مقایسه نمودار کنترل توأم شوارت و T^2 هتلینگ

فریده توانگر[۱] ، هوشنگ طالبی[۲] ، محمد حسین علامت ساز[۳]

چکیده:

وقتی در یک فرآیند کنترل کیفیت p مشخصه کیفیتی وابسته به طور همزمان اندازه گیری شوند، نمونه می‌تواند شامل بردارهای انفرادی یا گروهی از مشاهدات باشد. در چنین حالتی از روش های کنترل کیفیت چند متغیره استفاده می‌شود، که یکی از آنها روش آماره T^2 هتلینگ است. اما اگر هر یک از مشخصه‌های کیفیتی به طور مجزا در نظر گرفته شوند از نوعی نمودار کنترل شوارت به طور توأم استفاده می‌گردد. در این مقاله با مروری بر این دو روش و ارائه مثال‌هایی از داده‌های واقعی و شبیه سازی شده به مقایسه آن‌ها پرداخته نشان داده می‌شود که نمودار T^2 علاوه بر داشتن توان آزمون بالاتر، همبستگی بین متغیرها را نیز مورد توجه قرار داده و در نتیجه محافظه کارانه‌تر از نمودار شوارت عمل می‌کند.

واژه‌های کلیدی: آماره T^2 هتلینگ، توان آزمون، کنترل کیفیت چند متغیره، نمودار شوارت، نمودار کنترل.

۱ مقدمه

آماره T^2 هتلینگ یکی از معیارهای رایج در نمودارهای کنترل کیفیت چند متغیره در سال های اخیر است. این آماره نه تنها برای مرحله اول، در تهیه نمودار کنترل مبنا و یا کنار گذاشتن داده‌های پرت قابل استفاده است، بلکه برای مرحله دوم، یعنی برای کنترل کیفیت محصول تولید شده و تشخیص دور شدن سطح کیفیت از مقدار هدف، به کار برده می‌شود.

مروری بر کاربرد آماره T^2 را می‌توان در مقاله میسون و یانگ [۲] ملاحظه کرد. اخیراً این آماره، به عنوان تعمیمی از نمودار تک متغیره کنترل (شوارت) به چند متغیره، در کنترل فرآیند نمونه‌های گروهی مواد اولیه تولید مورد توجه قرار گرفته است. مروری بر این موضوع را می‌توان در مقاله وودال و توماس [۸] و مقاله نومیکاس و مک گریگر [۵] یافت.

در مرحله اول برای تهیه نمودار کنترل معمولا دو حالت مورد توجه قرار می‌گیرد. اول اینکه داده‌ها از جامعه‌ای با توزیع نرمال چند متغیره با میانگین‌های

---

[۱] دانشگاه آزاد اسلامی واحد شیراز
[۲] گروه آمار، دانشکده علوم، دانشگاه اصفهان
[۳] گروه آمار، دانشکده علوم، دانشگاه اصفهان

نابرابر ولی واریانس‌های برابر آمده باشند و دوم اینکه داده‌ها از جامعه‌ای با توزیع نرمال چند متغیره یکسان گرفته شده باشد. در مرحله دوم کنترل کیفیت محصول با مشاهدات جدید مورد ارزیابی قرار می‌گیرد. این در حالی است که در مرحله اول نموداری با میانگین و حدود اطمینان مشخص و استاندارد به دست آمده محاسبه و وجود داشته باشد.

در این مقاله برای مرحله اول، آماره $T^{۲}$ در حالت پارامترهای معلوم و نامعلوم فرمول‌بندی و ارائه می‌گردد. آنگاه توزیع نمونه‌ای این آماره و خواص آن که در تهیه نمودار کنترل در مرحله اول و در تشخیص مشاهده خارج از کنترل در مرحله دوم مورد استفاده قرار گرفته و قضایای مربوطه اثبات می‌شود که در پیوست آورده شده‌اند. سپس نمودارهای کنترل ساخته شده، بر مبنای آماره $T^{۲}$، با نمودارهای توأم که برای تک مؤلفه های بردار میانگین به طور مجزا در نظر گرفته شده است، مورد مقایسه قرار می‌گیرند. همچنین نشان می‌دهیم که آماره $T^{۲}$ نسبت به روش نمودارهای توأم مؤلفه‌ها در تشخیص مشاهدات خارج از کنترل محافظ کارانه‌تر رفتار می‌کند. از این رو با توجه به شباهت این نمودارها با فواصل اطمینان و تعمیم آن به آزمون فرض‌ها می‌توان نتیجه گرفت که خطای نوع اول در کاربرد آماره $T^{۲}$ کوچکتر و توان آن در تشخیص مشاهدات پرت و خارج از کنترل بیشتر است. این موارد را همراه با داده‌های شبیه سازی شده و ارائه داده‌های واقعی در قالب مثال ها مورد بحث و مقایسه قرار خواهیم داد. قضایای مربوط به مبنای نظری این مقاله در پیوست همراه با اثبات آن‌ها آمده است.

۲ مبانی نظری

فرض کنید p مشخصه کیفیتی را بخواهیم به طور همزمان اندازه گیری کنیم به طوری که بردار مشخصه‌ها دارای توزیع نرمال p—متغیره با بردار میانگین $\boldsymbol{\mu}= (\mu_{۱}, \mu_{۲}, ..., \mu_p)$ و ماتریس کوواریانس Σ باشد که در آن μ_j میانگین j—امین مشخصه و Σ ماتریسی $p\times p$ متشکل از کوواریانس های دو به دو بین p مشخصه است. در این بخش دو حالت مورد بررسی قرار می‌گیرد.

حالت اول: یکی از حالت‌هایی که می‌تواند در یک فرآیند کنترل کیفیت چند متغیره مورد نظر باشد، کنترل بردارهای انفرادی مشاهدات است. در این صورت برای یک نمونه n تایی از بردارهای تصادفی با میانگین $\boldsymbol{\mu}_i$ با چنین آزمون فرضی روبرو می‌باشیم:

$$H_o : \boldsymbol{\mu}_i = \boldsymbol{\mu} \quad \text{در برابر} \quad H_{۱} : \boldsymbol{\mu}_i \neq \boldsymbol{\mu}$$
$$\text{برای حداقل یک } i \qquad \forall\ i = ۱, ۲, \cdots, n$$

به عبارت دیگر مسئله این است که آیا بردار مشاهده‌ای $\boldsymbol{x}= (x_{۱}, x_{۲}, ..., x_p)$ رفتاری که دلالت بر انحراف مکانی آن از بردار میانگین $\boldsymbol{\mu}= (\mu_{۱}, \mu_{۲}, ..., \mu_p)$ داشته باشد را نشان می دهد؟

در حالت تک متغیره ($p = 1$)، این آزمون به وسیله نمودار کنترل شوارت در سطح معنی داری α انجام می‌گیرد. ممکن است تصور شود که بتوان در حالت چند متغیره نیز با تشکیل p نمودار شوارت برای هر یک از p متغیر در سطح α آزمون فوق را انجام داد ولی چنین روشی، همانطورکه نشان داده خواهد شد، همبستگی بین متغیرها را در نظر نمی‌گیرد بلکه موجب افزایش خطای نوع اول کل از مقدار اولیه هدف می‌شود.

یکی از خواص روش کنترل در حالت چند متغیره ثابت نگه داشتن خطای کل در سطح معین α است. به این منظور آماره نموداری چند متغیره زیر را که آماره T^2 هتلینگ می‌نامند، معرفی می‌کنیم.

$$\boldsymbol{T}^2 = (\boldsymbol{X} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\boldsymbol{X} - \boldsymbol{\mu}) \qquad (1)$$

این آماره با معلوم بودن پارامترها، وقتی که فرآیند تحت کنترل باشد یا $\boldsymbol{\mu}_i = \boldsymbol{\mu}$، دارای توزیع خی دو با p درجه آزادی است و بنابراین با احتمال α از نقطه بحرانی $\chi^2_{p,\alpha}$ فراتر می‌رود [۱]. از این رو، برای ناحیه بحرانی $T^2 > \chi^2_{p,\alpha}$ خطای نوع اول کل دقیقاً در سطح α باقی می‌ماند. این آزمون را می‌توان با یک نمودار از نوع شوارت که دارای حد کنترل بالایی $\chi^2_{p,\alpha}$ است نیز انجام داد. اما اگر $\boldsymbol{\mu}_i \neq \boldsymbol{\mu}$، آن‌گاه T^2 دارای توزیع خی دو غیرمرکزی با p درجه آزادی و پارامتر غیرمرکزی $\lambda = (\boldsymbol{\mu}_i - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu}_i - \boldsymbol{\mu})$ خواهد بود. در این صورت توان آزمون چنین نموداری به λ بستگی دارد.

اما تجربه نشان داده است که در عمل پارامترها نامعلوم‌اند و اغلب اطلاع دقیقی در مورد عملکرد فرآیند دردسترس نیست. در این شرایط، مطالعه و کاربرد توزیع آماره T^2 مفید است. برای این منظور در بخش ۵ مثالی ارائه خواهد شد. با این توضیحات، با فرض نامعلوم بودن پارامترهای جامعه، با استفاده از داده‌های اولیه، برآوردگرهای $\overline{\mathbf{X}}$ و $\mathbf{S}$ را برای $\boldsymbol{\mu}$ و $\boldsymbol{\Sigma}$ به دست می‌آوریم.

حالت دوم: فرض کنید k نمونه مستقل به حجم n، ($n > p$)، از $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ انتخاب شده باشند. با فرض $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \ldots = \boldsymbol{\mu}_k$ آماره T^2 مربوط به بردار میانگین هر نمونه، $\overline{\mathbf{X}}_i$، و توزیع متناظر با آن با استفاده از قضیه ۲ در پیوست عبارت است از:

$$T_i^2 = \frac{kn}{k-1}(\overline{\boldsymbol{X}}_i - \overline{\overline{\boldsymbol{X}}})' S^{-1} (\overline{\boldsymbol{X}}_i - \overline{\overline{\boldsymbol{X}}}) \sim (kn-1)\beta'_{(p,kn-p-1,\tau_i^2)} \qquad ; i = 1, 2, \cdots, k \qquad (2)$$

که در آن:

$$\mathbf{S} = \frac{1}{kn-1} \sum_{i=1}^{k} (\mathbf{X}_i - \overline{\mathbf{X}})(\mathbf{X}_i - \overline{\mathbf{X}})' , \quad \overline{\overline{\mathbf{X}}} = \frac{\sum_{i=1}^{k} \overline{\mathbf{X}}_i}{k}$$

و $kn \geq p + 2$ می باشند.

عبارت $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \ldots = \boldsymbol{\mu}_k$ بیان کننده تحت کنترل بودن فرآیند است. به بیان دیگر وقتی فرض $H_0: \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \ldots = \boldsymbol{\mu}_k$ درست باشد، تحت این فرض با احتمال α آماره T^2 از نقطه بحرانی، $UCL = (kn-1)\beta_{(\alpha, \frac{p}{2}, \frac{kn-p-1}{2})}$[4]، فراتر

[4] Upper Control Limit

می‌رود[۷]. نمودار کنترل $T^۲$ در این حالت مشابه یک نمودار شوارت در حالت تک متغیره است که دارای حد بالایی به اندازه UCL است. از آنجا که به علت ماهیت ساختاری $T^۲$ هر انحرافی از میانگین منجر به افزایش آماره $T^۲$ می‌شود معمولا حد پایین برابر صفر قرار داده می‌شود. توجه کنید که اگر فرآیند خارج از کنترل باشد، این آماره دارای توزیع بتا غیر مرکزی $\beta'_{(p,kn-p-۱,\tau_i^۲)}$ است که در آن $\tau_i^۲ = (\frac{kn}{k-۱})(\boldsymbol{\mu}_i - \overline{\boldsymbol{\mu}})'\boldsymbol{\Sigma}^{-۱}(\boldsymbol{\mu}_i - \overline{\boldsymbol{\mu}})$ پارامتر غیر مرکزی برای $\overline{\boldsymbol{\mu}}= \frac{۱}{k}\sum_{i=۱}^{k}\boldsymbol{\mu}_i$ است.

۳ مطالعه موردی با داده‌های واقعی

در خط مونتاژ بدنه اتومبیل ابتدا بخش‌های بدنه در کف ماشین نصب شده و سپس سقف اتومبیل بر مجموعه کف و بدنه نصب می‌گردد. هر یک از مراحل نصب با اهمیت هستند، زیرا هر گونه انحراف در این بخش از مونتاژ منجر به بروز مشکل در نصب سایر قطعات مانند درب‌ها، شیشه ها و جفت‌گیری آن‌ها می‌گردد.

هدف، ارزیابی و کنترل فرآیند مونتاژ بدنه می‌باشد. با بررسی انجام شده، فرآیند نیاز به کنترل میزان انحراف نصب در بخش پایین بدنه با کف و همچنین میزان انحراف نصب در بخش فوقانی بدنه با سقف اتومبیل دارد. از آنجا که این دو مشخصه همبستگی دارند و تعامل انحرافات ذکر شده آن‌ها منجر به کسب نتایج مختلف توأم در مونتاژ می‌شود، لازم است که از روش‌های کنترل کیفیت چند متغیره استفاده شود. به این منظور دو متغیر $X_۱$ و $X_۲$ را، که $X_۱$ میزان انحراف قطعه بدنه در قسمت سقف و $X_۲$ میزان انحراف قطعه بدنه در قسمت کف اتومبیل می‌باشند، تعریف می‌کنیم. سپس ۱۵ نمونه ۵ تایی به طور متوالی از هر دو مشخصه جمع آوری می‌گردد. این داده‌ها در جدول (۱) آورده شده‌اند. با استفاده از این داده‌ها محاسبات را انجام و نتایج مربوطه در جدول (۲) ارائه شده است. بنابراین براوردهای مورد نیاز، یعنی $\overline{\overline{\mathbf{X}}}$ و S برای براورد $\boldsymbol{\mu}$ و $\boldsymbol{\Sigma}$ عبارت‌اند از:

$$\hat{\mu} = \overline{\overline{\boldsymbol{X}}} = \begin{bmatrix} ۰/۳۲۹۶ \\ -۰/۲۳۴۱ \end{bmatrix}$$

$$\hat{\boldsymbol{\Sigma}} = \boldsymbol{S} = \begin{bmatrix} ۷/۰۲۸۵ & ۲/۱۳۲۹ \\ ۲/۱۳۲۹ & ۵/۰۵۶۶ \end{bmatrix} \quad و$$

برای رسم یک نمودار کنترل، با استفاده از داده های اولیه، مقادیر $T_i^۲$ متناظر با این داده ها از رابطه (۲) به دست آمده و باUCL مقایسه می‌شوند. هر یک از $T_i^۲$ ها که مقدار آن بزرگ‌تر از UCL باشد، بردار میانگین متناظر با آن حذف شده و براورد پارامترها مجدداً محاسبه می‌گردند. مقادیر $T_i^۲$ در جدول (۲) ارائه شده‌اند و مقدار عددی UCL برای $\alpha = ۰/۰۵$ برابر است با:

$$\begin{aligned} UCL &= (۷۵-۱)\times\beta_{۰/۰۵}(۲/۲,(۱۵\times \hookleftarrow \\ &\hookrightarrow ۵-۲-۱)/۲) \\ &= ۷۴\times ۰/۰۷۹۸ = ۵/۹۱ \end{aligned}$$

جدول ۱. میزان انحراف در نمونه ها

#م	انحراف از سقف X_1					انحراف از کف X_2				
	۱	۲	۳	۴	۵	۱	۲	۳	۴	۵
۱	−۰/۹۳۴	۱/۳۸	۲/۳۶	−۰/۶۳	۳/۱۸	−۱/۳۳	۴/۰۷۴	۰/۱۹۴	−۰/۵۴	۰/۴۲
۲	−۱/۴۲۱	−۰/۶۳	−۷/۴۱	−۴/۴۳	۲/۶۷	−۳/۲۵	−۰/۷۴	−۱/۳۱	−۲/۵۴	۰/۹۷
۳	۴/۹۰۱	−۲/۵۸	۲/۲	−۰/۵۹	۱/۷۷	۲/۸۸	−۰/۰۲	۲/۲۵۸	−۲/۹	۲/۴۵
۴	−۲/۱۳۱	−۰/۸۱	−۰/۵۸	−۴/۵۳	۱/۰۸	−۲/۴۲	−۲/۰۹	−۱/۷۱	−۱/۶۵	−۲/۹۹
۵	−۵/۴۴	−۰/۸	۳/۲۴	۰/۴	−۰/۳	−۳/۴۴	−۰/۴۶	۲/۲۹۴	۳/۳۸۵	−۲/۸۳
۶	۴/۵۳۱	−۱/۲۵	۰/۸۵	۱/۹۵	۲/۹	−۰/۸۱	−۰/۹۳	−۰/۰۴	۳/۹۱۵	۲/۲۳
۷	−۰/۴۰۶	−۵/۲	−۲/۵۷	۲/۴۷	−۰/۵	−۰/۲	−۲/۷	۰/۱۳۴	۰/۹۲	−۶/۷۲
۸	۱/۷۹۳	−۱/۱۲	۳/۰۹	−۱/۲۴	۰/۲۶	۰/۶۵	−۰/۳۶	−۲/۵۸	۰/۷۵۹	۱/۶
۹	۲/۸۳۶	−۰/۵۲	۶/۳۲	۱/۹۹	−۱/۵	−۱/۱۹	−۰/۵	۰/۹۷۴	−۱/۹۶	−۲/۸۱
۱۰	−۰/۴۳۳	۴/۶۳	−۱/۹۱	−۲/۵۲	۰/۱۷	۲/۲۸	۱/۹۴۳	۲/۶۶۱	−۳/۱۲	۰/۵۷
۱۱	۰/۶۸۹	۱/۲۶	۴/۴	−۲/۳۱	۲/۱۴	−۴/۲۶	۳/۴۹۵	۰/۹۴۶	۲/۴۵۱	−۰/۳۱
۱۲	۲/۱۵	۰/۴۴	۴/۱۶	−۰/۵۳	−۲/۹	−۲/۶۲	۳/۲۳۷	۰/۲۵۳	−۱/۲۳	۰/۲۱
۱۳	۰/۷۰۶	۰/۸۴	۲/۸۹	۵/۴	۲/۱۷	۰/۷۲	۱/۱۵۶	−۱/۲	۰/۵۳۳	۴/۰۴
۱۴	−۱/۶۱	−۱/۹۴	۱/۳۹	۰/۹۶	−۳/۳	−۱/۸۷	−۱/۷۱	−۳/۰۴	۱/۲۲۳	−۰/۸
۱۵	۲/۲۴۶	−۲/۴۱	۱/۳۹	۰/۹۹	۰/۶۴	۰/۹۲	۰/۱۱۱	−۳/۲۹	۳/۰۳۵	−۲/۹۸

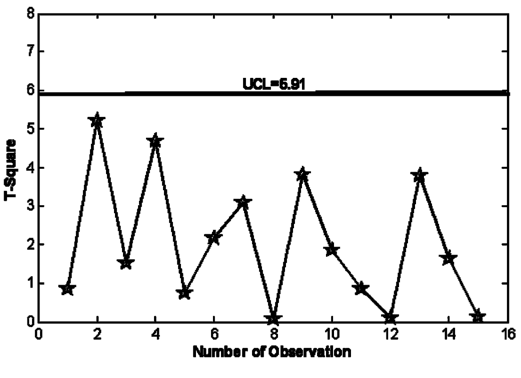
با توجه به مقدار عددی UCL تمامی ۱۵ مشاهده تحت کنترل هستند. نمودار کنترل T^2 که برای این مجموعه از داده‌ها در شکل (۱) رسم شده است مؤید همین معنااست.

جدول ۲. داده های لازم برای آنالیز

#م	$\overline{x}_1$	$\overline{x}_2$	T_i^2
۱	۱/۱۲۵۲	۰/۵۶۳۶	۰/۸۵۸
۲	−۲/۲۴۴	−۱/۳۷۴	۵/۲۰۵
۳	۱/۱۴۰۲	۰/۹۳۳۶	۱/۵۳۳
۴	−۱/۳۹۴	−۲/۱۷۲	۴/۶۹۷
۵	−۰/۵۸	−۰/۲۱	۰/۷۴۰
۶	۱/۷۹۶۲	۰/۸۷۳	۲/۱۷۲
۷	−۱/۲۴۱	−۱/۷۱۳	۳/۱۰۱
۸	۰/۵۵۶۶	۰/۰۱۳۸	۰/۰۷۸
۹	۱/۸۲۵۲	−۱/۰۹۷	۳/۸۱۲
۱۰	−۰/۰۱۳	۰/۸۶۶۸	۱/۸۵۳
۱۱	۱/۲۳۵۸	۰/۴۶۴۴	۰/۸۴۴
۱۲	۰/۶۶۴	−۰/۰۳	۰/۰۹۸
۱۳	۲/۴۰۱۲	۱/۰۴۹۸	۳/۷۹۳
۱۴	−۰/۹	−۱/۲۳۹	۱/۶۳۷
۱۵	۰/۵۷۱۲	−۰/۴۴۱	۰/۱۳۹

۱.۳ مقایسه روش‌ها

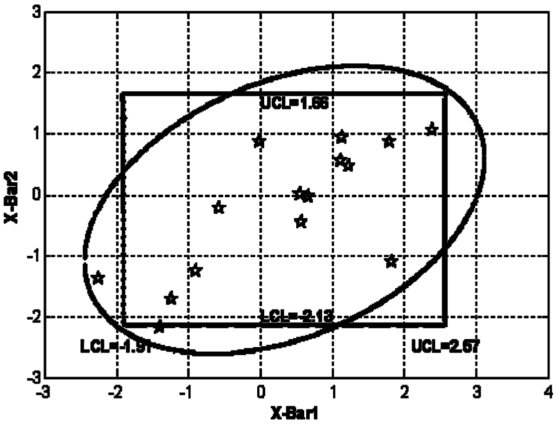
به منظور مقایسه دو روش ذکر شده در مقدمه، یعنی روش استفاده از p نمودار همزمان کنترل یک متغیره شوارت با روش استفاده از نمودار کنترل چند متغیره T^2 هتلینگ، بیضی کنترل مربوط به روش کنترل چند متغیره، به همراه نمودارهای کنترل یک متغیره و همچنین پراکنش داده‌ها در شکل (۲) نمایش داده شده‌اند.



شکل ۱. نمودار T^2

با انجام این مقایسه نشان خواهیم داد که چگونه یک نمودار کنترل T^2 شرایط تحت کنترل بودن سیستم را نشان می‌دهد در حالی‌که استفاده همزمان

دو نمودار کنترل یک متغیره $\overline{X}$، ناتوان از این حقیقت است. توجه شود، در این نمودار دو نقطه‌ای که یکی از نظر $x_{۱}$ و دیگری از نظر $x_{۲}$ خارج از حدود کنترل هستند، سیستم را خارج از کنترل تشخیص داده است.



شکل ۲. ناحیه کنترل تک متغیره و دو متغیره

لازم به ذکر است که از نظر ساختاری تفاوت بیضی و مستطیل به دلیل همبستگی دو متغیر است. زیرا ناحیه بیضی گون با در نظر گرفتن همبستگی دو متغیر به دست آمده در حالی‌که ناحیه مستطیل شکل از بعد نظری دو متغیر را مستقل فرض می‌کند. همان‌طور که مشاهده می‌شود دامنه عملکرد متغیرهای وابسته چند متغیره بزرگ‌تر از حدود نمودار کنترل متغیرهای مستقل است.

در حالت تک متغیره، $p = ۱$، حدود کنترل متناظر با رابطه (۲) برای $\overline{X}$ برابرند با:

$$\overline{\overline{\boldsymbol{X}}} \pm \boldsymbol{S} \quad \sqrt{\frac{(k-۱)(kn-۱)}{kn}}\sqrt{\beta_{(\alpha,\frac{۱}{۲},\frac{kn-۲}{۲})}} = \hookleftarrow$$
$$\hookrightarrow = \overline{\overline{\boldsymbol{X}}} \pm c_{۱}\frac{S}{\sqrt{n}}$$

این نمودار کنترل دقیقاً دارای سطح معنی داری α می‌باشد. در این مثال،

$$\hat{\boldsymbol{\Sigma}} = \boldsymbol{S} = \begin{bmatrix} ۷/۰۲۸۵ & ۲/۱۳۲۹ \\ ۲/۱۳۲۹ & ۵/۰۵۶۶ \end{bmatrix}$$

و

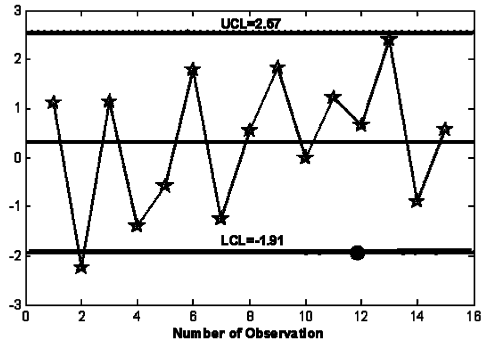
$$\hat{\mu} = \overline{\overline{\boldsymbol{X}}} = \begin{bmatrix} ۰/۳۲۹۶ \\ -۰/۲۳۴۱ \end{bmatrix}$$

بنابراین به منظور کنترل سیستم برای هر تک متغیر حدود کنترل مجزا برای $\overline{X}_{۱}$ و $\overline{X}_{۲}$ با انتخاب $\alpha = ۰/۰۵$ به ترتیب عبارت‌اند از:

$$\overline{\overline{x}}_{۱} \pm c_{۱}\sqrt{\frac{s_{۱۱}}{n}} = ۰/۳۲۹۶ \pm (۱/۸۸۸) \hookleftarrow$$
$$\hookrightarrow \sqrt{\frac{۷/۰۲۸۵}{۵}} = \begin{cases} ۲/۵۷ \\ -۱/۹۱ \end{cases}$$

و

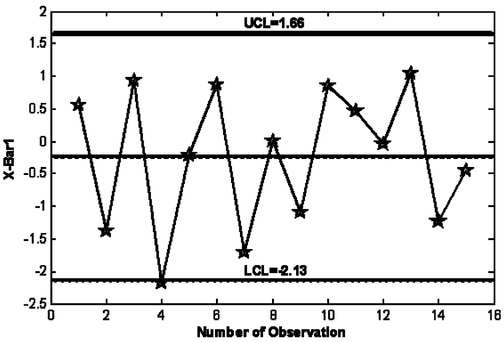
$$\overline{\overline{x}}_{۲} \pm c_{۱}\sqrt{\frac{s_{۲۲}}{n}} = -۰/۲۳۴۱ \pm (۱/۸۸۸) \hookleftarrow$$
$$\hookrightarrow \sqrt{\frac{۵/۰۵۶۶}{۵}} = \begin{cases} ۱/۶۶ \\ -۲/۱۳ \end{cases}$$



شکل ۳. نمودار کنترل $\overline{X}$ برای متغیر اول

نمودارهای تک متغیره فوق در شکل‌های (۳) و (۴) نمایش داده شده‌اند. همان‌طور که مشاهده

می‌شود هر یک از نمودارها یک نقطه خارج از کنترل دارند. البته باید توجه نمود که در اینجا آزمون تنها یک نقطه خارج از ناحیه کنترل مورد نظر بوده است. اطلاعات بیشتر درباره آزمون‌های دیگر در منابع [۳] و [۴] یافت می‌شود.



شکل ۴. نمودار کنترل $\overline{X}$ برای متغیر دوم

در قسمت قبل به دلیل اینکه داده‌های موجود شرایط داده‌های اولیه در مرحله اول یک فرآیند کنترل کیفیت آماری را داشتند، بردار $\overline{\boldsymbol{X}}_i - \overline{\overline{\boldsymbol{X}}}$ و S همبسته فرض شد و با استفاده از قضیه ۲، توزیع بتا برای آماره T^2 به کار برده شد. اما در صورتی که از این همبستگی چشم پوشی کنیم، مطابق قضیه ۱، آماره T^2 با ضریبی ثابت دارای توزیع F به صورت زیر می‌باشد.

$$T_i^2 \;=\; \frac{kn}{k-1}(\overline{\boldsymbol{X}}_i - \overline{\overline{\boldsymbol{X}}})'\boldsymbol{S}^{-1}(\overline{\boldsymbol{X}}_i - \overline{\overline{\boldsymbol{X}}})$$

$$\sim \quad \frac{(kn-1)p}{kn-p}F'_{(p,kn-p,\tau_i^2)} \qquad (3)$$

نادیده گرفتن این همبستگی منجر به حد کنترل تقریبی

$$UCL = \frac{(kn-1)p}{kn-p}F_{(\alpha,p,(kn-p))}$$

تحت فرض H_o برای T_i^2 در (۲) می شود (اثبات در پیوست).

با قرار دادن $p = 1$، حدود کنترل تک متغیره متناظر عبارت‌اند از:

$$\overline{\overline{\boldsymbol{X}}} \pm \sqrt{\frac{k-1}{kn}}\; t_{\frac{\alpha}{2},(kn-1)}\boldsymbol{S} = \overline{\overline{\boldsymbol{X}}} \pm c_2\frac{\boldsymbol{S}}{\sqrt{n}}$$

در جدول (۳) مقادیر c_1 و c_2 به ازای $\alpha = 0/0027$ مقایسه می‌شوند. ملاحظه می‌شود که در صورت استفاده از توزیع بتا به جای حدود تجربی سه انحراف معیار، عددی کوچک‌تر از ۳ باید در محاسبه منظور شود که این واقعیتی غیر محافظه کارانه را در خارج از کنترل تلقی نمودن نمونه (سیستم) بیان می‌کند. همچنین به ازاء هر n داده شده مقادیر c_1 از c_2 کوچک‌تر است.

جدول ۳. مقادیر c_1 و c_2 برای $\alpha = 0/0027$

	$n=1$		$n=2$		$n=3$		$n=4$		$n=5$	
k	c_1	c_2	c_1	c_2	c_1	c_2	c_1	c_2	c_1	c_2
۱۰	۲/۳۷	۳/۸۸	۲/۶۲	۳/۲۷	۲/۷	۳/۱۱	۲/۷۴	۳/۰۴	۲/۷۶	۳
۱۵	۲/۵۹	۳/۵۱	۲/۷۵	۳/۱۷	۲/۸	۳/۰۷	۲/۸۲	۳/۰۳	۲/۸۴	۳
۲۰	۲/۶۹	۳/۳۶	۲/۸۱	۳/۱۲	۲/۸۵	۳/۰۵	۲/۸۷	۳/۰۲	۲/۸۸	۳
۲۵	۲/۷۵	۳/۲۸	۲/۸۵	۳/۱	۲/۸۸	۳/۰۴	۲/۸۹	۳/۰۲	۲/۹	۳
۳۰	۲/۸	۳/۲۳	۲/۸۷	۳/۰۸	۲/۹	۳/۰۳	۲/۹۱	۳/۰۱	۲/۹۲	۳
۳۵	۲/۸۳	۳/۱۹	۲/۸۹	۳/۰۷	۲/۹۱	۳/۰۳	۲/۹۲	۳/۰۱	۲/۹۳	۳
۴۰	۲/۸۵	۳/۱۶	۲/۹۱	۳/۰۶	۲/۹۲	۳/۰۳	۲/۹۳	۳/۰۱	۲/۹۴	۳

جدول ۳. (ادامه) مقادیر $c_۱$ و $c_۲$ برای $\alpha = ۰/۰۰۲۷$

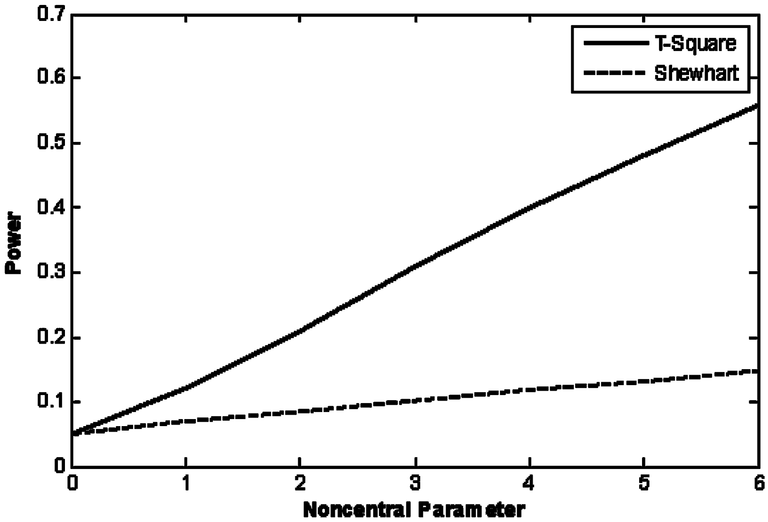
	$n=۶$		$n=۷$		$n=۸$		$n=۹$		$n=۱۰$	
k	$c_۱$	$c_۲$	$c_۱$	$c_۲$	$c_۱$	$c_۲$	$c_۱$	$c_۲$	$c_۱$	$c_۲$
۱۰	۲/۷۷	۲/۹۷	۲/۷۸	۲/۹۶	۲/۷۸	۲/۹۴	۲/۸	۲/۹۳	۲/۸	۲/۹۲
۱۵	۲/۸۴	۲/۹۸	۲/۸۶	۲/۹۸	۲/۸۶	۲/۹۶	۲/۸۶	۲/۹۶	۲/۸۷	۲/۹۵
۲۰	۲/۸۹	۲/۹۹	۲/۸۹	۲/۹۸	۲/۸۹	۲/۹۷	۲/۸۹	۲/۹۶	۲/۹۱	۲/۹۶
۲۵	۲/۹۱	۲/۹۹	۲/۹۱	۲/۹۸	۲/۹۱	۲/۹۸	۲/۹۲	۲/۹۷	۲/۹۲	۲/۹۷
۳۰	۲/۹۲	۲/۹۹	۲/۹۳	۲/۹۸	۲/۹۳	۲/۹۸	۲/۹۳	۲/۹۷	۲/۹۳	۲/۹۷
۳۵	۲/۹۳	۲/۹۹	۲/۹۴	۲/۹۹	۲/۹۴	۲/۹۹	۲/۹۴	۲/۹۹	۲/۹۴	۲/۹۸
۴۰	۲/۹۵	۲/۹۹	۲/۹۵	۲/۹۹	۲/۹۵	۲/۹۹	۲/۹۵	۲/۹۹	۲/۹۵	۲/۹۹

۴ مقایسه توان دو آزمون

در مرحله اول k نمونه n تایی که دارای توزیع نرمال با پارامترهای نامعلوم بودند گرفته شد. به عبارت دیگر

$$\left.\begin{array}{l} X_{۱۱}, ..., X_{۱n} \\ \cdots \\ X_{k۱}, ..., X_{kn} \end{array}\right\} N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

برای محاسبه توان آزمون فرض می‌کنیم نمونه‌ای n تایی مستقل از داده‌های اولیه در حین تولید گرفته شود. این نمونه را می‌توان به صورت $\boldsymbol{X}_{f۱}, ..., \boldsymbol{X}_{fn} \sim \boldsymbol{N}_p(\boldsymbol{\mu}_f, \boldsymbol{\Sigma}_f)$ نمایش داد که در آن $\boldsymbol{\mu}_f$ و $\boldsymbol{\Sigma}_f$ پارامترهای نامعلوم هستند. نمودار کنترل $T^۲$ برای فرآیند جدید (مرحله دوم) فرض ثابت ماندن بردار میانگین در طول زمان را آزمون می‌نماید. یعنی: $H_۰ : \boldsymbol{\mu}_f = \boldsymbol{\mu}$. با فرض $\boldsymbol{\Sigma}_f = \boldsymbol{\Sigma}$، تابع توان نمودار کنترل $T^۲$ احتمال رد شدن فرض $H_۰$ را در سطوح مختلف پارامترهای $\boldsymbol{\mu}$، $\boldsymbol{\mu}_f$ و $\boldsymbol{\Sigma}$ محاسبه می‌کند. البته این احتمال به مقادیر ثابت p, n, k و α بستگی دارد.



شکل ۵. توان آزمون نمودارهای $T^۲$ و توأم شوارت

آماره T^{2} و توزیع آن در این حالت برابر است با:

$$T^{2} = \frac{kn}{k+1}(\overline{\boldsymbol{X}}_i - \overline{\overline{\boldsymbol{X}}})'\boldsymbol{S}^{-1}(\overline{\boldsymbol{X}}_i - \overline{\overline{\boldsymbol{X}}}) \sim \frac{(kn-1)p}{kn-p}F'_{(p,kn-p,\tau^{2})} \qquad (4)$$

که در آن

$$\tau^{2} = (\frac{kn}{k+1})(\boldsymbol{\mu}_f - \overline{\boldsymbol{\mu}})'\Sigma^{-1}(\boldsymbol{\mu}_f - \overline{\boldsymbol{\mu}})$$

و $kn \geq p+1$ فرض می شود (اثبات در پیوست). شکل (۵) توان آزمون مربوط به نمودار T^{2} و نمودارهای توأم شوارت را نمایش می‌دهد.

۵ مطالعه موردی با داده‌های شبیه‌سازی شده

در این بخش، مثالی با ۱۰۰ داده شبیه‌سازی شده از توزیع نرمال دو متغیره ارائه شده که با استفاده از آن به مقایسه توان آزمون در هر دو روش پرداخته می‌شود. فرض کنید $\mathbf{X} \sim N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ به طوری‌که $\boldsymbol{\mu} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$ و $\boldsymbol{\Sigma} = \begin{bmatrix} 0/7983 & 0/6793 \\ 0/6793 & 0/7343 \end{bmatrix}$، که در این صورت ضریب همبستگی این ۱۰۰ داده برای هر زوج برابر ۸۸٪ می باشد.

برای اینکه مقایسه مناسبی انجام شود، مقدار خطای نوع اول را در هر دو روش یکسان قرار می‌دهیم، لذا به ازای α معلوم p نمودار شوارت را به صورت زیر تعریف می‌کنیم.

فرض کنید $p = 2$، مقدار c را طوری تعیین می‌کنیم که رابطه زیر برقرار شود:

$$\begin{aligned} \alpha &= 1 - p(-c \leq \frac{X_1 - \mu_1}{\sigma_{11}} \leq c \cap -c \leq \frac{X_2 - \mu_2}{\sigma_{22}} \leq c) \\ &= 1 - p(-c \leq Z \leq c \cap -c \leq Z \leq c) \\ &= 1 - (\int_{-c}^{c}\int_{-c}^{c} f(z_1, z_2)dz_1dz_2) \qquad (5) \end{aligned}$$

که در آن $f(z_1, z_2)$ تابع چگالی توأم نرمال استاندارد دو متغیره با همبستگی ρ می‌باشد. بنابراین حدود کنترل برای X_1 برابر $\mu_1 \pm c\sigma_{11}$ و برای X_2 برابر $\mu_2 \pm c\sigma_{22}$ است. جدول (۴) مقادیر c را به ازای α و ضریب همبستگی‌های مختلف ارائه می‌دهد.

جدول ۴. مقادیر c

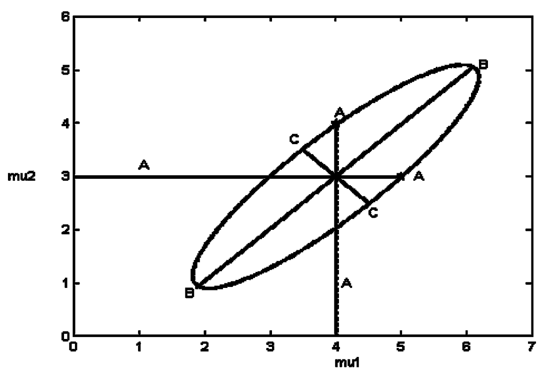
ρ	α			
	۰/۰۱	۰/۰۵	۰/۰۰۳	۰/۰۰۵
۰	۲/۸۰۶	۲/۲۳۶۵	۳/۱۷۴۵	۳/۰۲۳
۰/۱	۲/۸۰۵	۲/۲۳۶	۳/۱۷۴۳	۳/۰۲۲۸
۰/۲	۲/۸۰۵	۲/۲۳۳	۳/۱۷۳۸	۳/۰۲۲۱
۰/۳	۲/۸۰۳	۲/۲۲۸۵	۳/۱۷۲۸	۳/۰۲۰۸
۰/۴	۲/۸	۲/۲۲۲	۳/۱۷۰۹	۳/۰۱۸۳
۰/۵	۲/۷۹۵	۲/۲۱۲	۳/۱۶۷۴	۳/۰۱۴۲
۰/۶	۲/۷۸۵	۲/۱۹۹	۳/۱۶۱۵	۳/۰۰۷۳
۰/۷	۲/۷۷۴	۲/۱۸	۳/۱۵۱۵	۲/۹۹۶۲
۰/۸	۲/۷۵۳	۲/۱۵۲۵	۳/۱۳۴۴	۲/۹۷۷۸
۰/۹	۲/۷۱۵	۲/۱۰۸	۳/۱۰۲۱	۲/۹۴۳۷

همان‌طور که اشاره شد، توان آزمون در نمودار $T^{۲}$ به λ بستگی دارد در صورتی که در نمودارهای همزمان شوارت توان به نوع انحراف بستگی دارد. به عنوان مثال، اگر بخواهیم توان آزمون را برای هر دو روش با مقدار انحراف $\lambda = ۱$ مقایسه کنیم، با تغییر بردار میانگین تحت کنترل بی‌شمار بردار میانگین می‌توان دید که به مقدار $\lambda = ۱$ منتهی می شود. مثلاً بردارهای زیر همگی فاصله یکسان و برابر یک با $\boldsymbol{\mu}$ دارند:

$$\boldsymbol{\mu}_i = \begin{bmatrix} ۴/۴۱۲ \\ ۳ \end{bmatrix}$$
$$\boldsymbol{\mu}_i = \begin{bmatrix} ۴/۲۱۱ \\ ۲/۷۹۵ \end{bmatrix}$$
$$\boldsymbol{\mu}_i = \begin{bmatrix} ۴ \\ ۳/۳۹۶ \end{bmatrix}$$

توان آزمون نمودار کنترل $T^{۲}$ برای تمام این بردارهای خارج از کنترل یکسان است ولی نمودارهای همزمان شوارت توان‌های متفاوتی به ازای هر یک از این بردارها دارند. اما کدام‌یک از این بردارهای خارج از کنترل را می‌بایست برای مقایسه انتخاب نمود ؟ شکل (۶) مقادیر ممکن از هر دو میانگین مورد نظر در فرآیند را نشان می‌دهد. نقاطی که با A نامگذاری شده‌اند، نقاطی هستند که تنها در یکی از میانگین‌ها تغییری داده شده است تا فاصله آن از بردار میانگین تحت کنترل برابر یک شود. نقاط B مکان‌هایی از بیضی را نشان می‌دهد که بیشترین فاصله را با بردار میانگین دارند و نقاط C دارای کمترین فاصله هستند.

پس از بررسی این نقاط، ملاحظه شد که با استفاده از نمودارهای همزمان شوارت و وجود انحرافی مانند نقاط C، توان نمودار کنترل بدترین حالت را دارد. در مقابل، نقاط B بهترین توان را برای کشف انحرافی در میانگین با $\lambda = ۱$ دارند. بنا بر این برای اینکه بتوانیم از هر دو روش استفاده کنیم بهتر است نقاط A را انتخاب کنیم زیرا فاصله بین این نقاط و میانگین تحت کنترل حد وسط نقاط B و C می‌باشد.



شکل ۶. ناحیه کنترل دو متغیره

در این مثال به‌دلیل معلوم بودن پارامترها آماره $T^{۲}$ دارای توزیع خی دو با ۲ درجه آزادی است و توان آزمون عبارت‌است از:

$$\begin{aligned} \text{توان} &= P_{H_{۱}}(T^{۲} \geq UCL) \\ &= P(\chi^{۲}_{p,\lambda} \geq \chi^{۲}_{p,\alpha}) \\ &= P(\chi^{۲}_{۲,\lambda} \geq \chi^{۲}_{۲,۰/۰۵}) \\ &= P(\chi^{۲}_{۲,\lambda} \geq ۵/۹۹۱) \qquad (۶) \end{aligned}$$

توان حاصل از کاربرد نمودار شوارت همزمان ($p = ۲$) با استفاده از رابطه (۵) و تعیین مقدار c عبارت‌است از:

$$\text{توان} = 1 - \int_{\mu_2 - c\sigma_{22}}^{\mu_2 + c\sigma_{22}} \int_{\mu_1 - c\sigma_{11}}^{\mu_1 + c\sigma_{11}} f(x_1, x_2) dx_1 dx_2 \qquad (7)$$

که در آن $f(x_1, x_2)$ تابع چگالی توأم نرمال دو متغیره است.

۱.۵ مقایسه توان دو آزمون

به منظور مقایسه توان آزمون مربوط به هر یک از دو روش، با مقادیر معلوم $\boldsymbol{\mu}, \alpha, p$ و $\boldsymbol{\sigma}$ میزان انحرافی در میانگین را جستجو می کنیم که به ازای آن‌ها λ برابر مقدار معینی گردد. سپس با استفاده از روابط (۶) و (۷) توان آزمون را محاسبه می‌کنیم. جدول (۵) نتایج حاصل را در سطح $\alpha = ۰/۰۵$، برای مثال فوق نشان می‌دهد.

جدول ۵. مقادیر توان آزمون

	λ						
	۰	۱	۲	۳	۴	۵	۶
T^2	۰/۰۵	۰/۱۳۲۷	۰/۲۲۵۶	۰/۳۲۱۶	۰/۴۱۵۵	۰/۵۰۳۷	۰/۵۸۴۱
شوارت توأم							
$\rho = ۰$	۰/۰۵	۰/۰۶۵۷	۰/۰۸۲۳	۰/۰۹۹۸	۰/۱۱۸	۰/۱۳۶۸	۰/۱۵۵۹
$\rho = ۰/۲$	۰/۰۵	۰/۰۶۵۷	۰/۰۸۲۳	۰/۰۹۹۸	۰/۱۱۸۱	۰/۱۳۷	۰/۱۵۶۱
$\rho = ۰/۴$	۰/۰۵	۰/۰۶۵۸	۰/۰۸۲۵	۰/۱۰۰۱	۰/۱۱۸۶	۰/۱۳۷۷	۰/۱۵۷
$\rho = ۰/۶$	۰/۰۵	۰/۰۶۶۲	۰/۰۸۳۴	۰/۱۰۱۵	۰/۱۲۰۵	۰/۱۴۰۲	۰/۱۶
$\rho = ۰/۸$	۰/۰۵	۰/۰۶۷۸	۰/۰۸۶۶	۰/۱۰۶۴	۰/۱۲۶۹	۰/۱۴۸۱	۰/۱۶۹۳

۶ بحث و نتیجه‌گیری

با بررسی حالت‌های فوق نتایج زیر به‌دست می آیند (این نتایج به ازای مقادیر مختلف p و α بر قرارند):

۱— در نظر گرفتن توأم p مشخصه به صورت برداری p—متغیره، با مد نظر قرار دادن همبستگی متغیرها و استفاده از آماره T^2 هتلینگ، منتهی به نموداری چند متغیره به صورت بیضی گون می‌گردد. با ملاحظات مربوط به همبستگی متغیرها در خارج از کنترل تشخیص دادن نقطه نمونه این آماره محافظ کارانه‌تر از حالتی که هر مشخصه به صورت مجزا با نموداری تک متغیره (شوارت) خارج از کنترل نقطه را بیان نماید عمل می‌کند. از بعد نظری این موضوع مربوط به ساختار توزیعی آماره T^2 می شود که با نامعلوم بودن پارامترها و همبستگی $\overline{\boldsymbol{X}}_i - \overline{\overline{\boldsymbol{X}}}$ و $\mathbf{S}$ منجر به توزیع بتا می‌گردد.

۲— با افزایش همبستگی بین متغیرها توان آزمون T^2 نیز افزایش می‌یابد. بنا بر این در صورت وجود همبستگی زیاد انتظار توان بالایی را خواهیم داشت. اما با مقدار همبستگی‌های کوچک تفاوت معنی

داری در اندازه توان مشاهده نمی‌شود.

۳– در صورت ثابت بودن ضریب همبستگی، با افزایش میزان انحراف در میانگین یا λ توان آزمون نیز افزایش یافته به طوری که برای مقادیر بزرگ λ و ρ این افزایش بیش از سه برابرمی‌شود. این موضوع بدین جهت می‌تواند مهم باشد که معمولاً مقادیر بزرگ انحراف باعث عدم تطابق واحدها شده و چنین انحراف هایی می‌بایست هر چه سریع‌تر تشخیص داده شوند.

۴– در صورت وجود همبستگی شدید بین متغیرها و مقدار انحراف زیاد، نمودار T^2 توانمندتر از نمودارهای شوارت توأم می‌باشد. بنا بر این اگر چه نمودار کنترل T^2 پیچیده‌تر از شوارت است ولی روشی بهتر برای کنترل یک فرآیند محسوب می‌شود.

۷ پیوست

قضیه ۱ فرض کنید $T^2 = n\mathbf{X}'\mathbf{W}^{-1}\mathbf{X}$، به طوریکه $\mathbf{X}$ و $\mathbf{W}$ به ترتیب بردار و ماتریس‌های تصادفی مستقل با $\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ و $n > p$ ،$\mathbf{W} \sim \mathbf{W}_p(n, \boldsymbol{\Sigma})$، آنگاه: $T^2 \sim \frac{np}{n-p+1}F'_{(p,n-p+1,\tau^2)}$، که در آن $\tau^2 = \boldsymbol{\mu}'\Sigma^{-1}\boldsymbol{\mu}$ پارامتر غیر مرکزی است[6].

قضیه ۲ فرض کنید $T^2 = n\mathbf{X}'\mathbf{W}^{-1}\mathbf{X}$ که در آن $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ ،$\mathbf{W} \sim \mathbf{W}_p(n, \boldsymbol{\Sigma})$، به طوریکه $\mathbf{W} = \mathbf{W}' + \mathbf{X}\mathbf{X}'$، وقتی $\mathbf{W}' \sim \mathbf{W}_p(n-1, \Sigma)$ و $\mathbf{X}$ مستقل از یکدیگر باشند. آنگاه، $T^2 = n\beta'_{(p,n-p,\tau^2)}$ به طوریکه $\tau^2 = \boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$ پارامتر غیر مرکزی است.

اثبات:

$$\begin{aligned} T^2 &= n\boldsymbol{X}'\boldsymbol{W}^{-1}\boldsymbol{X} = n\boldsymbol{X}'(\boldsymbol{W}' + \boldsymbol{X}\boldsymbol{X}')^{-1}\boldsymbol{X} \\ &= n\boldsymbol{X}' \left(\boldsymbol{W}^{-1} - \tfrac{1}{1+X'W^{-1}X} \hookleftarrow \right. \\ &\hookrightarrow \left. \boldsymbol{W}^{-1}\boldsymbol{X}\boldsymbol{X}'\boldsymbol{W}^{-1} \right) \boldsymbol{X} \\ &= \tfrac{nX'W^{-1}X}{1+X'W^{-1}X} \end{aligned}$$

با استفاده از قضیه ۱ داریم:
$n\boldsymbol{X}'\boldsymbol{W}^{-1}\boldsymbol{X} \sim \frac{np}{n-p}F'_{(p,n-p,\tau^2)}$ که در آن $\tau^2 = \boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$. بنا بر این، توزیع T^2 برابر است با:

$$\begin{aligned} T^2 &\sim \frac{\frac{np}{n-p}F'_{(p,n-p,\tau^2)}}{1 + \frac{p}{n-p}F'_{(p,n-p,\tau^2)}} \\ &\sim \frac{n\chi'^2_{(p,\tau^2)}}{\chi^2_{(n-p)} + \chi'^2_{(p,\tau^2)}} \end{aligned}$$

لذا با توجه به اینکه دو متغیر تصادفی خی دو مذکور مستقل از یکدیگر هستند و با استفاده از تعریف توزیع بتا غیر مرکزی اثبات قضیه کامل می‌شود.

۳– اثبات رابطه (۳):

$$V(\overline{\boldsymbol{X}}_i - \overline{\overline{\boldsymbol{X}}}) = \frac{1}{n}\boldsymbol{\Sigma} + \frac{1}{kn}\boldsymbol{\Sigma} - \frac{2}{kn}\boldsymbol{\Sigma} = \frac{k-1}{kn}\boldsymbol{\Sigma}$$

زیرا $COV(\overline{\mathbf{X}}_i - \overline{\overline{\mathbf{X}}}) = \frac{1}{kn}\boldsymbol{\Sigma}$ و اگر $\overline{\boldsymbol{\mu}} = \frac{1}{k}\sum_{i=1}^{k}\boldsymbol{\mu}_i$ تعریف کنیم، داریم:

$$\sqrt{\frac{kn}{kn-1}}(\overline{\mathbf{X}}_i - \overline{\overline{\mathbf{X}}}) \sim N_p\left(\sqrt{\frac{kn}{kn-1}}(\overline{\boldsymbol{\mu}}_i - \overline{\overline{\boldsymbol{\mu}}}), \boldsymbol{\Sigma}\right)$$

و می دانیم: $(kn-1)\mathbf{S} \sim W_p(kn-1, \boldsymbol{\Sigma})$ که با استفاده از قضیه ۱ رابطه (۳) اثبات می‌شود.

۴- اثبات رابطه (۴):

$V(\overline{\mathbf{X}}_i - \overline{\overline{\mathbf{X}}}) = \frac{1}{n}\boldsymbol{\Sigma} + \frac{1}{kn}\boldsymbol{\Sigma} = \frac{k+1}{kn}\boldsymbol{\Sigma}$ بنا بر این $\sqrt{\frac{kn}{k+1}}(\overline{\mathbf{X}}_i - \overline{\overline{\mathbf{X}}}) \sim N_p\left(\sqrt{\frac{kn}{k+1}}(\overline{\boldsymbol{\mu}}_i - \overline{\overline{\boldsymbol{\mu}}}), \boldsymbol{\Sigma}\right)$ و مجدداً با استفاده از قضیه ۱ اثبات کامل است.

مراجع

[1] Mardia, K.V., Kent, J.T. and Bibby, J.M. (1979), *Multivariate Analysis*, Academic press, NewYork.

[2] Mason, R.L. and Young, J.C.(1998), Hotelling's T2:A Multivariate Atatistic for Industrial Process Control, *Processdingsoft the 52nd Annua Quality Congress*, 78-85.

[3] Nelson, L.S. (1984). The shewhart control chart-tests for special causes, *Journal of Quality Technology*, 16, 237-239.

[4] Nelson, L.S. (1985), Interpreting shewart $\overline{X}$ control chart, *Journal of Quality Technology*, 17, 114-116.

[5] Nomikos, P. and MacGregor, J.F.(1995), Multivariate SPC charts for monitoring batch processes, *Technometrics 37*, 41-59.

[6] Seber, G.A.F.(1984), *Multivariate Observations*, Wiley, NewYork.

[7] Tracy, N.D., Young, J.C., and Mason, R.L.(1992), Multivariate control chart for individual observations, *Journal of Quality Technology*, 24, 88-95.

[8] Woodall, W.H. and Thomas, E.V.(1995), Statistical process control with several components of common cause variability, *IIE Transactions* 27, 757-764.