوریتم افزایش می‌یابد. همچنین انحراف معیار نقاط تغییر شناسایی شده نسبت به تغییرات کوچک کم‌تر است. به‌عنوان مثال به‌ازای تغییرات  $\gamma_1 = 0.5$  و  $\gamma_2 = 0.7$  به‌ترتیب در نقاط ۱۰ و ۲۰، الگوریتم در ۹۴/۵۲٪ از تکرارهای شبیه‌سازی نقاط تغییر را در مکان‌های (۹۹۹۸, ۱۹۹۸۳) شناسایی می‌کند. به‌طور کلی می‌توان گفت، عملکرد روش پیشنهادی به‌ازای تغییرات کوچک نسبتاً قابل قبول است و به‌ازای تغییرات متوسط تا بزرگ، عملکرد خوبی دارد.

ادامه‌ی جدول ۳ شامل ۳ تغییر در عرض از مبدأ در نقاط ۱۰، ۱۵ و ۲۰ است. به‌ازای تغییرات پله‌ی چندگانه به‌صورت صعودی و نزولی، اگر اختلاف بین تغییرات خیلی کوچک باشد در اکثر موارد نقطه‌ی تغییر شناسایی نمی‌شود ولی اگر اختلاف بین تغییرات اعمال شده افزایش یابد حداقل یک نقطه‌ی تغییر شناسایی خواهد شد. به‌عنوان مثال به‌ازای تغییر کوچک  $\gamma_1 = 0.5$  و  $\gamma_2 = 1.0$  و به دنبال آن با ایجاد تغییرات در نقاط ۱۵ و ۲۰ به‌ترتیب با مقادیر  $\gamma_1 = 2$  و  $\gamma_2 = 1.25$  الگوریتم در ۶۳/۷۴٪ از تکرارهای شبیه‌سازی یک نقطه‌ی تغییر نزدیک به نقطه‌ی ۱۵ را شناسایی می‌کند. همچنین به‌ازای این تغییرات ۳۲/۹۳٪ دو نقطه‌ی تغییر نزدیک به مکان‌های ۱۰ و ۲۰ را شناسایی می‌کند. به‌علاوه، در ۲/۹۸٪ موارد، سه تغییر را نزدیک به مکان‌های ۱۰، ۱۵ و ۲۰ شناسایی خواهد کرد. این امر نشان می‌دهد که به‌ازای اعمال تغییرات کوچک با این که احتمال شناسایی سه تغییر اندک است ولی الگوریتم خوشه‌بندی تقریباً در ۹۹٪ از تکرارهای شبیه‌سازی حداقل یکی از تغییرات اعمال شده را شناسایی می‌کند. مزیت الگوریتم خوشه‌بندی در شناسایی تغییراتی است که به‌صورت نامنظم (صعودی و نزولی) اعمال می‌شوند و این تغییرات را به

## پانوشته‌ها

1. statistical process control
2. monotonic
3. step shift
4. linear trend
5. multiple change-point
6. exponentially weighted moving average
7. cumulative sum
8. clustering
9. multivariate exponentially weighted moving average
10. likelihood ratio test

## منابع (References)

1. Amiri, A. and Allahyari, S. "Change point estimation methods for control chart post signal diagnostics: A literature review", *Quality and Reliability Engineering International*, **28**, pp. 673-685 (2011).
2. Noorossana, R., Saghaei, A. and Amiri, A., *Statistical Analysis of Profile Monitoring*, John Wiley & Sons, Inc. (2011).

3. Woodall, W.H. "Current research in profile monitoring", *Revista Producao*, **17**, pp. 420-425 (2007).
4. Kang, L. and Albin, S.L. "On-line monitoring when the process yields a linear profile", *Journal of Quality Technology*, **32**(4), pp. 418-426 (2000).
5. Kim, K., Mahmoud, M.A. and Woodall, W.H. "On the monitoring of linear profiles", *Journal of Quality Technology*, **35**, pp. 317-328 (2003).
6. Chang, T.C. and Gan, F.F. "Monitoring linearity of measurement gauges", *Journal of Statistical Computation and Simulation*, **76**(10), pp. 889-911 (2006).
7. Zhang, J., Li, Z. and Wang, Z. "Control chart based on likelihood ratio for monitoring linear profiles", *Computational Statistics and Data Analysis*, **53**, pp. 1440-1448 (2009).
8. Saghaei, A., Mehrjoo, M. and Amiri, A. "A CUSUM-based method for monitoring simple linear profiles", *International Journal of Advanced Manufacturing Technology*, **45**, pp. 1252-1260 (2009).
9. Sullivan, J.H. "Detection of multiple change points from clustering individual observations", *Journal of Quality Technology*, **34**, pp. 371-383 (2002).



10. Zarandi, M.H.F. and Alaeddini, A. "A general fuzzy-statistical clustering approach for estimating the time of change in variable sampling control charts", *Information Sciences*, **180**(16), pp. 3033-3044 (2010).
11. Ghazanfari, M., Alaeddini, A., Akhavan Niaki, S.T., Aryanezhad, M.B. "A clustering approach to identify the time of a step change in shewhart control charts", *Quality and Reliability Engineering International*, **24**(7), pp. 765-778 (2008).
12. Movaffagh, A. "Estimation of change point in the mean vector and covariance matrix of multivariate normal processes", Master of Science Thesis, Shahed University, Tehran, Iran (2013).
13. Zou, C.L., Zhang, Y.J. and Wang, Z.J. "A control chart based on a change-point model for monitoring linear profiles", *IIE Transactions*, **38**(12), pp. 1093-1103 (2006).
14. Mahmoud, M.A., Parker, P.A., Woodall, W.H. and Hawkins, D.M. "A change point method for linear profile data", *Quality and Reliability Engineering International*, **23**, pp. 247-268 (2007).
15. Zou, C.L., Tsung, F. and Wang, Z.J. "Monitoring general linear profiles using multivariate exponentially weighted moving average schemes", *Technometrics*, **49**(4), pp. 395-408 (2007).
16. Zou, C.L., Qiu, P. and Hawkins, D. "Non-parametric control chart for monitoring profiles using change point formulation and adaptive smoothing", *Statistica Sinica*, **19**, pp. 1337-1357 (2009).
17. Eyvazian, M., Noorossana, R., Saghaie, A. and Amiri, A. "Phase II monitoring of multivariate multiple linear regression profiles", *Published online in Quality and Reliability Engineering International*, **27**(3), DOI:10.1002/qre.1119 pp. 281-296 (2010).
18. Sharafi, A., Amin nayeri, M. and Amiri, A. "A MLE approach for estimating the time of step changes in poisson profiles", *Scientia Iranica*, **20**(3), pp. 855-860 (2013).
19. Sharafi, A., Amin nayeri, M. and Amiri, A. "Identifying the time of step change in binary profiles", *The International Journal of Advance Manufacturing Technology*, **63**, pp. 209-214 (2013).
20. Sharafi, A., Amin nayeri, M., Amiri, A. and Rasouli, M. "Estimating the change point of binary profiles with a linear trend disturbance", *International Journal of Industrial Engineering & Production Research*, **24**(2), pp. 123-129 (2013).
21. Kazemzadeh, R.B., Noorossana, R. and Ayoubi, M. "Change point estimation of multivariate linear profiles under linear drift", *Communications in Statistics-Simulation and Computation, Published Online*, **44**(6), pp. 1570-1599 (2014).
22. Neter, J., Wasserman, W. and Kutner, M.H., *Applied Linear Statistical Models*, 3rd Edition, Richard D. Irwin, Inc, Boston, MA (1990).
23. Johnson, R.A. and Wichern, D.W., *Applied Multivariate Statistical Analysis*, 6th Edition, Pearson Education, Inc (2005).