کنترل مشخصه های وصفی وقتی همبستگی مارکوفی وجود دارد

محسن خسروی[۱]

چکیده

امروزه فرایندهای تولیدی از کیفیت بالایی برخوردارند. بویژه در صنعت الکترونیک و یا در صنایع دارویی کیفیت اقلام تولید شده بسیار بالاست و اغلب تولید به صورت خودکار صورت می پذیرد. به دلیل اینکه یک کالا باید دارای چندین استاندارد باشد، آنها را به صورت سالم و معیوب یا هماهنگ و ناهماهنگ دسته بندی می کنند. اگر یک کالا دارای تمام این استانداردها باشد سالم و در غیر این صورت معیوب به حساب می آید. در این گونه محیطها در اغلب موارد یک همبستگی مارکوفی بین واحدهای تولید شده دیده می شود و تکنیکهای کنترل کیفیت آماری استاندارد کاربرد چندانی ندارند. در این مقاله ابتدا یک توزیع را، تحت عنوان توزیع دوجمله ای منفی تعمیم یافته، به کمک زنجیرهای مارکوف دو حالتی مرتبهٔ اول مانا به دست آورده، سپس کاربرد آن را در موارد فوق نشان می دهیم.

۱ مقدمه

در بسیاری از موارد، آزمایشهای برنولی را مستقل فرض می کنیم در حالی که در عمل چنین نیست. در زمینه های مختلف یک همبستگی بین این آزمایشها وجود دارد و این همبستگی در زمینهٔ کنترل کیفیت آماری به وفور گزارش شده است. در فرایند تولید، کالاها ممکن است به صورت هماهنگ و ناهماهنگ یا سالم و معیوب دسته بندی شوند. در بسیاری از موارد بویژه هنگامی که تولید به صورت خودکار است، فرض الگوی مستقل و همتوزیع (i.i.d.) برنولی چندان واقعی به نظر نمی رسد و یک همبستگی پیاپی بین واحد های تولید شده وجود دارد. چنین فرایندهایی را می توان به وسیله یک زنجیر مارکوف دو حالتی مرتبهٔ اول نمایش داد.

در این مقاله ابتدا توزیع دوجمله ای منفی تعمیم یافته را به کمک زنجیرهای مارکوف دو حالتی مرتبهٔ اول مانا به دست آورده ، سپس کاربرد آن را در موارد فوق نشان می دهیم.

---

[۱] دانشجوی کارشناسی ارشد، گروه آمار دانشگاه شیراز

۲ توزیع دوجمله ای منفی تعمیم یافته

در یک واحد تولیدی با همبستگی پیاپی، احتمال یافتن یک واحد معیوب یا سالم به وضعیت واحد تولید شده قبلی بستگی دارد. به عبارت دیگر دو احتمال زیر را خواهیم داشت:

(سالم بودن کالای قبلی | معیوب بودن کالا)$a = P$

(معیوب بودن کالای قبلی | سالم بودن کالا)$b = P$

این فرایند را می توان به وسیلهٔ یک زنجیر مارکوف دو حالتی مرتبهٔ اول با ماتریس احتمال تغییر وضعیت زیر نمایش داد:

$$P = \begin{array}{c c} & \begin{array}{cc} S & F \end{array} \\ \begin{array}{c} S \\ F \end{array} & \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix} \end{array} \qquad (1)$$

که S (موفقیت) نشان دهندهٔ یک واحد سالم و F (شکست) نشان دهندهٔ یک واحد معیوب است. اگر در زنجیر مارکوف دو حالتی فوق، تعداد تغییر وضعیتهای لازم برای به دست آوردن k رخداد از یک پیشامد مورد نظر باشد، با توزیع دوجمله ای منفی تعمیم یافته سروکار خواهیم داشت.

یک روش کنترل مشترک برای فرایندهایی با کیفیت بالا بر اساس آمارهٔ X_1 است، که X_1 عبارت است از تعداد واحدهای مشاهده شده تا یافتن اولین واحد معیوب. اگر X_j تعداد واحدهای سالم بین (j-1) امین و j امین واحد معیوب باشد آنگاه $S_k = \sum_{j=1}^{k} X_j$ دارای توزیع دوجمله ای منفی تعمیم یافته است. بورک[2] (۱۹۹۱) در حالتی که آزمایشها مستقل باشند، یعنی $d = 1 - a - b = 0$، از S_2 به عنوان یک آمارهٔ کنترل استفاده کرده است.

فرض کنید یک دنباله آزمایش برنولی از یک زنجیر مارکوف با ماتریس احتمال تغییر وضعیت (۱) آمده و این زنجیر در آزمایش صفرم، یعنی در آغاز در حالت شکست باشد. اگر متغیر تصادفی Y را برابر تعداد موفقیتها بین آزمایش صفرم و اولین شکست در نظر بگیریم روابط زیر را خواهیم داشت:

$$\begin{cases} P(Y=0) = 1-b \\ P(Y=i) = b(1-a)^{i-1}a & ; i \geq 1 \end{cases}$$

از آنجایی که X، تعداد کل آزمایشهای لازم تا رسیدن به اولین شکست است، $X = Y + 1$. تابع احتمال X به صورت زیر است:

$$P(X=i) = \begin{cases} 1-b & ; i = 1 \\ ab(1-a)^{i-2} & ; i \geq 2 \end{cases} \qquad (2)$$

و تابع مولد احتمال آن برابر است با:

$$G_X(s) = \frac{s(1-b-ds)}{1-cs} \quad ; \quad c = 1-a \qquad (3)$$

میانگین و واریانس X عبارت اند از:

$$\mu = \frac{a+b}{a} = \frac{1}{p}$$

$$\sigma^2 = \frac{b(2-a-b)}{a^2} = \left(\frac{1-p}{p^2}\right)\left(\frac{1+d}{1-d}\right) \qquad (4)$$

وقتی که همبستگی پیاپی یعنی d برابر صفر باشد، توزیع فوق به توزیع هندسی تبدیل می شود. اگر $d > 0$، واریانس X از واریانس توزیع هندسی بزرگتر است و اگر $d < 0$ عکس آن صادق است.

اکنون فرض کنید که X_1 تعداد واحدهای بازرسی شده تا یافتن اولین واحد معیوب و X_2 تعداد این واحدها بعد از اولین واحد معیوب تا یافتن دومین واحد معیوب باشد. اگر $d = 0$، $Z = X_1 + X_2$ دارای توزیع دوجمله ای منفی است، اما اگر فرایند دارای یک همبستگی مارکوفی با ماتریس احتمال تغییر وضعیت (۱) باشد، تابع احتمال Z به صورت زیر است:

$$P(Z=n) = \qquad (5)$$

$$\begin{cases} (1-b)^2 & ; n = 2 \\ 2ab(1-b)c^{n-3} + (n-3)a^2b^2c^{n-4} & ; n \geq 3 \end{cases}$$

و تابع توزیع آن

---

[2] Bourke

$$F_Z(z) = \quad (۶)$$

$$\begin{cases} (۱-b)^۲ & ; \quad z=۲ \\ ۱+[(z-۱)b-۲]bc^{z-۲}-(z-۲)b^۲c^{z-۳} & ; \quad z \geq ۳ \end{cases}$$

که $c=۱-\alpha$.

از آنجایی که $X_۱$ و $X_۲$ مستقل اند (حتی اگر دنبالهٔ برنولی مستقل نباشد) تابع مولد احتمال Z به صورت زیر است:

$$G_Z(s) = \frac{s^۲(۱-b-ds)^۲}{(۱-cs)^۲} \; ; \; c=۱-a.$$

اکنون فرض کنید زنجیر مارکوف در حالت شکست آغاز شود. تابع احتمال تعداد آزمایشهای لازم تا یافتن k امین شکست را در نظر بگیرید. فرض کنید S_k متغیر تصادفی مربوط به تعداد این آزمایشها باشد. یعنی $S_k = \sum_{j=۱}^{k} X_j$. به روش مشابه استدلال بالا تابع مولد احتمال S_k به صورت زیر است:

$$G_k(s) = \frac{s^k(۱-b-ds)^k}{(۱-cs)^k}$$

میانگین و واریانس S_k عبارت اند از:

$$\mu = \frac{k(a+b)}{a} \quad , \quad \mu_۲ = \frac{kb(۲-a-b)}{a^۲}$$

در توزیع دوجمله ای منفی تعمیم یافته سه پارامتر k، a و b وجود دارند. از آن جایی که تعداد k داده می شود، فقط باید مقادیر a و b برآورد شوند. برآوردهای حداکثر درستنمائی a و b در سال ۱۹۸۷ توسط ساماپاس کومار[۳] و راجارشی[۴] به صورت زیر به دست آمده اند:

$$\hat{a} = \frac{N_{۰۱}}{N_{۰۱}+N_{۰۰}} \quad , \quad \hat{b} = \frac{N_{۱۰}}{N_{۱۰}+N_{۱۱}}$$

که در آن N_{ij} تعداد تغییر وضعیتها از حالت i به حالت j است. حالت موفقیت را با صفر و حالت شکست را با یک نمایش می دهیم.

۳ فرایندهای تولیدی با کیفیت بالا

وقتی که p ، نسبت اقلام معیوب، کوچک است و در مقیاس PPM (Parts-per-million) مطرح می شود، حتی اگر p معلوم باشد، یک روش نمونه گیری مناسب، نمونه گیری دو جمله ای وارون است که بر توزیع هندسی استوار است و نمودارهای کنترل مورد استفاده را نمودارهای کنترل هندسی می نامند.

اگر نسبت اقلام معیوب p به دست آمده باشد، تعداد واحدهای مشاهده شده تا یافتن یک کالای معیوب، یک متغیر تصادفی هندسی با تابع توزیع زیر است:

$$F(x;p) = \begin{cases} ۰ & ; \quad x<۱ \\ ۱-(۱-p)^{[x]} & ; \quad x \geq ۱ \end{cases}$$

اگر یک نمودار احتمالی (α_L, α_U) مورد نظر باشد می توان حدود بالا و پائین کنترل را با استفاده از توزیع فوق به دست آورد. اگر بازرسی را تا یافتن دومین واحد معیوب ادامه دهیم، می توان یک نمودار بر اساس دو طول گردش متوالی واحدهای سالم پایه گذاری کرد. به عبارت دیگر این نمودار بر اساس $S_۲ = X_۱ + X_۲$ قرار داده شده است. $S_۲$ دارای توزیع دوجمله ای منفی زیر است:

$$P(S_۲ = x) = \begin{cases} \binom{x+۱}{x} p^۲ q^x & ; \quad x = ۰,۱,۲,\ldots \\ ۰ & ; \quad \text{سایر جاها} \end{cases}$$

و می توان حدود کنترل بالا و پایین را با استفاده از توزیع فوق و مقادیر α_L و α_U و p به دست آورد. اگر قرار دهیم $\alpha_L = \alpha_U = ۰/۰۲۵$ و $p = ۰/۰۰۰۱$ خواهیم داشت:

---

[۳] Sampath Kumar

[۴] Rajarshi

$$\alpha_L = P(S_۲ \le LCL) = \sum_{x=۰}^{LCL} \binom{x+۱}{x} p^۲ q^x$$

$$= ۱-(LCL+۱)q^{LCL} + LCLq^{LCL+۱}q = ۱-p$$

و در نتیجه UCL = ۵۵۷۱۵ و LCL = ۲٤۲۳ .

اکنون فرض کنید فرایند در واقع بر اساس ماتریس (۱)، با همبستگی پیاپی d همبسته باشد. با استفاده از رابطهٔ (٦) برای مقادیر مختلف d و $\alpha = \alpha_L + \alpha_U = ۰/۰۵$ و p = ۰/۰۰۰۱ و به کار بردن $S_۲$ به عنوان آمارهٔ آزمون، می توان حدود کنترل داده شده در جدول (۱) را به دست آورد.

جدول ۱

d	LCL	UCL
۰/۰۰۰۱-	۲٤۲۵	۵۷۱۰
۰/۰۰۰	۲٤۲۳	۵۵۷۱۵
۰/۰۰۵	۲۳۳۵	۵۵۸۹٤
۰/۰۱۰	۲۲۵۰	۵٦۰۷۵
۰/۰۵۰	۱۵۸٤	۵۷۵۷۲
۰/۱۰۰	۸۲۵	۵۹۵۸۷
۰/۱۵۰	۱۱٦	٦۱۸۷۲
۰/۱۵۵	٤۵	٦۲۰۱۳

به عنوان مثال حدود کنترل بالا و پائین را بـرای d = ۰/۰۰۰۱ به دست می آوریم. با استفاده از تابع توزیع (٦) داریم:

$$P(S_۲ < LCL) =$$
$$+[(LCL-۱)b-۲]bc^{LCL-۲} - (LCL-۲)b^۲C^{LCL-۳}$$
$$;LCL \ge ۳$$

که در آن C = ۱ - a. از طرفی $\alpha_L = \alpha_U = ۰/۰۲۵$. بنابراین

$$۰/۰۲۵ = P(S_۲ < LCL) \qquad (۷)$$

مـقـادیر a و b با استـفـاده از مـقـادیـر p و d و روابط $p = \frac{a}{a+b}$ و $d = ۱-a-b$ به صـــورت a = ۰/۰۰۰۱۰۰۰۱ و b= ۰/۹۹۹۹۹۹۹۹ به دست آمده اند. با جـایگذاری این مقـادیر در معادلـهٔ (۷) مقـدار LCL = ۲٤۲۵ به دست می آید. همچنین با جـایگذاری مـقـادیـر a و b در مـعـادلهٔ $P(S_۲ \ge LCL) = ۰/۰۲۵$ ، مـقـدار UCL به صورت UCL = ۵۵۷۱۰ به دست می آید. برای سایر مقادیـر d ، مقادیر UCL و LCL محاسبه شـده و در جـدول (۱) داده شـده انـد. لازم به ذکر است که d نمی تواند از ۰/۰۰۰۱- کمتر باشد، زیـــرا b = (۱-p)(۱-d) و در این حالت برای p = ۰/۰۰۰۱ مقـدار b از ۱ بزرگتر می شود که غیر ممکن است.

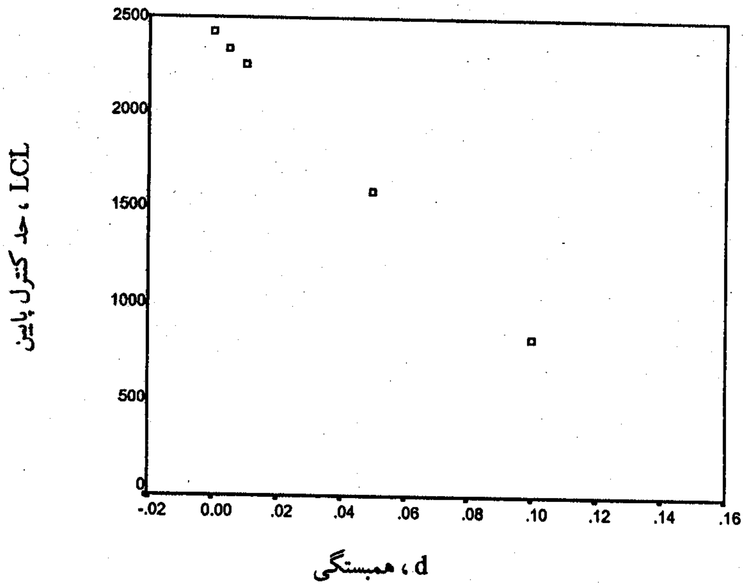
می توان ایدهٔ بورک را بـه نمودارهـای کنـترل کـه بـر اساس مجموع k طول گردش متوالی پایه گذاری شده اند، تعمیـم داد. اگر S_k را بـه صـورت $S_k = X_۱ + X_۲ + ... + X_k$ تعریـف کنیـم، واضح است که اگر فـــراینـــد بـه طـور متوالـی مستقل باشـد، S_k دارای توزیع دوجمله ای منفی است. این آماره ممکن است یک پایه برای ساختن یک نمودار میانگین متحرک از مرتبهٔ k تشکیل دهد. با این وجود براون[۵] و ودریل[٦] (۱۹۹۱) نشان دادنـد کـه یـافتن میـانگین طـول گردش (ARL) یـک نمودار مـیـانگیـــن متحـرک بـرای k > ۲ مشکل است.

---

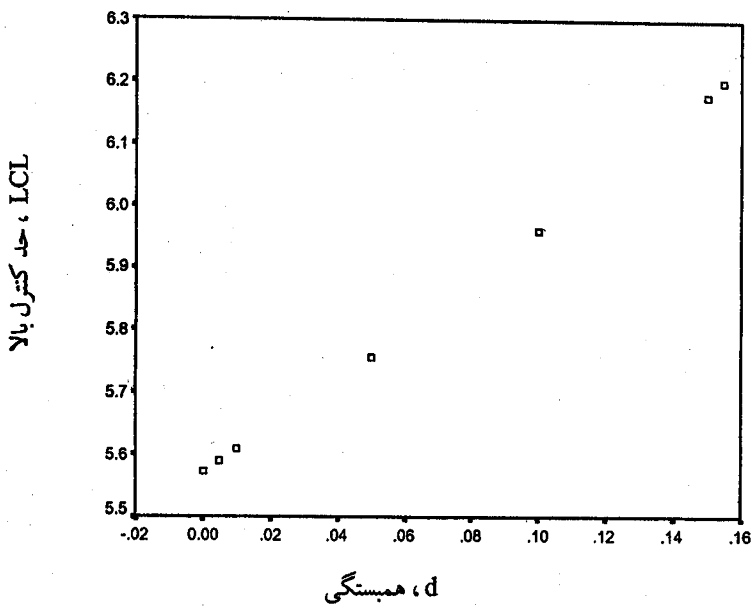
[۵] Brown

[٦] Wetherill

پراکندگی حد کنترل پایین با همبستگی پیاپی



پراکندگی حد کنترل بالا با همبستگی پیاپی



مراجع

[1] M.S. Bebbington And C. D. Lai (1998). *A Generalized Negative Binomial and Applications* . Commun. Statist-Theory . Meth. 27(0), 2515-2533.

[2] Charles P. Quesenberry (1995). *Geometric Q Chart for High Quality Process* . Journal of Quality Technology , 27(4), 304-315.

[3] William A. Stimson & Christian M. Mastrangelo. (1996). *Monitoring Serially-Dependent Processes With Atribute Data* . Journal of Quality Technology 28(3), 279-288.

[4] Wetherill, G.B. and Brown, D.W. (1991). *Statistical Process Control Theory and Practice,* Chapman and Hall, London.

[۵] خسروی، محسن، توزیع دوجمله ای منفی تعمیم یافته و کاربرد آن به کمک زنجیرهـای مـارکف دو حـالتی مرتبـهٔ اول، رسـاله کارشناسـی ارشد، گروه آمار دانشگاه شیراز ۱۳۷۹ .