



Profile Monitoring for Multinomial Responses

R. Noorossana* & H.R. Izadbakhsh

Rassoul Noorossana, Professor Iran University of Science and Technology, Faculty of Industrial Engineering, rassoul@iust.ac.ir

HamidReza Izadbakhsh, PhD student Iran University of Science and Technology, Faculty of Industrial Engineering, hizadbakhsh@iust.ac.ir

Keywords

Statistical Control Process,
Signal Probability,
Logistic Regression, Profile,
Control Chart, Hotelling,
Likelihood ratio test

ABSTRACT

In certain statistical process control applications, quality of a process or product can be characterized by a function between response variable and one or more independent variables. This function commonly referred to as profile. Response variable can be continuous or discrete. All of the research assumes that the response variable is continuous. Whereas, some of the potential applications of profile monitoring are cases where the output can be modeled using polytomous (especially multinomial) or binary logistic regression models. Polytomous response variables, especially multinomial variables, can have various applications especially in service industry. In this paper, we propose some methods for monitoring a profile when the process output is a multinomial response variable. Multinomial logistic regression (OLR) provides the basis for our profile model. Performances of the proposed methods in terms of the signal probability for different out-of-control scenarios is compared based on simulation studies.

© 2013 IUST Publication, IJIEPM. Vol. 23, No. 4, All Rights Reserved

*
Corresponding author. Rassoul Noorossana
Email: rassoul@iust.ac.ir

پایش پروفایل با پاسخ چند رسته‌ای اسمی

رسول نورالسنا* و حمیدرضا ایزدبخش

چکیده:

کلمات کلیدی

در بسیاری از فرآیندهای تولیدی و خدماتی، کیفیت محصول یا عملکرد فرآیند به کمک رابطه‌ای خطی یا غیرخطی، بین متغیر پاسخ و یک یا چند متغیر مستقل، بیان می‌گردد. به چنین روابطی در اصطلاح، پروفایل می‌گویند. متغیر پاسخ در رابطه پروفایل می‌تواند پیوسته یا گسسته باشد. تحقیقات نسبتاً گسترده‌ای تا بحال در رابطه با متغیرهای پاسخ پیوسته انجام شده ولی متغیرهای پاسخ گسسته کمتر مورد توجه قرار گرفته است. در حالی که کاربردهای بالقوه‌ای وجود دارد که متغیر پاسخ با مدل‌های رگرسیون لجستیک رسته‌ای یا دودویی مدلسازی می‌شوند. متغیرهای پاسخ رسته‌ای کاربردهای متنوعی به ویژه در صنعت خدمات دارند. در این مقاله ابتدا به منظور پایش پروفایل‌ها در فاز I، نمودارهای کنترل پیشنهادی برای پایش پروفایل‌هایی که دارای پاسخ چندرسته‌ای اسمی هستند توسعه داده شده‌اند. مدل مبنای پروفایل، رگرسیون لجستیک چندرسته‌ای اسمی در نظر گرفته شده است. عملکرد نمودارهای کنترل با استفاده از شبیه سازی و احتمال هشدار برای سناریوهای مختلف خارج از کنترل مقایسه شده‌اند.

کنترل کیفیت آماری،
احتمال هشدار،
رگرسیون لجستیک،
پروفایل،
نمودار کنترل T^2 هتیلینگ،
آزمون نسبت درستیابی

۱. مقدمه

در بسیاری از کاربردهای کنترل فرآیند آماری، از تابع توزیع یک یا چند مشخصه کیفی برای توصیف کیفیت محصول یا عملکرد فرایند استفاده می‌شود. مطالعات اخیر نشان می‌دهد که موارد متعددی را می‌توان یافت که عملکرد یک فرایند یا کیفیت یک محصول به وسیله رابطه بین متغیر پاسخ و یک یا چند متغیر مستقل توصیف می‌شود که این رابطه را پروفایل می‌نامند. به عنوان مثال، می‌توان به تحقیقات انجام شده توسط واسرمن و همکاران [۱]، مستک و همکاران [۲]، استور و بریل [۳]، کنگ و آلباین [۴]، نورالسنا و همکاران [۵] و کاظم‌زاده و همکاران [۶] اشاره کرد که اغلب در برگیرنده پروفایل‌های خطی هستند. آلت [۷] از دو مرحله یا فاز تحت عنوان فاز I و فاز II در طراحی نمودارهای کنترل نام می‌برد. در فاز I، نخست مشاهدات گذشته مورد مطالعه قرار می‌گیرند تا در مورد تحت کنترل بودن فرآیند

در زمان جمع آوری مشاهدات اطمینان حاصل شود. در صورتی که فرآیند تحت کنترل باشد آنگاه پارامترهای توزیع مشخصه یا مشخصه‌های مورد نظر برآورد و نمودار یا نمودارهای کنترل جهت کنترل فرآیند در مراحل آتی طراحی می‌شود. در فاز II، نمودارهای کنترل طراحی شده به منظور تشخیص هر گونه تغییر در پارامترهای برآورد شده استفاده می‌شود و عملکرد آنها مورد ارزیابی قرار می‌گیرد. در این مقاله تلاش می‌شود تا روش‌های پایش پروفایل‌هایی که دارای پاسخ چند رسته‌ای اسمی هستند در فاز I توسعه داده شود.

در دهه گذشته، پایش پروفایل‌ها به عنوان یکی از حوزه‌های کنترل فرآیند آماری مورد استقبال و توجه محققین و کاربران این حوزه از علم آمار قرار گرفته است. در اغلب مطالعات انجام شده، مدل رگرسیون خطی جهت توصیف پروفایل در نظر گرفته شده است. به عنوان مثال، می‌توان در فاز I به تحقیقات انجام شده توسط مستک و همکاران [۲]، استور و بریل [۳]، محمود و وودال [۸]، محمود و همکاران [۹] اشاره نمود. همچنین در بین مطالعات انجام شده در فاز II می‌توان به تحقیقات انجام شده توسط کنگ و آلباین [۴]، کیم و همکاران [۱۰]، گیوپتا و همکاران [۱۱]، زو و همکاران [۱۲] اشاره نمود.

تاریخ وصول: ۸۹/۵/۱۳

تاریخ تصویب: ۹۰/۶/۲۳

*نویسنده مسئول مقاله: دکتر رسول نورالسنا، دانشکده مهندسی صنایع، دانشگاه علم و صنعت ایران، rassoul@iust.ac.ir
حمیدرضا ایزدبخش، دانشکده مهندسی صنایع، دانشگاه علم و صنعت ایران، hizadbakhsh@iust.ac.ir

نخواهد داشت. در این میان می‌توان به فرایندهایی اشاره نمود که مشخصه کیفی مورد نظر به صورت وصفی و از نوع برنولی بوده (معیوب و نامعیوب) و این مشخصه به متغیرهای مستقل دیگری وابسته باشد. یه و همکاران [۲۲] روش‌های مبتنی بر آماره T^2 هتیلینگ معرفی شده در پایش پروفایل‌های غیرخطی را برای پایش پروفایل‌های با متغیر پاسخ غیر نرمال و حالت خاصی از آن یعنی دوتایی بکار گرفته‌اند.

متغیر پاسخ در نظر گرفته شده در تحقیق آنها، دو سطحی و مبنای پروفایل ارایه شده، مدل لجیت لجستیک است. آنها ۵ روش موجود در ادبیات موضوع برای تخمین ماتریس واریانس-کواریانس آماره T^2 را استفاده کردند و یکی از نتایج تحقیق آنها ضعف عملکرد دو روش تخمینگر شامل مینیمم حجم بیضی و مینیمم تعیین کننده کواریانس بوده است. لازم به ذکر است که ضعف این دو روش در ادبیات موضوع نیز به آن اشاره شده است. در این مقاله از رگرسیون لجستیک چند رسته‌ای اسمی جهت توصیف پروفایل مورد نظر استفاده شده است. این مدل اجازه می‌دهد تا متغیر پاسخ بیش از دو سطح داشته باشد. به عنوان مثال می‌توان از سلیقه مشتریان در انتخاب خدمات/محصولات و عادت کاربران در استفاده از خدمات اینترنتی در طول شبانه‌روز به عنوان مثال‌هایی که دارای متغیرهای پاسخ چندسطحی هستند نام برد. در بخش ۲ مدل مبنای پروفایل لجستیک چند رسته‌ای اسمی به همراه الگوریتمی برای برآورد ضرایب مدل آورده شده است. نمودارهای کنترل توسعه داده شده برای پایش پروفایل مورد نظر در بخش ۳ تشریح می‌شود. در بخش ۴ با استفاده از یک مثال عددی روش‌های توسعه داده شده مورد ارزیابی قرار می‌گیرند.

۲. مدل مبنا و برآورد ضرایب

بردار $p+1$ بُعدی $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{ip}, x_{ip+1})^T$ شامل p متغیر مستقل و عرض از مبدا که در آن $i=1, \dots, n$ نشانگر حالت‌های مستقل آزمایش است را در نظر بگیرید. بردار $\mathbf{y}_i = (y_{i1}, y_{i2}, \dots, y_{ij})^T$ متغیر پاسخ متناظر با مشاهدات است و تعداد رسته‌های اسمی متغیر پاسخ می‌باشد. مثلاً برای حالتی که متغیر پاسخ وضعیت محصولی را به صورت سالم، معیوب و یا از نوع دوباره‌کاری نشان دهد آنگاه $J=3$ و بردار $\mathbf{y}_i$ دارای توزیع چندجمله‌ای^۱ با بردار احتمالات $\boldsymbol{\pi}_i = (\pi_{i1}, \pi_{i2}, \dots, \pi_{ij})^T$ خواهد بود. بنابراین،

مثال‌های متعددی وجود دارد که در آنها کیفیت محصول یا فرایند بوسیله رابطه غیرخطی بین دو یا چند متغیر مشخص شده است. برای نمونه می‌توان به مطالعه موردی بریل [۱۳]، والکر و رایت [۱۴]، جی‌آنگ و همکاران [۱۵]، و ویلیامز و همکاران [۱۶ و ۱۷] اشاره کرد. نظر به این که پروفایل‌های غیرخطی از پیچیدگی نسبتاً بیشتری برخوردار هستند تحقیقات کمتری در راستای توسعه رویکردهایی جهت پایش اینگونه پروفایل‌ها انجام شده است. جین و شی [۸] از مویک برای ارزیابی عملکرد فرایند همراه با نمودار کنترل T^2 استفاده کرده‌اند. بریل [۱۳] آماره T^2 آنها را برای تشخیص سریع تغییرات فرایند یا محصول مناسب ندانست. برای رفع این مشکل، جی‌آنگ و همکاران [۱۵] از یک آماره T^2 مبتنی بر زیر مجموعه‌ای از مشاهدات حاصل از تبدیل مویک استفاده نمودند. ویلیامز و همکاران [۱۶ و ۱۷] نمودارهای کنترل چند متغیره T^2 را برای پایش پروفایل‌های غیرخطی ارائه نمودند. واکر و رایت [۱۴] رویکردی ناپارامتری برای مقایسه پروفایل‌ها با استفاده از مدل‌های جمع پذیر ارائه داده‌اند.

این قبیل مدل‌ها، فرم تابعی خاصی ندارند و همچنین پارامتری برای برآورد در آنها وجود ندارد. روش‌های رگرسیون ناپارامتریک، انعطاف پذیری بیشتری را در مدل سازی پروفایل‌ها فراهم می‌سازد ولی بدیهی است که توان آزمون این روش‌ها به مراتب کمتر از رویکرد های پارامتریک است. از جمله این روش‌ها می‌توان به روش‌های مختلف هموارسازی اشاره نمود. یک عیب روش‌های هموارسازی ناپارامتریک این است که تفسیر منحنی ناپارامتریک برآورد شده ممکن است دشوار باشد و کاربر را در پی بردن به عوامل غیرتصادفی که منجر به حالت خارج از کنترل می‌گردد دچار مشکل سازد. دینگ و همکاران [۱۹] از انواع مختلفی از روشهای کاهش داده‌ها از جمله آنالیز اجزای اصلی و آنالیز اجزای مستقل، برای پایش داده‌هایی که در قالب یک پروفایل غیرخطی ظاهر می‌شوند استفاده می‌کنند. جنسن [۲۰] از مدل ترکیبی غیرخطی برای پایش پروفایل‌های غیرخطی به منظور با هدف لحاظ نمودن ساختار همبستگی استفاده کرد. واقفی و همکاران [۲۱] دو رویکرد برای پایش پروفایل‌های غیرخطی در فاز II ارائه نموده‌اند. رویکرد اول شامل روش‌های نمودار کنترل با بکارگیری برآوردهای پارامتریک مدل رگرسیون به منظور پایش پروفایل‌های غیرخطی و رویکرد دوم بر اساس اندازه انحراف از منحنی مرجع بوده است.

در تمامی روش‌های فوق، فرضیات رگرسیون یعنی فرض استقلال و نرمال بودن توزیع مشاهدات و همچنین ثابت بودن واریانس آنها، به عنوان فرضیات اصلی این روشها مطرح می‌گردد. اما در واقعیت موارد متعددی می‌توان یافت که در آنها این فرضیات برقرار نبوده و بنابراین روش‌های ارایه شده در پایش اینگونه پروفایل‌ها کاربرد

¹ Multinomial

$$E(y_i) = (\pi_{i1}, \pi_{i2}, \dots, \pi_{iJ})^T, \text{ var}(y_i) = (m_i \pi_{i1}(1 - \pi_{i1}), m_i \pi_{i2}(1 - \pi_{i2}), \dots, m_i \pi_{iJ}(1 - \pi_{iJ}))^T, \text{ cov}(y_{ij}, y_{ik}) = -m_i \pi_{ij} \pi_{ik}$$

$$\log\left(\frac{\pi_{ij}}{\pi_{iJ}}\right) = \sum_{k=1}^{p+1} \beta_{jk} x_{ijk}, j = 1, 2, \dots, J-1 \quad (۱)$$

در رابطه فوق $x_{ijk} \equiv x_{ijk}$ و β_{jk} پارامترهای مدل است. به سادگی می‌توان نشان داد که

$$\pi_{ij} = \frac{\exp(x_i^T \beta_j)}{1 + \sum_{j=1}^{J-1} \exp(x_i^T \beta_j)} = \frac{\exp(\eta_{ij})}{1 + \sum_{j=1}^{J-1} \exp(\eta_{ij})}, j = 1, \dots, J-1 \quad (۲)$$

که در آن

$$\eta_{ij} = x_i^T \beta_j$$

ماکزیمم لگاریتم درستنمایی مدل لوجیت چندرسته‌ای اسمی به صورت زیر بدست می‌آید:

$$l(\beta_j; y_j) = \sum_{i=1}^n \sum_{j=1}^{J-1} \sum_{k=1}^{p+1} y_{ij} \beta_{jk} x_{ijk} - \sum_{i=1}^n m_i \log\left(1 + \sum_{j=1}^{J-1} \exp\left(\sum_{k=1}^{p+1} \beta_{jk} x_{ijk}\right)\right) \quad (۳)$$

با هدف برآورد ضرایب مدل، از مدل (۳) به ترتیب دو بار به صورت

زیر نسبت به β_{jk} مشتق گرفته می‌شود:

$$\begin{aligned} \frac{\partial l}{\partial \beta_{jk}} &= \sum_{i=1}^n y_{ij} x_{ijk} - \sum_{i=1}^n m_i \frac{x_{ijk} \exp\left(\sum_{k=1}^{p+1} \beta_{jk} x_{ijk}\right)}{1 + \sum_{j=1}^{J-1} \exp\left(\sum_{k=1}^{p+1} \beta_{jk} x_{ijk}\right)} = \sum_{i=1}^n y_{ij} x_{ijk} - \sum_{i=1}^n m_i x_{ijk} \pi_{ij} \\ &= \sum_{i=1}^n x_{ijk} (y_{ij} - m_i \pi_{ij}) = \sum_{i=1}^n x_{ijk} (y_{ij} - \mu_{ij}) \end{aligned} \quad (۴)$$

برای $j \neq r$ داریم:

$$\frac{\partial^2 l}{\partial \beta_{jk} \partial \beta_{rs}} = - \sum_{i=1}^n m_i x_{ijk} \exp\left(\sum_{k=1}^{p+1} \beta_{jk} x_{ijk}\right) \frac{x_{irs} \exp\left(\sum_{k=1}^{p+1} \beta_{sk} x_{isk}\right)}{\left(1 + \sum_{j=1}^{J-1} \exp\left(\sum_{k=1}^{p+1} \beta_{jk} x_{ijk}\right)\right)^2} = - \sum_{i=1}^n m_i x_{ijk} x_{irs} \pi_{ij} \pi_{ir} \quad (۵)$$

به همین ترتیب برای $j = r$ داریم:

$$\frac{\partial l}{\partial \beta} = \tilde{X}^T (y - \mu) \quad (۷)$$

$$\frac{\partial^2 l}{\partial \beta \partial \beta^T} = - \tilde{X}^T W \tilde{X} \quad (۸)$$

$$\frac{\partial^2 l}{\partial \beta_{jk} \partial \beta_{rs}} = - \sum_{i=1}^n m_i x_{ijk} x_{irs} \pi_{ij} (1 - \pi_{ij}) \quad (۶)$$

فرم ماتریسی روابط ۴ تا ۶ به این صورت است:

که نمادهای بکار رفته عبارتند از:

۳. نمودارهای کنترل برای پایش پروفایل لجستیک

اسمی

در این بخش نمودارهای کنترل مبتنی بر T^2 هتیلینگ و نمودار مبتنی بر آزمون نسبت درستیابی ارائه می‌شود. همانطور که قبلاً اشاره شد نمودارهای مبتنی بر T^2 برای پایش پروفایل‌های از نوع مدل‌های رگرسیون خطی و غیرخطی با پاسخ‌های پیوسته و همچنین برای مدل رگرسیونی با پاسخ دوتایی مورد مطالعه و بررسی قرار گرفته شده است. در اینجا نیز به توسعه نمودارهای کنترل مبتنی بر T^2 توصیه شده برای مورد چند رسته‌ای اسمی پرداخته می‌شود. از طرف دیگر نیز به نمودار کنترل مبتنی بر آزمون نسبت درستیابی به عنوان روش دیگری در این خصوص پرداخته می‌شود. در زیر به فرضیات اصلی مورد استفاده برای طراحی نمودارها اشاره می‌گردد:

۱. نمودارهای کنترل برای فاز I طراحی شده‌اند.
۲. بر پایه مطالعه محمود و همکاران [۹] k نمونه مستقل برای ایجاد k معادله رگرسیونی نمونه در نظر گرفته شده است. در ایجاد هر معادله، از $M = \sum_{i=1}^n m_i$ نمونه استفاده شده است. i حالت‌های مختلف برای مقادیر متغیرهای مستقل است.
۳. ماتریس مقادیر متغیرهای مستقل (X) برای همه معادلات ثابت و یکسان در نظر گرفته شده است.
۴. برای معادله t ام، $t = 1, 2, \dots, k$ ، برآورد ضرایب β_t بر اساس روش آورده شده در بخش ۲ است.
۵. آماره مورد استفاده در نمودارهای کنترل مبتنی بر T^2 به صورت زیر است:

$$T_t^2 = (\hat{\beta}_t - \beta_0)^T \sum_0^{-1} (\hat{\beta}_t - \beta_0)$$

$\sum_0$ ماتریس کوواریانس $\hat{\beta}_t$ در زمان تحت کنترل بودن فرآیند به ازاء نمونه معادله t ام است.

۳-۱. نمودارهای کنترل مبتنی بر T^2

همانطور که قبلاً اشاره شد در فاز I، نیاز به برآورد β_0 و $\sum_0$ است و عملکرد نمودار کنترل به نحوه برآورد بستگی دارد. ویلیامز و همکاران [۱۷، ۱۶] چندین روش مختلف برای برآورد ماتریس کوواریانس در T^2 که برای پایش میانگین فرایند نرمال چندمتغیره استفاده شد، ارائه دادند. این روش‌ها با بکارگیری سناریوهای متعدد خروج از کنترل شامل وجود دورافتاده‌ها، تغییرات پله‌ای و تدریجی مورد ارزیابی قرار گرفته‌اند. به و

$$\tilde{X}_{n(J-1),(J-1)(p+1)} = \begin{pmatrix} X_{n,(p+1)} & 0 & \dots & 0 \\ 0 & X_{n,(p+1)} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & X_{n,(p+1)} \end{pmatrix}$$

$$W_{n(J-1),n(J-1)} = \begin{bmatrix} w_{11} & w_{12} & \dots & w_{1(J-1)} \\ w_{21} & w_{22} & \dots & w_{2(J-1)} \\ \dots & \dots & \dots & \dots \\ w_{(J-1)1} & w_{(J-1)2} & \dots & w_{(J-1)(J-1)} \end{bmatrix}$$

که w_{km} برای $1 \leq k, m \leq J-1$ یک ماتریس قطری $n \times n$ است. $\hat{I}$ آمین عنصر قطری در W_{km} به صورت زیر محاسبه می‌شود:

$$\begin{cases} -m_i \pi_{ik} (1 - \pi_{ik}) & k = m \\ -m_i \pi_{ik} \pi_{im} & k \neq m \end{cases}$$

برآورد ضرایب به روش ماکزیمم درستیابی از تساوی عبارت (۷) با بردار صفر بدست می‌آید. از آنجائیکه این دستگاه معادلات غیرخطی است لذا بایستی از روش‌های تکراری عددی استفاده شود. روش نیوتن رافسون برای حل این دستگاه در این مقاله بکار گرفته شده است. الگوریتم محاسبه شامل مراحل زیر بدست است:

گام اول: به بردار β مقدار اولیه $\beta^{(0)}$ تخصیص داده می‌شود.

گام دوم: با استفاده از $\hat{\beta}^{(1)}, \hat{\pi}^{(1)}, \hat{\mu}^{(1)}$ و $\hat{W}^{(1)}$ بدست می‌آید.

گام سوم: $\hat{z}^{(1)} = \tilde{X} \hat{\beta}^{(1)} + (\hat{W}^{(1)})^{-1} (y - \hat{\mu}^{(1)})$ محاسبه می‌شود.

گام چهارم: تخمین β با استفاده از رابطه $\hat{\beta}^{(i+1)} = (\tilde{X}' \hat{W}^{(i)} \tilde{X})^{-1} \tilde{X}' \hat{W}^{(i)} \hat{z}^{(i)}$ بهبود می‌یابد.

گام پنجم: گام‌های ۲ تا ۴ را ۱ بار تکرار می‌شود تا شرط زیر برقرار گردد:

$$\left\| \beta^{(i)} - \beta^{(i-1)} \right\| / \left\| \beta^{(i-1)} \right\| \leq \alpha$$

که $\|v\|$ فاصله اقلیدسی از بردار v و α یک ثابت به اندازه کافی کوچک مثلاً 10^{-5} است.

$$T_{II,t}^2 = (\hat{\beta}_t - \bar{\beta}) S_{II}^{-1} (\hat{\beta}_t - \bar{\beta})^T$$

که

$$S_{II} = \frac{1}{2(k-1)} \sum_{t=1}^{k-1} (\hat{\beta}_{t+1} - \hat{\beta}_t)(\hat{\beta}_{t+1} - \hat{\beta}_t)^T$$

این نمودار هنگامی که T_{II}^2 از حد بالای کنترل UCL_{II} تجاوز کند هشدار می‌دهد.

۳-۱-۳. T^2 مبتنی بر میانگین نمونه و ادغام درون پروفایلی

همانطور که در بخش ۲ آورده شد $\hat{\beta}$ ماتریس کوواریانس متقارنی معادل با $\tilde{X}^T W \tilde{X}$ دارد. برای هر t ، $\text{var}(\hat{\beta}_t)$ به صورت زیر تخمین زده می‌شود:

$$\text{var}(\hat{\beta}_t) = (\tilde{X}^T W_t \tilde{X})^{-1}$$

لذا، تخمینی از $\sum_0$ با استفاده از میانگین $\text{var}(\hat{\beta}_t)$ ها بدست می‌آید. نمودارهای کنترل برای $t=1, 2, \dots, k$ به صورت زیر می‌باشد:

$$T_{III,t}^2 = (\hat{\beta}_t - \bar{\beta})^T S_{III}^{-1} (\hat{\beta}_t - \bar{\beta})$$

که در آن

$$S_{III} = \frac{1}{k} \sum_{t=1}^k \text{var}(\hat{\beta}_t) = \frac{1}{k} \sum_{t=1}^k (\tilde{X}^T W_t \tilde{X})^{-1}$$

این نمودار هنگامی که T_{III}^2 از حد بالای کنترل UCL_{III} تجاوز نماید هشدار می‌دهد.

۳-۲. نمودارهای کنترل مبتنی بر آزمون نسبت درستنمایی

آماره‌ای که برای نمودار کنترل پیشنهادی استفاده می‌شود بر مبنای مقایسه میزان تشابه الگوی برآورد شده و الگوی مرجع است که به صورت زیر می‌باشد:

$$\lambda_t = -2 \ln \left[\frac{L(\mathbf{y}_t, \boldsymbol{\beta})}{L(\mathbf{y}_t, \mathbf{b}_t)} \right] = 2 [\ln L(\mathbf{y}_t, \mathbf{b}_t) - \ln L(\mathbf{y}_t, \boldsymbol{\beta})], \quad t = 1, 2, \dots, k \quad (9)$$

شده و مقادیر مورد انتظار است. نحوه محاسبه این مقدار برای هر کدام از حالات ذکر شده به قرار زیر است:

$$l(\mathbf{y}_t, \boldsymbol{\beta}) = \sum_{i=1}^n \sum_{j=1}^{J-1} \sum_{k=1}^{p+1} y_{ij} \beta_{jk} x_{ijk} - \sum_{i=1}^n m_i \log \left(1 + \sum_{j=1}^{J-1} \exp \left(\sum_{k=1}^{p+1} \beta_{jk} x_{ijk} \right) \right)$$

$$l(\mathbf{y}_t, \mathbf{b}_t) = \sum_{i=1}^n \sum_{j=1}^{J-1} \sum_{k=1}^{p+1} y_{ij} \hat{\beta}_{jk} x_{ijk} - \sum_{i=1}^n m_i \log \left(1 + \sum_{j=1}^{J-1} \exp \left(\sum_{k=1}^{p+1} \hat{\beta}_{jk} x_{ijk} \right) \right)$$

همکاران [۲۲] نیز این نمودارها را بر اساس سناریوهای مختلف به منظور بررسی اثربخشی در پایش پروفایل‌های با پاسخ دوتایی مورد مطالعه قرار دادند. نتیجه بررسی این بوده است که روش T^2 که در آن ماتریس کوواریانس به طریق میانگین گرفتن از برآوردهای ماتریس کوواریانس نمونه معادله‌ها بدست می‌آید در ساخت نمودارهای کنترل قوی‌تر به طور موثر عمل می‌کند. همچنین در بعضی از حالت‌های خروج از کنترل آنها نشان دادند که روش T^2 که در آن ماتریس کوواریانس از میانگین تفاضل‌های متوالی از بردار ضرایب برآورد شده حاصل می‌شود نیز موثر عمل می‌کند. نتیجه این که در این مقاله، این دو روش به همراه روش دیگری مبتنی بر T^2 مورد ارزیابی قرار می‌گیرند که در ادامه به شرح هر سه روش مبتنی بر T^2 پرداخته می‌شود.

۳-۱-۱. T^2 مبتنی بر میانگین نمونه و ماتریس کوواریانس

در این نمودار میانگین بردارهای ضرایب برآورد شده هر k نمونه معادله‌ای مبنای محاسبه ماتریس کوواریانس مورد نیاز T^2 شده است. به عبارت دیگر:

$$T_{I,t}^2 = (\hat{\beta}_t - \bar{\beta})^T S_I^{-1} (\hat{\beta}_t - \bar{\beta})$$

که $\bar{\beta} = 1/k \sum_{t=1}^k \hat{\beta}_t$ و در T^2 ،

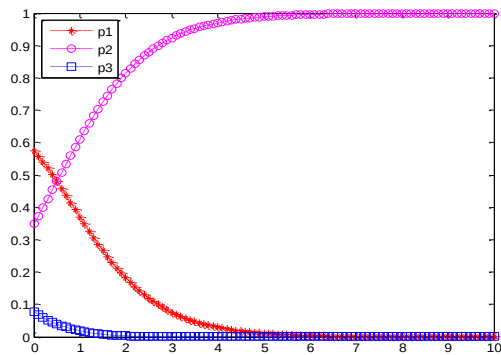
$$S_I = \frac{1}{k-1} \sum_{t=1}^k (\hat{\beta}_t - \bar{\beta})(\hat{\beta}_t - \bar{\beta})^T$$

این نمودار هنگامی که T_I^2 از حد بالای کنترل UCL_I تجاوز کند هشدار می‌دهد.

۳-۱-۲. T^2 مبتنی بر میانگین نمونه و دامنه متحرک

در این نمودار، به منظور برآورد ماتریس کوواریانس از دامنه‌های متحرک نمونه‌های مجاور استفاده شده است. T^2 برای $t=1, 2, \dots, k$ به صورت زیر محاسبه می‌شود:

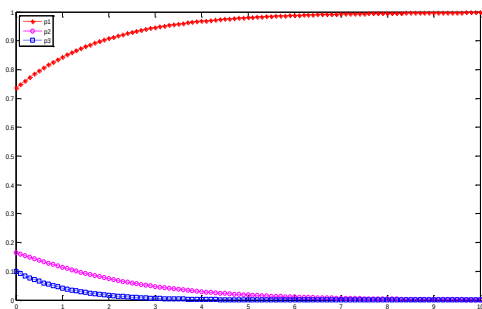
$\hat{\beta}$ بردار ضرایب مدل مرجع و $\mathbf{b}_t$ برای $t=1, 2, \dots, k$ بردار ضرایب برآورد شده از نمونه t ام و $\mathbf{y}_t$ بردار مشاهدات نمونه t ام هستند. همانطور که مشاهده می‌شود، این آماره تابعی از اختلاف بین لگاریتم تابع درستنمایی الگوی لجستیک به ازای مقادیر برآورد



شکل ۱. نمودار احتمالات رشته‌ها بر اساس مدل رگرسیونی مبنا

حال فرض نمائید که بهبودی در ساخت این قطعات صورت پذیرفته است و بر اساس داده‌های بدست آمده معادله رگرسیونی جدید به صورت زیر درآمده است. نمودار احتمالات در شکل ۲ آورده شده است.

$$\begin{cases} \log\left(\frac{\pi_{i1}}{\pi_{i3}}\right) = 2 + 1\log(x_i) \\ \log\left(\frac{\pi_{i2}}{\pi_{i3}}\right) = 0.5 + 0.5\log(x_i) \end{cases} \quad (12)$$



شکل ۲. نمودار احتمالات رشته‌ها بعد از تغییر ضرایب

همانطور که مشاهده می‌شود با تغییر ضرایب، مدل احتمالات تغییر می‌کند. به عبارت دیگر در این وضعیت احتمال سالم ماندن قطعات افزایش یافته است. لذا در ادامه با استفاده از نمودارهای کنترل پیشنهاد شده به قدرت تشخیصی آنها در انواع تغییر ضرایب مدل ارایه شده در رابطه (۱۰) پرداخته می‌شود.

۵. ارزیابی و مقایسه عملکردها

به منظور ارزیابی عملکرد نمودارهای کنترل در فاز I از شاخص مرسوم احتمال هشدار^۲ استفاده می‌شود. احتمال هشدار، احتمال

مقدار آماره آزمون نسبت درستنمایی (۹) در بهترین شرایط برابر صفر است و هر چه از آن مقدار فاصله بگیرد بیانگر این واقعیت خواهد بود که انحراف الگوی برازش شده از الگوی مورد انتظار در حال افزایش است. از آنجایی که نتایج حاصل از آماره (۹) در شرایط تحت کنترل بودن فرآیند، با تقریبی مناسب از توزیع χ^2 تبعیت می‌کنند [۲۳]، می‌توان حد کنترل بالایی بر این اساس برای نمودار مورد نظر محاسبه نمود. پیشنهاد می‌شود که این حد برابر با مقدار تعیین شده در رابطه (۱۰) باشد.

$$UCL = \chi^2_{\alpha} \quad (10)$$

$\chi^2_{(J-1)(p+1), \alpha}$ بیانگر متغیر توزیع χ^2 با $p+1$ (تعداد پارامترهای موجود در الگوی لجستیک) و J تعداد رشته است که α درصد از این توزیع را در سمت راست خود جای داده است. این نمودار زمانی هشدار می‌دهد که برآورد آماره آن از $\chi^2_{(J-1)(p+1), (1-\alpha)^{1/k}}$ بیشتر شود.

۴. مثال عددی

مثال عددی که در ادامه آورده می‌شود توسعه یافته مثال عددی استفاده شده در مقاله^۱ به و همکاران [۲۲] است. در این مثال عددی متغیر مستقل، فشار بارگذاری بر حسب psi^۱ و متغیر پاسخ قدرت قطعه فلزی در مقابل فشار است. ده سطح از محدوده x در فاصله ۲۵۰۰ تا ۴۳۰۰ psi انتخاب شده و در هر سطحی از x تعدادی بست مورد آزمایش قرار گرفته‌اند. متغیر پاسخ برای هر سطح معینی از آزمایش فشار، یک بردار سه تایی بیانگر تعداد بست‌هایی که نتیجه آزمایش آنها در آن سطح فشار سالم، معیوب و یا نیازمند دوباره‌کاری بوده است. بر اساس داده‌های گذشته، کفایت مدل لجستیک در شرایط تحت کنترل با مشخصات زیر را نشان می‌دهد. به دلیل بزرگ بودن محدوده x از $\log(x)$ به جای x استفاده شده است [۲۴].

$$\begin{cases} \log\left(\frac{\pi_{i1}}{\pi_{i3}}\right) = 2 + \log(x_i) \\ \log\left(\frac{\pi_{i2}}{\pi_{i3}}\right) = 1.5 + 2\log(x_i) \end{cases} \quad (11)$$

در معادلات فوق π_1 ، π_2 و π_3 به ترتیب احتمال سالم بودن، نیاز به دوباره کاری و معیوب بودن در نظر گرفته شده است. در شکل ۱ نمودار احتمالات وقوع سه رشته بر اساس معادلات (۱۰) آورده شده است. محور افقی $\log(x_i)$ می‌باشد. همانطور که در شکل مشاهده می‌شود در شرایط عادی با افزایش مقدار $\log(x_i)$ احتمال نیاز به دوباره کاری قطعه افزایش می‌یابد.

^۲ Signal Probability

^۱ pounds per square inch

UCL به ازاء مقدار مشخصی از m با افزایش k افزایش می‌یابد. از طرف دیگر به ازاء مقدار مشخصی از k با افزایش m کاهش می‌یابد.

جدول ۱. مقادیر شبیه‌سازی شده محاسبه UCL برای نمودارهای کنترل

T_{III}^2	T_{II}^2	T_I^2	کل نمونه معادله‌ها (k)	تعداد تکرار (m)
20.51	18.73	14.95	30	
24.22	21.78	19.41	60	30
26.13	24.11	22.52	90	
19.21	18.25	14.48	30	
22.01	20.31	18.35	60	50
23.99	22.03	20.71	90	
17.99	17.85	14.01	30	
20.20	19.62	17.48	60	100
21.46	20.75	19.37	90	

بر مبنای β_0 و $\tilde{X}$ تعریف شده در بخش 1.4، زمانی که فرایند تحت کنترل است β به طور مجانبی دارای توزیع نرمال چندمتغیره با بردار میانگین β و ماتریس کواریانس $(\tilde{X}^T W \tilde{X})^{-1}$ است [۲۵].

روشی که برای تعیین میزان تغییر در پارامترهای مدل استفاده شده است به صورت زیر می باشد [۲۲]:

$$\tilde{\beta} = \beta_0 + \Delta = \beta_0 + (\delta_1 \sigma_1, \delta_2 \sigma_2, \delta_3 \sigma_3, \delta_4 \sigma_4)^T$$

$$(\tilde{X}^T W \tilde{X})^{-1} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} & \sigma_{24} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 & \sigma_{34} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_4^2 \end{bmatrix} = \begin{bmatrix} 0.0882 & 0.0558 & 0.0739 & 0.0468 \\ 0.0558 & 0.0494 & 0.0453 & 0.0373 \\ 0.0739 & 0.0453 & 0.1279 & 0.1070 \\ 0.0468 & 0.0373 & 0.1070 & 0.1478 \end{bmatrix}$$

مقادیر k_1 را ۱، ۲، $k/10$ ، $k/5$ ، $k/3$ ، $2k/5$ و $k/2$ در نظر گرفته شده‌اند. k_1 نمونه دورافتاده در نمونه‌هایی با اندیس‌های $(2 \times i)$ ($k_1, i=1, 2, \dots$) مشخص شده‌اند. مثلاً برای $k_1=2$ ، دورافتاده‌ها عبارتند از نمونه‌های ۲ و ۴. شبیه‌سازی‌ها برای ترکیبات مختلف از m و k انجام شده است که در جدول ۲ خلاصه‌ای از نتایج احتمالات هشدار آورده شده است. برای n_{cp} های معین کوچک احتمال‌های هشدار نمودار T_I^2 و T_{II}^2 با افزایش k_1 ، افزایش و

خروج از کنترل حداقل یک نمونه از k نمونه تعریف می‌شود. هنگامی که فرایند خارج از کنترل است احتمال هشدار بزرگتر نشاندهنده توانایی بهتر نمودار کنترل در تشخیص خروج از کنترل فرایند است. اما، هنگامی که فرایند تحت کنترل است، مقدار بیشتر احتمال هشدار نشاندهنده نرخ هشدارهای اشتباهی بیشتر نمودار کنترل است.

ابتدا در این مقاله، همه نمودارها طوری تنظیم شده‌اند که خطای نوع اول آنها به طور یکسان 0.05 باشد. سپس نمودارها برای سناریوهای متفاوت اجرا شده و در نهایت نتایج حاصل مقایسه می‌شود. از آنجائیکه محاسبه توزیع نمودارهای مبتنی بر T^2 پیچیده است لذا مقدار حدود کنترل بالا از طریق شبیه‌سازی مونت کارلو تعیین می‌شود.

با توجه به مثال عددی بخش قبل، $p = 2$ و $J = 3$. تعداد تکرار در هر آزمایش (i) را ثابت و برابر m در نظر گرفته شده است.

برای هر نمودار کنترل، k نمونه مستقل را ایجاد و ماکزیمم T^2 که با $T_{\max}^2$ نشان داده می‌شود محاسبه می‌گردد. سپس $T_{\max}^2$ با حدود کنترل مقایسه می‌شود.

برای هر چهار نمودار کنترل، شبیه‌سازی‌ها برای ترکیب‌های متفاوت از مقادیر m و k به ازاء $m=30, 50, 100$ و همچنین $k=30, 60, 90$ اجرا می‌شود.

حدود کنترل بالای محاسبه شده از طریق شبیه‌سازی مونت کارلو برای سه نمودار کنترل مبتنی بر T^2 در جدول ۱ ارائه شده است. بر اساس جدول ۱، مشاهده می‌شود که حد کنترل بالا با

که در آن δ_i مقادیر ثابت است. مقدار تغییر توسط پارامتر غیر مرکزی^۱ یا n_{cp} زیر مشخص می‌شود:

$$ncp = \Delta^T \Sigma_0^{-1} \Delta$$

به منظور ارزیابی نمودارهای کنترل در حضور مشاهدات دورافتاده، k_1 نمونه از k نمونه بر اساس $\tilde{\beta}$ (ncp>0) تولید می‌شود.

¹ non-centrality parameter

افزایش ncp در بین ۴ نمودار کنترل، نمودار کنترل LRT بهترین عملکرد را به دلیل داشتن بالاترین مقدار احتمال هشدار دارد. نمودار کنترل T_{III}^2 نیز عملکرد مناسبی دارد. این نتایج با نتایج حاصل در حالت پیوسته و با پاسخ دوتایی سازگار است. نمودارهای کنترل T_I^2 و T_{II}^2 تنها در حضور یک دورافتاده عملکرد مناسبی دارند. نتیجه اینکه، عملکرد نمودارهای کنترل LRT و T_{III}^2 در این تشخیص وجود دورافتاده‌ها مناسب هستند.

سپس کاهش می‌یابند. در حالی که برای ncp‌های بزرگتر احتمال‌های هشدار این نمودارها با افزایش k_1 افزایش می‌یابد. این مطلب در خصوص نمودارهای کنترل T_{III}^2 و LRT به صورت افزایش k_1 همراه با افزایش احتمال هشدار می‌باشد. برای یک k_1 معین، انتظار می‌رود که احتمال هشدار در هر سه نمودار با افزایش ncp افزایش یابد. این موضوع برای نمودارهای کنترل T_{III}^2 و LRT به وضوح دیده می‌شود. تا بدین‌جا دو نمودار کنترل LRT و T_{III}^2 از عملکرد مناسبی برخوردار هستند. اما در زمان

جدول ۲. احتمال‌های هشدار زمانی که k_1 نمونه از k نمونه دورافتاده هستند ($m=50, k=30$)

ncp	$(\delta_1, \delta_2, \delta_3, \delta_4)$	k_1	T_I^2 (۱۴,۴۸)	T_{II}^2 (۱۸,۲۵)	T_{III}^2 (۱۸,۸)	LRT (۱۷,۳۷)
0	(۰,۰,۰,۰)	1	0.0447	0.0488	0.0524	0.0481
		2	0.0462	0.047	0.0558	0.0447
		3	0.0467	0.0492	0.052	0.0474
		6	0.0455	0.0525	0.0542	0.0514
		10	0.0494	0.0525	0.0546	0.0524
		12	0.0482	0.0487	0.05	0.0496
		15	0.0465	0.0498	0.0575	0.0474
1.4505	(۱,۱,۱,۱)	1	0.0533	0.0523	0.0646	0.0559
		2	0.0547	0.052	0.0665	0.0605
		3	0.0529	0.0508	0.0764	0.0669
		6	0.0471	0.0435	0.0865	0.0905
		10	0.0488	0.0427	0.0958	0.1188
		12	0.054	0.0475	0.092	0.1269
		15	0.046	0.0348	0.0912	0.1422
4.1013	(۱,۱,۱,۰)	1	0.065	0.0548	0.0851	0.0969
		2	0.0666	0.0665	0.1112	0.1314
		3	0.059	0.0558	0.1325	0.1579
		6	0.0487	0.0468	0.1446	0.2738
		10	0.0484	0.0372	0.1581	0.3824
		12	0.0464	0.0345	0.1547	0.4285
		15	0.0425	0.031	0.1467	0.4997
8.2612	(۱,۱,۲,۰)	1	0.1245	0.0903	0.2001	0.1998
		2	0.1139	0.0847	0.2964	0.3323
		3	0.0982	0.0747	0.3577	0.4228
		6	0.0673	0.0448	0.4241	0.6539
		10	0.0529	0.0347	0.389	0.8257
		12	0.0435	0.0297	0.3639	0.8746
		15	0.047	0.0313	0.2981	0.9246
14.6926	(۲,۲,۲,۰)	1	0.2712	0.157	0.4424	0.4296
		2	0.2072	0.1235	0.6291	0.6528
		3	0.1467	0.0907	0.7088	0.7935
		6	0.0718	0.0518	0.7879	0.9575
		10	0.051	0.034	0.7056	0.9934
		12	0.0492	0.0275	0.6487	0.9982
		15	0.0394	0.027	0.5424	0.9997
16.4051	(۲,۲,۲,۰)	1	0.3013	0.1698	0.4941	0.5268
		2	0.191	0.11	0.6787	0.7638
		3	0.1071	0.0865	0.7674	0.8852
		6	0.0462	0.0477	0.8208	0.9832
		10	0.039	0.0365	0.7448	0.9988
		12	0.0367	0.0315	0.6887	0.9999
		15	0.037	0.033	0.5976	1
24.5335	(۱,۳,۱,۰)	1	0.5385	0.2872	0.762	0.8006
		2	0.3243	0.1727	0.9204	0.9563
		3	0.1729	0.0975	0.9622	0.9931

ادامه جدول ۲. احتمال‌های هشدار زمانی که k_1 نمونه از k نمونه دورافتاده هستند ($m=50, k=30$)

ncp	$(\delta_1, \delta_2, \delta_3, \delta_4)$	k_1	T_I^2	T_{II}^2	T_{III}^2	LRT
			(۱۴,۴۸)	(۱۸,۲۵)	(۱۸,۸)	(۱۷,۳۷)
25.6329	(۲,۵۰ و ۲,۵۰)	6	0.0777	0.053	0.9852	1
		10	0.0562	0.0363	0.9635	1
		12	0.0493	0.0375	0.9255	1
		15	0.0461	0.0298	0.8562	1
		1	0.5483	0.3072	0.7823	0.8115
		2	0.2885	0.154	0.9315	0.9622
		3	0.1311	0.0847	0.9654	0.9921
		6	0.0479	0.0458	0.9853	0.9999
		10	0.039	0.0313	0.9624	1
		12	0.0404	0.0317	0.928	1
		15	0.0384	0.0315	0.8667	1
		1	0.6687	0.38	0.8746	0.886
		2	0.3418	0.1632	0.9741	0.9875
		3	0.1364	0.0858	0.9919	0.9978
		6	0.0472	0.0443	0.9974	1
30.6937	(۳ و ۳ و ۳ و ۳)	10	0.0418	0.0335	0.99	1
		12	0.042	0.0327	0.9767	1
		15	0.0383	0.0362	0.9313	1
		1	0.7886	0.4667	0.9461	0.9568
		2	0.3941	0.1907	0.9941	0.9982
		3	0.1492	0.0842	0.9985	0.9998
		6	0.0468	0.0472	1	1
		10	0.0402	0.037	0.9993	1
		12	0.0421	0.0368	0.9954	1
		15	0.0393	0.0305	0.979	1

است. برخی از نتایج حاصل به ترتیب برای ($m=30$ و $k=30$) و ($m=50, k=30$) در جدول ۳ و ۴ آورده شده است. همانگونه که ملاحظه می‌شود روش T_I^2 در تشخیص تغییرات پله‌ای بسیار ضعیف عمل نموده است. T_{II}^2 عملکرد بهتری از T_{III}^2 در تشخیص تغییرات پله‌ای برای مقادیر کوچک ncp دارد. در حالی که برای مقادیر بزرگتر ncp عملکرد نمودار کنترل T_{III}^2 در تشخیص این شیفت‌ها بهتر است. نمودار کنترل LRT برای تمامی مقادیر ncp، k و m بهترین عملکرد را داشته است.

به منظور ارزیابی نمودارهای کنترل در حالت تغییرات پله‌ای، فرض گردیده که در فرآیند از نقطه‌ای به بعد (k_2) تغییر به میزان زیر ایجاد شده است:

$$\beta_t = \begin{cases} \beta_0 & t = 1, 2, \dots, k_2 - 1 \\ \beta = \beta_0 + \Delta & t = k_2, k_2 + 1, \dots, k \end{cases}$$

در شبیه‌سازی‌ها، k_2 برابر $\frac{k}{2} + 1$ در نظر گرفته شده است. به عبارت دیگر، فرض گردیده که در نیمه دوم نمونه‌ها تغییر وجود دارد. شبیه‌سازی به ازای ترکیبات مختلفی از m و k انجام شده

جدول ۳. احتمال هشدار برای وقوع شیفت پله‌ای در نیمه از نمونه‌ها ($m=30, k=30$)

ncp	$(\delta_1, \delta_2, \delta_3, \delta_4)$	T_I^2	T_{II}^2	T_{III}^2	LRT
		(۱۴,۹۵)	(۱۸,۷۲)	(۲۰,۷۱)	(۱۷,۳۶)
0	(۰ و ۰ و ۰ و ۰)	0.05	0.0505	0.0507	0.0507
1.4505	(۱ و ۱ و ۱ و ۱)	0.0485	0.1005	0.0823	0.144
4.1013	(۱ و ۱ و ۰ و ۰)	0.0453	0.1925	0.1225	0.4947
8.2612	(۱ و ۱ و ۲ و ۲)	0.0443	0.3548	0.2515	0.9077
14.6926	(۲ و ۲ و ۳ و ۳)	0.0472	0.5215	0.4392	0.9995
16.4051	(۲ و ۲ و ۰ و ۰)	0.0453	0.5523	0.4688	1
24.5335	(۱ و ۳ و ۱ و ۳)	0.0578	0.7042	0.7202	1
25.6329	(۲, ۵۰ و ۲, ۵۰)	0.0427	0.6955	0.7275	1
30.6937	(۳ و ۳ و ۳ و ۳)	0.0455	0.742	0.8083	1
36.9114	(۳ و ۳ و ۰ و ۰)	0.052	0.7791	0.916	1

جدول ۴. احتمال هشدار برای وقوع شیفت پله‌ای در نیمی از نمونه‌ها ($m=50, k=30$)

nep	$(\delta_1, \delta_2, \delta_3, \delta_4)$	T_I^2 (۱۴,۴۸)	T_{II}^2 (۱۸,۲۵)	T_{III}^2 (۱۹,۲۱)	LRT (۱۷,۲۶)
0	(۰,۰,۰,۰)	0.0463	0.0495	0.0462	0.0495
1.4505	(۱,۱,۱,۱)	0.0456	0.0962	0.079	0.1602
4.1013	(۱,۱,۱,۰)	0.041	0.2015	0.1253	0.5125
8.2612	(۱,۲,۲,۱)	0.0452	0.3603	0.2828	0.9248
14.6926	(۲,۳,۳,۲)	0.0399	0.5435	0.4968	0.9997
16.4051	(۲,۰,۰,۲)	0.0338	0.5613	0.5603	1
24.5335	(۱,۳,۳,۱)	0.0445	0.7193	0.8287	1
25.6329	(۲,۵,۰,۲)	0.0378	0.7093	0.8401	1
30.6937	(۳,۰,۳,۰)	0.0365	0.7585	0.9093	1
36.9114	(۳,۰,۰,۳)	0.0396	0.8058	0.9755	1

نتایج حاصل برای حالت ($m=30, k=30$) در جدول ۵ آمده است. رویکرد پیشنهادی نیز در این حالت بهترین عملکرد را داشته است. اما برای سه نمودار کنترل دیگر همانگونه که در جدول ۵ ملاحظه می‌شود T_I^2 عملکرد مناسبی ندارد. عملکرد T_{II}^2 بهتر از عملکرد T_{III}^2 در تشخیص شیفت‌های تدریجی داشته است.

در نهایت به منظور ارزیابی نمودارهای کنترل سناریویی برای حالت تغییرات تدریجی آغاز شده از نمونه k_3 ($1 < k_3 < k$) در نظر گرفته شده است. این تغییر تدریجی در فرایند به صورت تغییر در بردار پارامترهای فرایند تعریف شده است.

$$\beta_t = \begin{cases} \beta_0 & t = 1, 2, \dots, k_3 - 1 \\ \tilde{\beta} = \beta_0 + \frac{t-1}{k-1} \Delta & t = k_3, k_3 + 1, \dots, k \end{cases}$$

جدول ۵. احتمال هشدار برای وقوع شیفت تدریجی آغاز شده از نمونه دوم ($m=30, k=30$)

nep	$(\delta_1, \delta_2, \delta_3, \delta_4)$	T_I^2 (14.95)	T_{II}^2 (۱۸,۷۲)	T_{III}^2 (۲۰,۷۱)	LRT (17.36)
0	(۰,۰,۰,۰)	0.0501	0.0498	0.0496	0.0533
1.4505	(۱,۱,۱,۱)	0.05	0.0745	0.0601	0.1057
4.1013	(۱,۱,۱,۰)	0.0476	0.1182	0.0764	0.3015
8.2612	(۱,۲,۲,۱)	0.0507	0.2257	0.1182	0.6648
14.6926	(۲,۳,۳,۲)	0.0552	0.3527	0.1881	0.948
16.4051	(۲,۰,۰,۲)	0.0436	0.3952	0.1881	0.9845
24.5335	(۱,۳,۳,۱)	0.0536	0.5827	0.3265	1
25.6329	(۲,۵,۰,۲)	0.0454	0.5844	0.3221	1
30.6937	(۳,۰,۳,۰)	0.0455	0.6578	0.3879	1
36.9114	(۳,۰,۰,۳)	0.0467	0.7444	0.4901	1

سازی نشان داد که نمودار کنترل پیشنهادی مبتنی بر آزمون درستنمایی بهترین عملکرد را در تمامی سناریوها دارد. همچنین نمودار کنترل T_I^2 که ماتریس کواریانس ضرایب آن با بکارگیری ضرایب برآورد شده از نمونه‌ها بدست آمده در هیچ یک از

۶. نتیجه گیری

در این مقاله نمودارهای کنترل مختلفی برای پایش پروفایل‌ها با پاسخ چندرسته‌ای اسمی در فاز I برای سناریوهای مختلف شامل نقاط دورافتاده، تغییرهای پله‌ای و تدریجی ارائه شد. نتایج شبیه

- [10] Kim, K., Mahmoud, M.A., Woodall, W.H., "On the Monitoring of Linear Profiles", J. Qual. Technol., Vol. 35, 2003, pp. 317–328.
- [11] Gupta, S., Montgomery, D.C., Woodall, W.H., "Performance Evaluation of Two Methods for Online Monitoring of Linear Calibration Profiles", Int. J. Product. Res., Vol. 44, 2006, pp. 1927–1942.
- [12] Zou, C., Zhang, Y., Wang, Z., "Control Chart Based on Change-Point Model for Monitoring Linear Profiles", IIE Transactions, Vol. 38, 2006, pp. 1093–1103.
- [13] Brill, R.V., "A Case Study for Control Charting a Product Quality Measure that is a Continuous Function Over Time", Presentation at the 45th Annual Fall Technical Conference, Toronto, Ontario, 2001.
- [14] Walker, E., Wright, S.P., "Comparing Curves Using Additive Models", J. Qual. Technol., Vol. 34, 2002, pp. 118–129.
- [15] Jeong, M.K., Lu, J.C., Wang, N., "Wavelet-Based SPC Procedure for Complicated Functional Data", Int. J. Product. Res., 2006, Vol. 44, pp. 729–744.
- [16] Williams, J.D., Woodall, W.H., Birch, J.B., "Statistical Monitoring of Nonlinear Product and Process Quality Profiles", Quality and Reliability Engineering International, Vol. 23, 2007, pp. 925–941.
- [17] Williams, J.D., Birch, J.B., Woodall, W.H., Ferry, N.M., "Statistical Monitoring of Heteroscedastic Dose-Response Profiles From High Throughput Screening", Journal of Agricultural, Biological, and Environmental Statistics, Vol. 12, 2007, pp. 216–235.
- [18] Jin, J., Shi, J., "Automatic Feature Extraction of Waveform Signals for In-Process Diagnostic Performance Improvement", Journal Intelligent Manufacturing, 2001, Vol. 12, pp. 257–268.
- [19] Ding, Y., Zeng, L., Zhou, S., "Phase I Analysis for Monitoring Nonlinear Profiles in Manufacturing Processes", Journal of Quality Technology, Vol. 38, 2006, pp. 199–216.
- [20] Jensen, W.A., Birch, J.B., "Profile Monitoring via Nonlinear Mixed Models", Journal Quality Technology, 2009, Vol. 41, pp. 18–34.
- [21] Vaghefi, A., Tajbakhsh, S.D., Noorossana, R., "Phase II Monitoring of Nonlinear Profiles", Communications in Statistics—Theory and Methods, Vol. 38, 2009, pp. 1834–1851.
- [22] Yeh, A.B., Huwang, L., Li, Y.M., "Profile Monitoring for a Binary Response", IIE Transactions, Vol. 41, 2009, pp. 931–941.

سناریوها عملکرد مناسبی نداشته است. نمودار کنترل T_{II}^2 که ماتریس کواریانس آن از تفاضلات متوالی ضرایب برآورد شده بدست آمده برای حالت شیفتهای پله‌ای به ازاء ncpهای کوچک و برای حالت شیفتهای پله‌ای تدریجی به ازاء کلیه مقادیر ncp نسبت به T_{III}^2 عملکرد بهتری داشته است.

نمودار کنترل T_{III}^2 که ماتریس کواریانس آن از میانگین ماتریس کواریانس نمونه‌ها بدست آمده عملکرد نسبتاً ضعیفی در حالت شیفتهای تدریجی داشته است. پیشنهاد می‌شود که متغیرهای پاسخ ترتیبی، بکارگیری دیگر توابع لینک مانند پروبیت و بررسی عملکرد در فاز II مورد تحقیق قرار گیرد.

مراجع

- [1] Wasserman, W., Neter, J., Kunter, H.M., *Applied Linear Statistical Models*, 4 Edition, Richard D. Irwin, Inc, Boston, MA, 1990.
- [2] Mestek, O., Pavlik, J., Suchánek, M., "Multivariate Control Charts: Control Charts for Calibration Curves", Fresenius J. Analyt. Chem, Vol. 350, 1994, pp. 344–351.
- [3] Stover, F.S., Brill, R.V., "Statistical Quality Control Applied to Ion Chromatography Calibrations", J. Chromatography A, Vol. 804, 1998, pp. 37–43.
- [4] Kang, L., Albin, S.L., "On-line Monitoring When the Process Yields a Linear Profile", J. Qual. Technol., Vol. 32, 2000, pp. 418–426.
- [5] Noorossana, R., Eyvazian, M., Amiri, A., Mahmoud, A.M., "Statistical Monitoring of Multivariate Multiple Linear Regression Profiles in Phase I with Calibration Application", Quality and Reliability Engineering International, Vol. 26, No. 3, 2010, pp. 291 – 303.
- [6] Kazemzadeh, R.B., Noorossana, R., Amiri, A., "Monitoring Polynomial Profiles in Quality Control Applications", International Journal Advanced Manufacturing Technology, Vol. 42, 2009, pp. 703–712.
- [7] Alt, F.B., "Multivariate Quality Control in Encyclopedia of Statistical Sciences", John Wiley & Sons, New York, NY, 1985.
- [8] Mahmoud, M.A., Woodall, W.H., "Phase I Monitoring of Linear Profiles with Calibration Application", Technometrics, Vol. 46, 2004, pp. 380–391.
- [9] Mahmoud, M.A., Parker, P.A., Woodall, W.H., Hawkins, D.M., "A Change Point Method for Linear Profile Data", Quality and Reliability Engineering International, Vol. 23, No. 2, 2007, pp. 247–268.

- [23] Agresti, A., *Categorical Data Analysis*, John Wiley & Sons, Inc, 2 Edition, 2002.
- [24] Finney, D.J., *Probit Analysis*, Cambridge University Press, Cambridge, UK, 1950.
- [25] McCullagh, P., Nelder, J.A., *Generalized Linear Models*, 2 ed, Chapman & Hall, London, UK, 2 Edition, 1992.

