



## یک نمودار کنترلی چندمتغیره ژرفا مبناي جديد در فاز II

علی هدایتی<sup>(۱)</sup>، محمدرضا فریدروحانی<sup>(۱)</sup> و سکینه دهقان<sup>(۱)</sup>

<sup>(۱)</sup> گروه آمار، دانشکده علوم ریاضی، دانشگاه شهید بهشتی، تهران، ایران

دبیر مسئول: بهزاد منصوری

تاریخ پذیرش: ۱۴۰۳/۱۰/۱۸

تاریخ دریافت: ۱۴۰۳/۰۲/۳۱

**چکیده:** در این مقاله یک نمودار جدید کنترلی ناپارامتری چندمتغیره در فاز دو بر اساس تابع ژرفا ارائه می‌شود. آماره پیشنهادی ناوردای آفین و مجاناً آزاد توزیع است. در واقع، توزیع مجانبی آماره پیشنهادی در حالت تحت کنترل بودن فرآیند، حاصل می‌شود. بر اساس مطالعات شبیه‌سازی، عملکرد آماره پیشنهادی بر اساس چندین تابع ژرفا محاسبه و با سه آماره رقیب مقایسه شده است. نتایج نشان می‌دهد که آماره معرفی شده دارای عملکرد قابل قبولی بوده و در برخی حالات بهتر از آماره‌های رقیب است. بر اساس نمودار پیشنهادی، یک مجموعه داده واقعی مورد تحلیل قرار گرفته است.

**واژه‌های کلیدی:** نمودارهای کنترل کیفیت چندمتغیره، ناوردای آفین، تابع ژرفا، فاز دو.

**رده‌بندی ریاضی:** 62P30; 62G10; 13C15; 62H15

### ۱ مقدمه

نمودارهای کنترل کیفیت، ابزارهای بسیار مفیدی برای پایش فرآیندهای تولید هستند. اکثر نمودارها معرفی شده در حالت یک متغیره و چند متغیره بر فرض نرمال بودن استوارند که در عمل کم اتفاق می‌افتد. به همین دلیل نمودارهای کنترلی ناپارامتری که نسبت به تغییر توزیع استوار بوده و طیف وسیعی از خانواده توزیع‌ها را در بر می‌گیرند، بسیار مورد توجه هستند. چاکرابورتی و همکاران [۶] یک مرور از نمودارهای کنترل کیفیت ناپارامتری یک متغیره تا سال ۲۰۰۰ انجام داده‌اند و چاکرابورتی با گراهام [۷] آن را تا ۲۰۰۷ و چاکرابورتی و گراهام [۸] تا ۲۰۱۶ به‌روز کرده‌اند.

واضح است که عوامل متعددی بر تولید یک محصول موثر هستند، بنابراین نمودارهای کنترل کیفیت چندمتغیره عملکرد بهتری نسبت به نمودارهای یک متغیره در فرآیند پایش تولید دارند. در نمودارهای چند متغیره ناپارامتری نیز مانند حالت یک متغیره بیشتر از رتبه و علامت برای معرفی نمودارها استفاده می‌شود، مانند لیو [۱۶]، کیو و هاوکینز [۲۵] و [۲۶]،

<sup>۱</sup> نویسنده مسئول مقاله (M.R. Faridrohani) [m\\_faridrohani@sbu.ac.ir](mailto:m_faridrohani@sbu.ac.ir)

<sup>۱</sup> نویسنده مسئول مقاله

سان و تسونگ [۳۰]، لیو و همکاران [۱۷]، کیو [۲۷]، زو و تسونگ [۳۳]، زو و همکاران [۳۴] و بون و چاکرابورتی [۳]. بون [۴] مروری بر نمودارهای کنترل چندمتغیره ناپارامتری ارائه داده است. مفهوم ژرفا نخستین بار در سال ۱۹۷۵ توسط توکی [۳۱] معرفی شد که یک ابزار تصویر سازی برای داده‌های دو متغیره بود و داناها و گاسکو [۱۱] آن را به چند متغیره تعمیم دادند. ژرفای یک نقطه، اندازه‌ای برای سنجش فاصله آن نقطه از توده داده‌های موجود است. این یک روش ناپارامتری رتبه‌بندی از مرکز به بیرون برای بردارهای با بعد دلخواه است. پس از معرفی ژرفای نیم‌فضا توسط توکی [۳۱] توابع ژرفای زیادی مانند تابع ژرفای سادگی [۱۸]، تابع ژرفای ماهالانوبیس [۱۹]، تابع ژرفای تصویرمنا، [۲۰] تابع ژرفای زونوئید [۱۴] و ... معرفی شدند. زو و سرفلینگ [۳۶] تعریف رسمی تابع ژرفا را ارائه دادند. اگر  $X = (X_1, \dots, X_p)^T$  یک بردار تصادفی بر فضای احتمال  $(\Omega, \mathcal{F}, P)$  باشد، به ازای  $F \in \mathcal{F}$  دلخواه، تابع ژرفا به صورت:  $D(\cdot, F) : \mathbb{R}^p \rightarrow \mathbb{R}$  تعریف می‌گردد که دارای چهار ویژگی ناوردایی آفین، بیشینه شدن در مرکز، یکنوایی نسبت به ژرفترین نقطه و صفر شدن در بینهایت است. اولین نمودارهای کنترل کیفیت ژرفا مبنا بر اساس توابع ژرفای سادگی و ماهالانوبیس توسط لیو [۱۶] ارائه شد. تا کنون چندین نمودار کنترل کیفیت ژرفا مبنا معرفی شده‌است که می‌توان به [۵] در فاز یک، [۲۳] برای میانگین وزن داده شده نمایی، [۱] برای پایش داده‌های انفرادی، [۲] برای پایش میانگین زیرگروه‌ها، [۱۵]، [۹] برای نمودار تجمعی و [۲۲] برای فاز دو با میانگین معلوم اشاره کرد. دهقان و فرید روحانی [۱۰] یک آزمون ناوردای آفین ژرفا مبنا برای آزمون اختلاف مکان یک نمونه‌ای چندمتغیره معرفی کردند. در این مقاله یک نمودار ناپارامتری چندمتغیره براساس آماره دهقان و فرید روحانی [۱۰] ارائه می‌شود که نمودار پیشنهادی ناوردای آفین و مجانباً آزاد توزیع است. در بخش دوم مقاله آماره و توزیع مجانبی آن معرفی می‌شود. در بخش سوم نمودار کنترل کیفیت ناپارامتری چندمتغیره براساس آماره بخش ۲ ارائه می‌گردد. سرانجام در بخش چهارم مطالعات شبیه‌سازی برای بررسی عملکرد نمودار پیشنهادی ارائه می‌شود.

## ۲ نمودار کنترلی جدید ژرفا مبنا

تابع علامت فضایی چندمتغیره برای هر  $x \in R^p$ ، به صورت زیر تعریف می‌شود:

$$U(x) = \begin{cases} \frac{x}{\|x\|}; & x \neq 0, \\ 0; & x = 0. \end{cases} \quad (1.2)$$

که  $\|X\| = \sqrt{X^T X}$  [۱۳] مانن و اجا [۲۱] بر اساس تابع علامت فضایی یک آماره برای آزمون برای مکان چندمتغیره مطرح کردند. واضح است که آزمون مبتنی بر علامت فضایی به دلیل استفاده از نرم اقلیدسی، ناوردای آفین نیست. هاسجر و کرو [۱۲] و رندلس [۲۹] تبدیلاتی را بر اساس ماتریس واریانس-کوواریانس نمونه‌ای و ماتریس تیلر [۲۲] برای رفع این مشکل پیشنهاد کردند. دهقان و فرید روحانی [۱۰] با جایگزینی یک مقدار دورافتادگی مبتنی بر ژرفا به جای نرم اقلیدسی در تعریف تابع علامت فضایی، یک خانواده‌ی ناوردای آفین از آزمونهای ناپارامتری چند متغیره ژرفا مبنا برای مسئله مکان معرفی کردند. فرض کنید  $\{X_1, \dots, X_m\}$  یک نمونه تصادفی مستقل و هم‌توزیع از بردار  $p$  بعدی اکیداً پیوسته  $X = (X_1, \dots, X_p)^T$ ، که حول  $\mu$  متقارن مرکزی است، باشند. بردار تصادفی  $X$  دارای توزیع متقارن مرکزی حول  $\mu$  است هرگاه  $X - \mu \stackrel{d}{=} \mu - X$ . تابع ژرفای نمونه‌ای  $D(\cdot, \hat{F}_m^s)$  را در نظر بگیرید که  $\hat{F}_m^s$  توزیع تجربی نمونه‌ی متقارن شده  $\{\pm X_1, \dots, \pm X_m\}$  است. متناظر با نرم اقلیدسی که اندازه‌ای از میزان دورافتادگی نقاط است تابع دورافتادگی براساس تابع ژرفا برای نقاط  $x \in R^p$ ، به صورتی که مقدار آن در بازه  $[0, \infty)$  باشد، به صورت زیر تعریف می‌شود:

$$O(X, \hat{F}_m^s) = \frac{1}{D(X, \hat{F}_m^s)} - \frac{1}{a},$$

که  $a \in (0, 1]$  و  $D(\cdot, \hat{F}_m^s) \in [0, a]$  تعریف می‌شود:

$$Q_{mi} = \frac{X_i}{\sqrt{1 + O(X_i, \hat{F}_m^s)}}.$$

دهقان و فریدروhani [۱۰] برای آزمون  $H_0: \mu = \mu_0$  در مقابل  $H_1: \mu \neq \mu_0$  یک رده از آماره های آزمون را به صورت زیر تعریف کردند:

$$T_m = m \bar{Q}_m^T \left( m^{-1} \sum_{i=1}^m Q_{mi} Q_{mi}^T \right)^{-1} \bar{Q}_m, \quad (2.2)$$

که  $\bar{Q}_m = \frac{1}{m} \sum_{i=1}^m Q_{mi}$  آنها نشان دادند، توزیع مجانبی  $T_m$ ، تحت فرض صفر،  $\chi_p^2$  است.

## ۱.۲ نمودار کنترل پیشنهادی در فاز دو

فرض کنید  $X_{ij}$  ها،  $i = 1, \dots, n$  و  $j = 1, \dots, m$ ، نمونه های تصادفی  $p$  بعدی مستقل پیوسته مرجع با اندازه  $n$ ، از زیر گروه هایی با اندازه  $m$ ، باشند. همچنین فرض کنید  $\hat{F}_m^i$ ، تابع توزیع تجربی نمونه ی متقارن شده ی  $\{\pm X_1, \dots, \pm X_m\}$  و  $\hat{F}_m^i$ ، تابع ژرفای نمونه ای باشند. تعریف کنید:

$$Q_{mij} = \frac{X_{ij}}{\sqrt{1 + O(X_{ij}, \hat{F}_m^i)}}, \quad \bar{Q}_{mi} = \frac{1}{m} \sum_{j=1}^m Q_{mij}, \quad \bar{Q} = \frac{1}{n} \sum_{i=1}^n \bar{Q}_{mi} \quad (3.2)$$

که

$$O(X, \hat{F}_m^i) = \frac{1}{D(X, \hat{F}_m^i)} - \frac{1}{a},$$

که  $a \in (0, 1]$  و  $D(., \hat{F}_m^s) \in [0, a]$ ، بنابراین از نمونه های مرجع،  $\bar{Q}$  حاصل می شود که برای سادگی آن  $\theta_0$  نامیده می شود.

فرض کنید  $Y_k$  ها  $k = 1, \dots, m$ ، یک دنباله از بردارهای تصادفی پیوسته  $p$  بعدی از فاز دو فرآیند باشند. فرض کنید  $\hat{F}_m^s$ ، تابع توزیع تجربی نمونه متقارن شده  $\{\pm Y_1, \dots, \pm Y_m\}$ ، باشد. تعریف کنید:

$$R_{mk} = \frac{Y_k}{\sqrt{1 + O(Y_k, \hat{F}_m^s)}}, \quad \bar{R}_m = \frac{1}{m} \sum_{l=1}^m R_{ml}$$

که

$$O(Y, \hat{F}_m^s) = \frac{1}{D(Y, \hat{F}_m^s)} - \frac{1}{a},$$

که  $a \in (0, 1]$  و  $D(., \hat{F}_m^s) \in [0, a]$ ، اکنون یک رده از آماره های کنترلی بر اساس  $\bar{R}_m$ ،  $k = 1, \dots, m$ ، به صورت زیر تعریف می شود:

$$T_m = m(\bar{R}_m - \theta_0)^T \left( m^{-1} \sum_{k=1}^m (R_k - \theta_0)(R_k - \theta_0)^T \right)^{-1} (\bar{R}_m - \theta_0), \quad (4.2)$$

بردارهای  $Q_{mij}$  و  $R_{qk}$ ، بزرگی و جهت نقاط اولیه را حفظ می کنند، بنابراین ساختار هندسی داده ها حفظ می شود. از طرف دیگر چون مخرج آنها بزرگتر از یک است این نقاط نسبت به نقاط اولیه متمرکزتر بوده و در بسیاری موارد از جمله توزیع های دم سنگین که داده های پراکنده تولید می کنند، موجب افزایش توان می شوند [۱۰]. ویژگی مطلوب ناوردایی آفین  $T_m$  در قضیه زیر ارائه می شود.

**قضیه ۱.۲.** اگر تابع ژرفای نمونه ای ناوردای آفین باشد، آنگاه  $T_m$  تحت تبدیلات آفین ناوردا است یعنی

$$T_m(Y_1, \dots, Y_m) = T_m(AY_1, \dots, AY_m),$$

که  $A$  یک ماتریس  $p \times p$ ، نامنفرد دلخواه است.

اثبات: در پیوست ارائه شده است.

حال باید حدود کنترل تعیین شوند و بدین منظور باید توزیع مجانبی آماره حاصل شود. قبل از بیان توزیع  $T_m$  در حالت تحت کنترل به بیان لم زیر نیاز است. بردارهای تصادفی  $\bar{W}_m, W_{mk}$ ؛  $k = 1, \dots, m$  را به صورت زیر تعریف کنید:

$$W_{mk} = \frac{Y_k}{\sqrt{1 + O(Y_k, F)}}, \bar{W}_m = \frac{1}{m} \sum_{k=1}^m W_{mk},$$

$W_{mk}$  همان  $R_{mk}$  است هرگاه به جای تابع توزیع تجربی  $\hat{F}_m^s$  از تابع توزیع  $F$  استفاده شود.

لم ۲.۲. فرض کنید  $Y_k$  ها،  $k = 1, \dots, m$ ، بردارهای  $p$  بعدی مستقل پیوسته از بردار تصادفی  $p$  بعدی  $Y$  با تابع توزیع  $F(y - \theta)$  باشد. همچنین فرض کنید گشتاور دوم مولفه های  $Y_k$  وجود داشته و متناهی است. اگر تابع ژرفای نمونه ای دارای ویژگی ناوردایی آفین باشد و در شرط سازگاری زیر صدق کند

$$\lim_{N \rightarrow \infty} \sup_{y \in R^p} |D(y, \hat{F}_m^s) - D(y, F)| = 0. \quad (5.2)$$

آنگاه

$$(\bar{R}_m - \theta_0) - (\bar{W}_m - \theta_0) = (\bar{R}_m - \bar{W}_m) = o_p(1). \quad (6.2)$$

آنگاه

$$\frac{1}{m} \sum_{k=1}^m (R_{mk} - \theta_0)(R_{mk} - \theta_0)^T - \frac{1}{m} \sum_{k=1}^m (W_{mk} - \theta_0)(W_{mk} - \theta_0)^T = o_p(1). \quad (7.2)$$

اثبات: از لم ۱ دهقان و فریدروحانی [۱۰] نتیجه حاصل می شود.

قضیه ۳.۲. فرض کنید شرایط لم ۲.۲ برقرار باشد، آنگاه وقتی فرآیند تحت کنترل است  $T_m$  دارای توزیع مجانبی  $\chi^2$  با  $p$ ، درجه آزادی است.

اثبات: از قضیه ۲ دهقان و فریدروحانی [۱۰] حاصل می شود.

حال یک نمودار ناپارامتری چندمتغیره براساس آماره  $T_m$  معرفی می شود. برای هر زیرگروه در فاز دو،  $T_m$  محاسبه می شود و هرگاه  $T_k > \chi_{p, 1-\alpha}^2$  فرآیند خارج از کنترل است که  $\chi_{p, 1-\alpha}^2$  چنک  $1 - \alpha$  از توزیع  $\chi^2$  با  $p$ ، درجه آزادی است. مانند بسیاری از نمودارهای کنترل کیفیت چندمتغیره، این نمودار کنترل کیفیت نیز فقط دارای حد بالایی است.

### ۳ مطالعات شبیه سازی

در این بخش عملکرد نمودار کنترل پیشنهادی برای پارامتر مکان در فاز دو با نمودار  $T^2$  هتلینگ [۱۳]، نمودار  $Q$  [۱۶] و نمودار  $NBF$  [۲۲] مقایسه شده است که نمودارهای  $Q$  و  $NBF$  نیز مبتنی بر ژرفا هستند. از توزیع نرمال چند متغیره و توزیع  $t(3)$  چند متغیره با  $p = 2$ ، نمونه هایی به حجم  $20$  و زیرگروه هایی با حجم  $20$  و  $50$  برای فاز یک و دو در  $10000$  تکرار، تولید شده است. به دلیل اینکه آماره  $T_m$  ناوردای آفین است بدون از دست رفتن کلیت  $\Sigma = I_p$  در نظر گرفته می شود. خطای نوع اول اسمی  $5\%$  در نظر گرفته شده است. عملکرد نمودار کنترل پیشنهادی و نمودارهای  $Q$  و  $NBF$  تحت پنج تابع ژرفای ماهالانوبیس، توکی (نیم فضا)، سادکی، تصویر پایه و زونوئید که آنها را به ترتیب با  $T_P, T_S, T_T, T_M$  و  $T_Z$  نشان می دهیم، بررسی می شود. ژرفای نقاط نمونه با استفاده از بسته  $ddalpha$  در نرم افزار  $R$  [۲۴] محاسبه شده اند.

فرض کنید  $X_{ij}$  ها،  $i = 1, \dots, n$  و  $j = 1, \dots, m$ ، نمونه های تصادفی  $p$ ، بعدی مستقل پیوسته مرجع به حجم  $n$ ، از

زیر گروه هایی به حجم  $m$ ، باشند که دارای تابع توزیع تجربی  $\hat{F}_m^i$  و تابع ژرفای  $D(\cdot, \hat{F}_m^i)$ ،  $i = 1, \dots, n$ ، در فاز اول هستند.  $Y_k$  ها  $k = 1, \dots, m$ ، یک دنباله از بردارهای  $p$ ، بعدی مستقل پیوسته، با حجم  $m$ ، در فاز دو فرآیند باشند که دارای تابع توزیع تجربی  $\hat{F}_m^s$ ، هستند آنگاه برای نمودار  $T^2$  هتلینگ دارید:

$$\mu_0 = \frac{1}{mn} \sum_{i=1}^n \sum_{j=1}^m X_{ij} \quad (1.3)$$

سپس تعریف کنید:

$$Z_k = Y_k - \mu_0 \quad (2.3)$$

$$\bar{Z}_m = \frac{1}{m} \sum_{l=1}^m Z_k, \quad (3.3)$$

آنگاه

$$T^2 = (\bar{Z}_m)^T S^{-1} (\bar{Z}_m) \quad (4.3)$$

که  $S$  ماتریس کوواریانس نمونه ای است. حد پائین این آماره  $LCL_{T^2} = 0$  و حد بالای آن به صورت زیر است [۱۳].

$$UCL_{T^2} = \frac{p(m-1)(m+1)}{(m-p)m} F_{1-\alpha, p, m-p} \quad (5.3)$$

که  $F_{1-\alpha, p, M-p}$ ، چندک  $1 - \alpha$  از توزیع  $F$  با  $m - p$  درجه آزادی است. برای نمودار  $Q$  تعریف کنید:

$$r_{\hat{F}_m^i}(y) = \#\{Y_k | D(\cdot, \hat{F}_m^i)(Y_k) \leq D(\cdot, \hat{F}_m^i)(y), k = 1, \dots, m\} / m \quad (6.3)$$

$$Q(\hat{F}_m^i, \hat{F}_m^s) = \frac{1}{m} \sum_{l=1}^m r_{\hat{F}_m^i}(Y_k), \quad (7.3)$$

و هرگاه مینیم  $(m, n)$  به سمت  $\infty$  می رود، توزیع  $[Q(\hat{F}_m^i, \hat{F}_m^s) - \frac{1}{4}]$  به  $N(0, [\frac{1}{m} + \frac{1}{n}])$  میل می کند. در این نمودار حد پائین به صورت زیر است:

$$LCL_Q = 0.5 - Z_\alpha \left( \sqrt{\frac{1}{12} \left[ \frac{1}{m} + \frac{1}{n} \right]} \right) \quad (8.3)$$

و مانند اکثر نمودارهای چندمتغیره، فقط دارای یک حد است [۱۶].

آماره  $NBF$  که بر اساس تابع ژرفا تعریف شده است نیز دارای توزیع مجانبی  $\chi^2$  است و هرگاه مقدار این آماره بزرگتر از  $\chi^2_{p, 1-\alpha}$  چندک  $1 - \alpha$  از توزیع  $\chi^2$  با  $p$ ، درجه آزادی است باشد، آنگاه فرایند خارج از کنترل است.

### ۱.۳ حالت تحت کنترل

عملکرد نمودار کنترل فاز دو معمولاً با طول متوسط اجرا  $ARL$  اندازه گیری می شود. وقتی فرایند در حالت  $IC$  است تعداد نمونه ها تا وقوع اولین هشدار  $OC$  را طول اجرا و میانگین این طول اجراها را  $ARL_0$  می نامند.  $ARL_0$  با خطای نوع اول رابطه دارد. این رابطه اگر نمونه ها از یکدیگر مستقل باشند به صورت  $ARL_0 = \frac{1}{\alpha}$  است [۲۸] و احتمال خطای نوع اول تجربی، نسبت هشدارهای نادرستی است که در ۱۰۰۰۰ بار تکرار مشاهده شده است. چون در این مقاله خطای

نوع اول اسمی را برابر ۵٪ در نظر گرفته‌ایم انتظار داریم تقریباً  $ARL$  برابر با ۲۰ باشد. جدول ۱ عملکرد نمودارهای پیشنهادی،  $Q$  و  $NBF$  را تحت پنج تابع ژرفا و نمودار  $T^2$  هتلینگ را وقتی فرآیند تحت کنترل است نشان می‌دهد که در اینجا نمونه‌های تولیدشده در فاز یک و دو دارای میانگین برابر با بردار صفر هستند. همانگونه که مشاهده می‌شود تحت توزیع نرمال هرگاه  $n = 20$  مقادیر  $ARL$  نمودار پیشنهادی تحت تمام توابع ژرفا نزدیک به مقدار اسمی است. مقادیر  $ARL$  نمودار  $NBF$  نیز تحت هر تابع ژرفا نزدیک به مقدار اسمی ولی کمی کمتر از نمودار پیشنهادی است و عملکرد نمودار  $Q$  تحت تابع ژرفای سادگی مناسب نیست. مقدار  $ARL$ ،  $T^2$  هتلینگ بیشتر از مقدار اسمی است. وقتی مقدار  $m$  به ۵۰ افزایش می‌یابد خطای نوع اول نمودارهای پیشنهادی و  $NBF$  تقریباً صفر می‌شود و  $ARL$  به شدت زیاد می‌شود، خطای آماره  $Q$  همچنان حدود ۵٪ است ولی تحت تابع ژرفای سادگی عملکرد نامناسبی دارد.

نتایج تحت توزیع  $t(3)$  با  $m = 20$  و  $m = 50$  در جدول ۲ ارائه شده‌است. این نتایج نشان می‌دهد که عملکرد نمودار پیشنهادی تحت توزیع دم‌سنگین  $t(3)$  همچنان مناسب و در برخی حالات اندکی دقیق‌تر از نمودار  $NBF$  است. نمودار  $Q$  کمی ضعیف‌تر از نمودار پیشنهادی و همچنان تحت تابع ژرفای سادگی نامناسب است. نتایج نشان می‌دهد نمودار  $T^2$  هتلینگ در این حالت نتوانسته خطای نوع اول را به‌خوبی کنترل کند.

داده‌ها تحت توزیع‌های نرمال و  $t(3)$  با  $m = 70$  نیز تولید و همه آماره‌ها محاسبه شد و تغییر محسوسی در عملکرد آنها نسبت به حالت  $m = 50$  مشاهده نگردید که برای رعایت اختصار گزارش نشد. همه حالت‌های فوق با خطای نوع اول اسمی ۱٪ نیز تولید و همه آماره‌ها محاسبه شد که عملکرد همه آنها مخصوصاً آماره پیشنهادی مشابه عملکردشان با خطای نوع اول اسمی ۵٪ بود و به دلیل کمتر بودن دقت و رعایت اختصار گزارش نشد.

| $T_Z$ | $T_P$ | $T_S$ | $T_T$ | $T_M$ | $m$ |                      |
|-------|-------|-------|-------|-------|-----|----------------------|
| ۲۲/۱۹ | ۲۰    | ۸۱/۲۳ | ۲۸/۲۱ | ۶۱/۱۹ | ۲۰  | آماره پیشنهادی       |
| ۱۰۰۰  | ۱۰۰۰  | ۵۰۰   | ۵۰۰   | ۱۰۰۰  | ۵۰  |                      |
| ۵۴/۱۷ | ۸۷/۱۸ | ۴۵/۶  | ۱۸/۱۸ | ۶۱/۱۹ | ۲۰  | آماره Q              |
| ۷۲/۲۲ | ۷۴/۲۱ | ۹۹/۸  | ۴۱/۲۰ | ۲۲/۲۲ | ۵۰  |                      |
| ۲۴/۱۷ | ۸۷/۱۸ | ۲۸/۲۱ | ۸۶/۱۷ | ۵۲/۱۸ | ۲۰  | آماره NBF            |
| ۱۰۰۰  | ۱۰۰۰  | ۵۰۰   | ۵۰۰   | ۱۰۰۰  | ۵۰  |                      |
| ۷۲/۲۷ |       |       |       |       | ۲۰  | آماره T <sup>2</sup> |
| ۴۵/۴۵ |       |       |       |       | ۵۰  |                      |

جدول ۱. مقایسه مقادیر  $ARL$  وقتی تعداد نمونه  $n = 20$  و اندازه زیر گروه  $m$  است، از توزیع  $N((0, \dots, 0)^T, I_p)$  با  $p = 2$  و  $\alpha = 0.05$ .

| $T_Z$ | $T_P$ | $T_S$ | $T_T$ | $T_M$ | $m$ |                      |
|-------|-------|-------|-------|-------|-----|----------------------|
| ۵۲/۱۸ | ۲۲/۱۹ | ۷۲/۲۲ | ۲۰    | ۵۲/۱۸ | ۲۰  | آماره پیشنهادی       |
| ۱۰۰۰  | ۱۰۰۰  | ۱۰۰۰  | ۱۰۰۰  | ۱۰۰۰  | ۵۰  |                      |
| ۸۶/۱۷ | ۸۷/۱۸ | ۷۵/۷  | ۲۴/۱۷ | ۵۴/۱۷ | ۲۰  | آماره Q              |
| ۷۲/۲۲ | ۲۲/۲۲ | ۱۱/۱۱ | ۸۲/۲۰ | ۲۲/۲۲ | ۵۰  |                      |
| ۸۶/۱۷ | ۵۲/۱۸ | ۸۷/۱۸ | ۸۶/۱۷ | ۳۹/۱۶ | ۲۰  | آماره NBF            |
| ۱۰۰۰  | ۱۰۰۰  | ۱۰۰۰  | ۱۰۰۰  | ۱۰۰۰  | ۵۰  |                      |
| ۶۲/۱۵ |       |       |       |       | ۲۰  | آماره T <sup>2</sup> |
| ۸۷/۱۸ |       |       |       |       | ۵۰  |                      |

جدول ۲. مقایسه مقادیر  $ARL$  وقتی تعداد نمونه  $n = 20$  و اندازه زیر گروه  $m$  است، از توزیع  $t(3)$  با  $\mu = (0, \dots, 0)^T$  و  $p = 2$  و  $\alpha = 0.05$ .

## ۲.۳ حالت خارج از کنترل

وقتی فرایند در حالت  $OC$  است تعداد نمونه‌ها تا وقوع اولین هشدار  $OC$  را طول اجرا و میانگین این طول اجراها را  $ARL_1$  می‌نامند.  $ARL_1$  با خطای نوع دوم رابطه دارد. این رابطه اگر نمونه‌ها از یکدیگر مستقل باشند به صورت  $ARL_1 = 1/(1 - \beta)$  است. [۲۸] و احتمال خطای نوع دوم تجربی، نسبت هشدارهای نادرستی است که در ۱۰۰۰۰ بار تکرار مشاهده شده است. جدول‌های ۳ و ۴، عملکرد نمودارهای پیشنهادی،  $Q$  و  $NBF$  تحت پنج تابع ژرفای ماهالانویس، توکی (نیم فضا)، سادگی، تصویر مینا و زونوئید و نمودار  $T^2$  هتلینگ را وقتی فرایند خارج از کنترل است، نشان می‌دهند که در اینجا میانگین نمونه‌های تولیدشده در فاز یک، برداری با همه مولفه‌های برابر با صفر و میانگین نمونه‌های تولید شده در فاز دو، برداری با همه مولفه‌های برابر با ۵٪ است، در واقع به‌اندازه نیم‌واحد، میانگین‌ها در فاز

دو افزایش داده شده است. همان گونه که مشاهده می شود در جدول ۳، داده ها تحت توزیع نرمال با  $m = ۲۰$  تولید شده اند و با وجود آنکه تغییر، اندک است عملکرد نمودار پیشنهادی مناسب و تغییر به سرعت تشخیص داده شده است. همان گونه که نتایج نشان می دهد عملکرد نمودار پیشنهادی مانند نمودار  $T^2$  هتلینگ که مبتنی بر فرض نرمال بودن است می باشد و اندکی بهتر از نمودار  $NBF$  است. همچنین نتایج نشان می دهد که عملکرد نمودار  $Q$  مناسب نیست. زمانی که مقدار  $m$  به  $۵۰$  افزایش می یابد، عملکرد نمودار پیشنهادی و  $NBF$  در مقایسه با حالت  $m = ۲۰$  تغییر چندانی نمی کند ولی عملکرد  $Q$  بهتر می شود و عملکرد  $T^2$  هتلینگ نیز بهبود می یابد. در جدول ۴ داده ها تحت توزیع  $t(۳)$  با  $m = ۲۰$  تولید شده است با توجه به آنکه تغییر میانگین اندک است اما مقادیر  $ARL_1$  نمودار پیشنهادی نسبتاً مناسب و اندکی بهتر از نمودار  $NBF$  است ولی همچنان نمودار  $Q$  عملکرد مناسب ندارد و نمودار  $T^2$  هتلینگ هم به دلیل دم سنگین بودن توزیع، دارای عملکرد ضعیفتری نسبت به نمودار پیشنهادی و  $NBF$  است. وقتی مقدار  $m$  به  $۵۰$  افزایش می یابد، عملکرد نمودار پیشنهادی و  $NBF$  افت می کند که با توجه به اینکه بر اساس توزیع مجانبی تعریف شده اند غیرعادی است ولی عملکرد  $Q$  تقریباً ثابت اما نامطلوب است و عملکرد  $T^2$  هتلینگ به نحو چشمگیری نسبت به حالت  $m = ۲۰$  بهبود می یابد.

در جدول های ۵ و ۶، عملکرد نمودارهای پیشنهادی،  $Q$  و  $NBF$ ، تحت پنج تابع ژرفای ماهالانویس، توکی (نیم فضا)، سادکی، تصویر پایه و زونوئید و نمودار  $T^2$  هتلینگ را وقتی میانگین نمونه های تولید شده در فاز یک، برداری با همه مولفه های برابر با صفر و در فاز دو برداری با همه مولفه های برابر با  $۰/۵$  است، را نشان می دهند. همان گونه که مشاهده می شود در جدول ۵، داده ها تحت توزیع نرمال با  $m = ۲۰$  تولید شده اند و چون تغییر میانگین بیشتر است عملکرد نمودارهای پیشنهادی،  $NBF$  و  $T^2$  هتلینگ خیلی خوب (بلافاصله تغییر میانگین را اعلام می کنند) و بهتر از نمودار  $Q$  است. وقتی مقدار  $m$  به  $۵۰$  افزایش می یابد عملکرد همه نمودارها باز هم بهتر می شود. در جدول ۶، داده ها تحت توزیع  $t(۳)$  تولید شده اند. عملکرد نمودار پیشنهادی مناسب و کمی بهتر از نمودار  $NBF$  است. از طرفی عملکرد نمودار  $Q$  و  $T^2$  هتلینگ ضعیفتر از نمودار پیشنهادی و  $NBF$  است. با افزایش  $m$  به  $۵۰$  عملکرد همه به خصوص نمودار  $Q$  بهبود می یابد ولی همچنان عملکرد نمودار پیشنهادی بهتر از بقیه است.

| $T_Z$ | $T_P$ | $T_S$ | $T_T$ | $T_M$ | $m$ |                |
|-------|-------|-------|-------|-------|-----|----------------|
| ۴۵/۱  | ۳۸/۱  | ۴۰/۱  | ۴۴/۱  | ۴۱/۱  | ۲۰  | آماره پیشنهادی |
| ۵۲/۱  | ۳۹/۱  | ۲۶/۱  | ۴۹/۱  | ۴۰/۱  | ۵۰  |                |
| ۰۸/۴  | ۳۳/۴  | ۲۶/۲  | ۰۸/۴  | ۳۵/۴  | ۲۰  | آماره Q        |
| ۳۱/۲  | ۳۵/۲  | ۷۲/۱  | ۲۰/۲  | ۳۳/۲  | ۵۰  |                |
| ۵۷/۱  | ۴۰/۱  | ۴۱/۱  | ۵۱/۱  | ۴۱/۱  | ۲۰  | آماره NBF      |
| ۶۰/۱  | ۴۰/۱  | ۴۵/۱  | ۶۳/۱  | ۴۲/۱  | ۵۰  |                |
| ۴۵/۱  |       |       |       |       | ۲۰  | آماره $T^2$    |
| ۰۱/۱  |       |       |       |       | ۵۰  |                |

جدول ۳. مقایسه مقادیر  $ARL_1$  وقتی تعداد نمونه  $n = ۲۰$  و اندازه زیر گروه  $m$  است، از توزیع  $N((۰, ..., ۰)^T, I_p)$  در فاز دو و  $N((۰/۵, ..., ۰/۵)^T, I_p)$  در فاز یک با  $p = ۲$  و  $\alpha = ۰/۵$ .

| $T_Z$ | $T_P$ | $T_S$ | $T_T$ | $T_M$ | $m$ |                |
|-------|-------|-------|-------|-------|-----|----------------|
| ۴۶/۱  | ۵۴/۱  | ۵۶/۱  | ۴۸/۱  | ۵۰/۱  | ۲۰  | آماره پیشنهادی |
| ۴۸/۱  | ۷۴/۱  | ۶۸/۱  | ۵۱/۱  | ۵۵/۱  | ۵۰  |                |
| ۹۴/۳  | ۸۶/۳  | ۱۵/۳  | ۷۷/۳  | ۱۱/۴  | ۲۰  | آماره Q        |
| ۷۹/۳  | ۸۰/۳  | ۴۴/۲  | ۷۰/۳  | ۷۰/۳  | ۵۰  |                |
| ۴۸/۱  | ۵۶/۱  | ۵۸/۱  | ۴۹/۱  | ۵۱/۱  | ۲۰  | آماره NBF      |
| ۷۲/۱  | ۷۷/۱  | ۶۷/۱  | ۷۷/۱  | ۸۲/۱  | ۵۰  |                |
| ۶۱/۲  |       |       |       |       | ۲۰  | آماره $T^2$    |
| ۵۲/۱  |       |       |       |       | ۵۰  |                |

جدول ۴. مقایسه مقادیر  $ARL_1$  وقتی تعداد نمونه  $n = ۲۰$  و اندازه زیر گروه  $m$  است، از توزیع  $t(۳)$  با  $\mu = (۰/۵, ..., ۰/۵)^T$  در فاز دو و  $\mu = (۰, ..., ۰)^T$  در فاز یک با  $p = ۲$  و  $\alpha = ۰/۵$ .

| $T_Z$ | $T_P$ | $T_S$ | $T_T$ | $T_M$ | $m$ |                |
|-------|-------|-------|-------|-------|-----|----------------|
| ۰۰/۱  | ۰۰/۱  | ۰۰/۱  | ۰۰/۱  | ۰۰/۱  | ۲۰  | آماره پیشنهادی |
| ۰۰/۱  | ۰۰/۱  | ۰۰/۱  | ۰۰/۱  | ۰۰/۱  | ۵۰  |                |
| ۱۰/۱  | ۱۰/۱  | ۰۴/۱  | ۱۰/۱  | ۱۰/۱  | ۲۰  | آماره Q        |
| ۰۰/۱  | ۰۰/۱  | ۰۰/۱  | ۰۰/۱  | ۰۰/۱  | ۵۰  |                |
| ۰۰/۱  | ۰۰/۱  | ۰۰/۱  | ۰۰/۱  | ۰۰/۱  | ۲۰  | آماره NBF      |
| ۰۰/۱  | ۰۰/۱  | ۰۰/۱  | ۰۰/۱  | ۰۰/۱  | ۵۰  |                |
| ۰۰/۱  |       |       |       |       | ۲۰  | آماره $T^2$    |
| ۰۰/۱  |       |       |       |       | ۵۰  |                |

جدول ۵. مقایسه مقادیر  $ARL_1$  وقتی تعداد نمونه  $n = ۲۰$  و اندازه زیرگروه  $m$  است، از توزیع  $N((۱, ..., ۱)^T, I_p)$  در فاز دو و  $N((۰, ..., ۰)^T, I_p)$  در فاز یک با  $p = ۲$  و  $\alpha = ۰/۰۵$ .

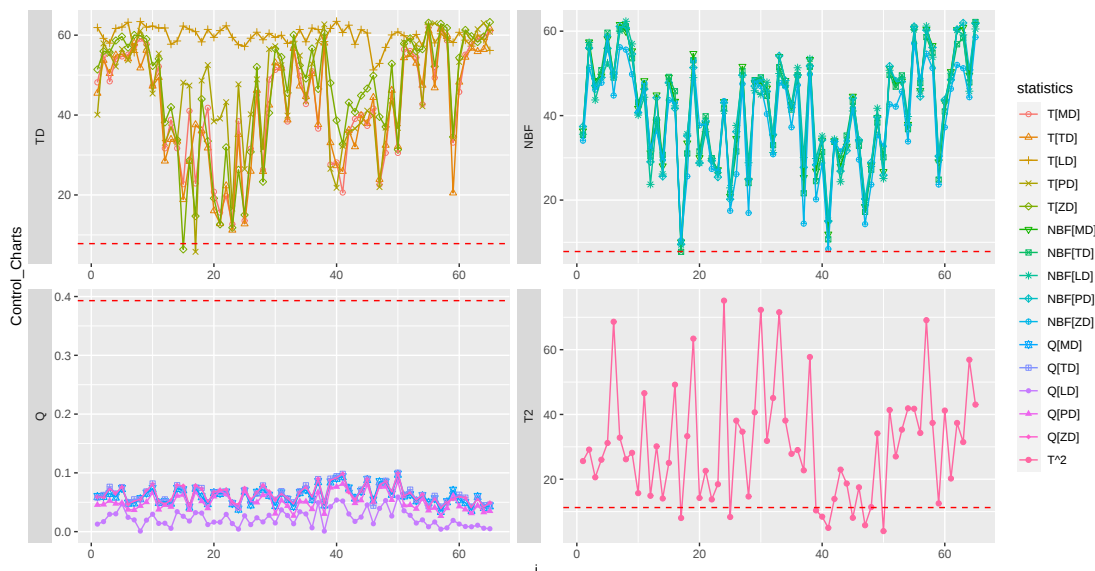
| $T_Z$ | $T_P$ | $T_S$ | $T_T$ | $T_M$ | $m$ |                |
|-------|-------|-------|-------|-------|-----|----------------|
| ۰۲/۱  | ۰۵/۱  | ۰۶/۱  | ۰۲/۱  | ۰۴/۱  | ۲۰  | آماره پیشنهادی |
| ۰۰/۱  | ۰۲/۱  | ۰۰/۱  | ۰۱/۱  | ۰۱/۱  | ۵۰  |                |
| ۳۷/۱  | ۳۶/۱  | ۱۵/۱  | ۳۵/۱  | ۳۸/۱  | ۲۰  | آماره Q        |
| ۰۳/۱  | ۰۳/۱  | ۰۱/۱  | ۰۲/۱  | ۰۳/۱  | ۵۰  |                |
| ۰۳/۱  | ۰۶/۱  | ۰۷/۱  | ۰۴/۱  | ۰۴/۱  | ۲۰  | آماره NBF      |
| ۰۱/۱  | ۰۳/۱  | ۰۳/۱  | ۰۱/۱  | ۰۱/۱  | ۵۰  |                |
| ۱۵/۱  |       |       |       |       | ۲۰  | آماره $T^2$    |
| ۰۳/۱  |       |       |       |       | ۵۰  |                |

جدول ۶. مقایسه مقادیر  $ARL_1$  وقتی تعداد نمونه  $n = ۲۰$  و اندازه زیرگروه  $m$  است، از توزیع  $t(۳)$  با  $\mu = (۱, ..., ۱)^T$  در فاز دو و  $\mu = (۰, ..., ۰)^T$  در فاز یک با  $p = ۲$  و  $\alpha = ۰/۰۵$ .

### ۳.۳ مثال کاربردی

در این بخش، نمودارهای کنترلی پیشنهادی و نمودارهای رقیب برای مجموعه داده‌ی کیفیت یک نوشیدنی به کار می‌روند. مجموعه داده در آدرس <http://archive.ics.uci.edu/ml/datasets/Wine+Quality> در دسترس است. این داده‌ها شامل ۱۱ ویژگی پیوسته است که از اندازه‌گیری‌های فیزیکی و شیمیایی به عنوان شاخص‌های احتمالی کیفیت نوشیدنی به دست آمده‌اند. متغیر ترتیبی کیفیت (دسته‌بندی شده) از ۰ (بسیار بد) تا ۱۰ (عالی) کیفیت را تعیین می‌کند. زو و همکاران [۳۵] روش مبتنی بر نمودار کنترلی را بر برای این داده‌ها به کار بردند. هدف آن‌ها مدل‌سازی و نظارت بر کیفیت بر اساس تست‌های فیزیک و شیمیایی بود. آن‌ها داده‌های سطح هفت (سطح استاندارد کیفیت) را به عنوان داده‌های تحت کنترل در نظر گرفتند و دریافتند که متغیرهای کلرید، غلظت و الکل، مسئول تغییرات کیفیت از سطح «هفت» به «شش» هستند. این نشان می‌دهد که این سه متغیر می‌توانند به عنوان داده‌های مرجع تحت کنترل در نظر گرفته شوند. داده‌ها مربوط به دسته «هفت» به عنوان داده‌های تحت کنترل (فاز یک) و مشاهدات مربوط به سطح کیفیت «شش» و «پنج» به عنوان داده‌های خارج از کنترل استفاده می‌شوند. در این مقاله مشاهدات به زیرگروه‌هایی به اندازه ۲۰ تقسیم شده‌اند. در این صورت، داده‌های فاز یک شامل ۱۰ زیرگروه ۲۰ تایی و داده‌های فاز ۲ شامل ۶۵ زیرگروه ۲۰ تایی است.

نمودارهای کنترلی معرفی شده در بخش شبیه‌سازی همراه با حدود کنترلی متناظر با آنها در شکل ۱ ارائه شده‌اند. در هر پنل در شکل ۱ خط نقطه‌چین قرمز رنگ، حد کنترلی متناظر با آن آماره‌ها است که در واقع برای آماره‌های پیشنهادی، آماره NBF و آماره  $T^2$  هتلینگ نشان‌دهنده‌ی مقدار UCL و برای آماره‌ی  $Q$  مقدار LCL متناظر است. همان‌گونه که مشاهده می‌شود، از ۶۵ نمونه کیفیت سطوح «پنج» و «شش»، نمودارهای پیشنهادی، نمودارهای NBF و نمودارهای  $Q$ ، به درستی یا همه‌ی ۶۵ زیرگروه یا ۶۴ زیرگروه را به عنوان نمونه‌های خارج از کنترل شناسایی کرده‌اند. این درحالی است که نمودار  $T^2$  هتلینگ ۵۷ نمونه را به عنوان خارج از کنترل به درستی شناسایی کرده است. بنابراین، مشاهده می‌کنیم که نمودار پیشنهادی عملکرد قابل قبولی برای این داده‌ها دارد.



شکل ۱: نمودارهای کنترلی در فاز ۲ برای داده‌های کیفیت نوشیدنی، ردیف بالا سمت چپ، نمودارهای پیشنهادی بر اساس توابع ژرفای مختلف، ردیف بالا سمت راست، نمودارهای NBF، ردیف پایین سمت چپ، نمودار Q و ردیف پایین سمت راست نمودار  $T^2$  هتلینگ.

## ۴ نتیجه‌گیری

در این مقاله، یک نمودار کنترلی چندمتغیره ناپارامتری ژرفا مبنا برای پایش مکان در فاز دو معرفی شده است. این نمودار، ناوردای آفین و مجانباً آزاد توزیع است. در مطالعات شبیه‌سازی، عملکرد نمودار پیشنهادی تحت توابع ژرفای مختلف مورد بررسی قرار گرفت. همانگونه که مشاهده شد، عملکرد نمودار پیشنهادی برای داده‌هایی با توزیع نرمال و داده‌هایی با توزیع دم‌سنگین در برخی موارد بهتر از نمودارهای رقیب و در برخی دیگر عملکردی مشابه با آنها دارد. دقت توزیع مجانبی آماره پیشنهادی برای اندازه نمونه کوچک نیز دارای عملکرد مناسبی است و مقادیر بدست آمده برای ARL در حالت تحت کنترل به مقدار مورد انتظار نزدیک هستند. همچنین نتایج نشان داد عملکرد نمودار پیشنهادی، تحت تاثیر تابع ژرفای به‌کار رفته در تعریف آن نمودار نیست. بنابراین می‌توان این نمودار را تابع ژرفای فضایی و تابع ژرفای تصویر مبنا که دارای پیچیدگی محاسباتی نیستند، تعریف کرد و از توابع ژرفای پیچیده که در داده‌هایی با بعدهای بالا قابل استفاده نیستند، صرف‌نظر کرد. در پایان برای یک مجموعه داده واقعی نمودار کنترلی پیشنهادی و نمودارهای رقیب به کار رفتند، که این مثال نیز عملکرد قابل قبول آماره پیشنهادی را نشان داد.

## پیوست

**اثبات قضیه ۱.۲.** فرض کنید  $X_{ij}^* = AX_{ij}$  و  $i = 1, \dots, n$  و  $j = 1, \dots, m$ ، و همچنین فرض کنید  $Y_k^* = AY_k$  و  $O(X_{ij}, \hat{F}_m^i) = O(X_{ij}^*, \hat{F}_m^{i*})$  از ویژگی ناوردای آفین برای تابع ژرفا بدست می‌آید:  $O(Y_k, \hat{F}_m^s) = O(Y_k^*, \hat{F}_m^{s*})$  که  $\hat{F}_m^{s*}$  و  $\hat{F}_m^{i*}$  به ترتیب متناظر با  $\hat{F}_m^s$  و  $\hat{F}_m^i$  و براساس  $X_{ij}^*$  و  $Y_k^*$  تعریف می‌شوند. بنابراین بر اساس داده‌های جدید  $\theta_0$  نیز به  $A\theta_0$  تبدیل می‌شود، پس:

$$T_m(AY_1, \dots, AY_m) = m(A\bar{R}_m - A\theta_0)^T \left( \frac{1}{m} \sum_{k=1}^m (AR_{mk} - A\theta_0)(AR_{mk} - A\theta_0)^T \right)^{-1} (A\bar{R}_m - A\theta_0)$$

$$= m(\bar{R}_{mk} - \theta_0)^T A^T A^{T^{-1}} \left( \frac{1}{m} \sum_{k=1}^m (R_m - \theta_0)(R_m - \theta_0)^T \right)^{-1} A^{-1} A(\bar{R}_{mk} - \theta_0)$$

$$= m(\bar{\mathbf{R}}_{mk} - \boldsymbol{\theta}_o)^T \left( \frac{1}{m} \sum_{k=1}^m (\mathbf{R}_{mk} - \boldsymbol{\theta}_o)(\mathbf{R}_{mk} - \boldsymbol{\theta}_o)^T \right)^{-1} (\bar{\mathbf{R}}_{mk} - \boldsymbol{\theta}_o) = T_m(\mathbf{Y}_1, \dots, \mathbf{Y}_m)$$

## فهرست منابع

- [1] Barale, M.S. and Shirke, D.T., 2023. A control chart based on data depth for monitoring the variability in a multivariate process. *Communications in Statistics-Simulation and Computation*, pp.1-15. <http://dx.doi.org/10.1080/03610918.2023.2185932>
- [2] Bell, R.C., Jones-Farmer, L.A. and Billor, N., 2014. A distribution-free multivariate phase I location control chart for subgrouped data from elliptical distributions. *Technometrics*, 56(4), pp.528-538. <http://dx.doi.org/10.1080/00401706.2013.879264>
- [3] Boone, J.M. and Chakraborti, S., 2012. Two simple Shewhart-type multivariate nonparametric control charts. *Applied Stochastic Models in Business and Industry*, 28(2), pp.130-140. <https://doi.org/10.1002/asmb.900>
- [4] Boone, J.M., 2010. Contributions to multivariate control charting: studies of the z chart and four nonparametric charts. The University of Alabama.
- [5] Capizzi, G. and Masarotto, G. (2024). Distribution-Free Multivariate Phase I Shewhart Control Charts: Analysis, Comparisons and Recommendations. In *Advanced Statistical Methods in Process Monitoring, Finance, and Environmental Science: Essays in Honour of Wolfgang Schmid* (pp. 59-81). Cham: Springer Nature Switzerland.
- [6] Chakraborti, S., Van der Laan, P. and Bakir, S.T., 2001. Nonparametric control charts: an overview and some results. *Journal of quality technology*, 33(3), pp.304-315. <https://doi.org/10.1080/00224065.2001.11980081>
- [7] Chakraborti, S. and Graham, M.A., 2007. Nonparametric control charts. *Encyclopedia of statistics in quality and reliability*, 1, pp.415-429. <https://doi.org/10.1002/9781118445112.stat02699>
- [8] Chakraborti, S. and Graham, M.A., 2019. Nonparametric (distribution-free) control charts: An updated overview and some results. *Quality Engineering*, 31(4), pp.523-544. <https://doi.org/10.1080/08982112.2018.1549330>
- [9] Dai, Y., Zhou, C. and Wang, Z., 2004. Multivariate CUSUM Control Charts Based on Data Depth For Preliminary Analysis. *The Natural Sciences Foundation of Tianjin*, 33603111.
- [10] Dehghan, S. and Faridrohani, M.R., 2019. Affine invariant depth-based tests for the multivariate one-sample location problem. *Test*, 28(3), pp.671-693. <http://dx.doi.org/10.1007/s11749-018-0593-3>
- [11] Donoho, D.L. and Gasko, M., 1992. Breakdown properties of location estimates based on halfspace depth and projected outlyingness. *The Annals of Statistics*, pp.1803-1827. <http://dx.doi.org/10.1214/aos/1176348890>
- [12] Hössjer, O. and Croux, C., 1995. Generalizing univariate signed rank statistics for testing and estimating a multivariate location parameter. *Journal of Nonparametric Statistics*, 4(3), pp.293-308. <http://dx.doi.org/10.1080/10485259508832620>

- [13] Hotelling, H., 1947. Multivariate quality control, techniques of statistical analysis. Eisenhart C, Has-tay HW, Wallis WA, editors.
- [14] Koshevoy, G. and Mosler, K., 1997. Zonoid trimming for multivariate distributions. The Annals of Statistics, 25(5), pp.1998-2017. <http://dx.doi.org/10.1214/aos/1069362382>
- [15] Li, J., Zhang, X. and Jeske, D.R., 2013. Nonparametric multivariate CUSUM control charts for location and scale changes. Journal of Nonparametric Statistics, 25(1), pp.1-20. <http://dx.doi.org/10.1080/10485252.2012.726992>
- [16] Liu, R.Y., 1995. Control charts for multivariate processes. Journal of the American Statistical Association, 90(432), pp.1380-1387.
- [17] Liu, R.Y., Singh, K. and Teng\*, J.H., 2004. DDMA-charts: nonparametric multivariate moving average control charts based on data depth. Allgemeines Statistisches Archiv, 88(2), pp.235-258. <http://dx.doi.org/10.1007/s101820400170>
- [18] Liu, R.Y., 1988. On a notion of simplicial depth. Proceedings of the National Academy of Sciences, 85(6), pp.1732-1734. <http://dx.doi.org/10.1073/pnas.85.6.1732>
- [19] Liu, R.Y. and Singh, K., 1993. A quality index based on data depth and multivariate rank tests. Journal of the American Statistical Association, 88(421), pp.252-260.
- [20] Liu, R.Y., 1992. Data depth and multivariate rank tests, in (ed.) Y. Dodge, L1-Statistical Analysis and Related Methods.
- [21] Möttönen, J. and Oja, H., 1995. Multivariate spatial sign and rank methods. Journal of Nonparametric Statistics, 5(2), pp.201-213.
- [22] Nasrollahzadeh, S., Moghadam, M.B. and Farnoosh, R., 2023. A Shewhart-type nonparametric multivariate depth-based control chart for monitoring location. Communications in Statistics-Theory and Methods, 52(20), pp.7385-7404. <http://dx.doi.org/10.1080/03610926.2022.2045023>
- [23] Nasrollahzadeh, S., Bameni Moghadam, M. and Bayati, M. (2024). An efficient multivariate depth-based EWMA control chart for monitoring location. Journal of Control and Decision, 1-11.
- [24] Pokotylo, O., Mozharovskyi, P. and Dyckerhoff, R., 2016. Depth and depth-based classification with R-package ddalpha. arXiv preprint arXiv:1608.04109. <https://doi.org/10.48550/arXiv.1608.04109>
- [25] Qiu, P. and Hawkins, D., 2001. A rank-based multivariate CUSUM procedure. Technometrics, 43(2), pp.120-132. <https://doi.org/10.1198/004017001750386242>
- [26] Qiu, P. and Hawkins, D., 2003. A nonparametric multivariate cumulative sum procedure for detecting shifts in all directions. Journal of the Royal Statistical Society Series D: The Statistician, 52(2), pp.151-164. <https://doi.org/10.1111/1467-9884.00348>
- [27] Qiu, P., 2008. Distribution-free multivariate process control based on log-linear modeling. IIE Transactions, 40(7), pp.664-677. <http://dx.doi.org/10.1080/07408170701744843>
- [28] Qiu, P., 2013. Introduction to statistical process control. CRC press. <https://doi.org/10.1201/b15016>
- [29] Randles, R.H., 2000. A simpler, affine-invariant, multivariate, distribution-free sign test. Journal of the American Statistical Association, 95(452), pp.1263-1268. <http://dx.doi.org/10.1080/01621459.2000.10474326>

- [30] Sun, R. and Tsung, F., 2003. A kernel-distance-based multivariate control chart using support vector methods. *International Journal of Production Research*, 41(13), pp.2975-2989. <http://dx.doi.org/10.1080/1352816031000075224>
- [31] Tukey, J.W., 1975. Mathematics and the picturing of data. In *Proceedings of the International Congress of Mathematicians, Vancouver, 1975* (Vol. 2, pp. 523-531). [https://doi.org/10.20551/jscswabun.23.2\\_81](https://doi.org/10.20551/jscswabun.23.2_81)
- [32] Tyler, D.E., 1987. A distribution-free M-estimator of multivariate scatter. *The annals of Statistics*, pp.234-251. <http://dx.doi.org/10.1214/aos/1176350263>
- [33] Zou, C. and Tsung, F., 2011. A multivariate sign EWMA control chart. *Technometrics*, 53(1), pp.84-97. <http://dx.doi.org/10.2307/40997295>
- [34] Zou, C., Wang, Z. and Tsung, F., 2012. A spatial rank-based multivariate EWMA control chart. *Naval Research Logistics (NRL)*, 59(2), pp.91-110. <https://doi.org/10.1002/nav.21475>
- [35] Zou, C., W. Jiang, and F. Tsung. 2011. A LASSO-based diagnostic framework for multivariate statistical process control. *Technometrics* 53 (3):297–309. doi:10.1198/TECH.2011.10034.
- [36] Zuo, Y. and Serfling, R., 2000. General notions of statistical depth function. *Annals of statistics*, pp.461-482. <http://dx.doi.org/10.1214/aos/1016218226>



## A New Depth-based Multivariate Control Chart in Phase II

A. hedayati <sup>(1)</sup>, M.R. Faridrohani <sup>(1) 2</sup>, S. Dehghan <sup>(1)</sup>.

<sup>(1)</sup> Department of Statistics, Faculty of Mathematical Sciences, Shahid Beheshti University, Tehran, Iran

Communicated by: Behzad Mansouri

Received: 20/05/2024

Accepted: 07/01/2025

**Abstract:** This article presents a new multivariate non-parametric control chart in Phase II based on depth functions. The proposed statistic is an affine invariant and asymptotically free distribution. Indeed, the asymptotic distribution of the proposed statistic is derived under the in-control process. Based on simulation studies, the performance of the proposed statistic, using several depth functions, has been evaluated and compared with three competing statistics. The results show that the introduced statistic performs adequately and, in some cases, outperforms the competing statistics. A real dataset has been analyzed based on the proposed chart.

**Keywords:** Nonparametric multivariate control chart, Affine invariant, Depth function, Phase II.

---

<sup>2</sup>Corresponding author

E-mail addresses: (M.R. Faridrohani) [m\\_faridrohani@sbu.ac.ir](mailto:m_faridrohani@sbu.ac.ir)