

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 خود کار 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 مرتبه 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 زمینه 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 خود کار است، فرض الگوی مستقل و هم توزیع (i.i.d.) برنولی چندان واقعی به نظر نمی رسد و یک همبستگی پیاپی بین واحدهای تولید شده وجود دارد. چنین فرایندهایی را می توان به وسیله یک زنجیر مارکوف دو حالتی مرتبه اول نمایش داد. در این مقاله ابتدا توزیع دوجمله ای منفی تعمیم یافته را به کمک زنجیرهای مارکوف دو حالتی مرتبه 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:

$$a = P(\text{سالم بودن کالای قبلی} | \text{معیوب بودن کالای})$$

$$b = P(\text{معیوب بودن کالای قبلی} | \text{سالم بودن کالای})$$

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

$$P = \begin{matrix} & \begin{matrix} S & F \end{matrix} \\ \begin{matrix} S \\ F \end{matrix} & \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix} \end{matrix} \quad (1)$$

که S (موفقیت) نشان دهنده یک واحد سالم و F (شکست) نشان دهنده 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 آماره X_1 است، که عبارت است از تعداد واحدهای مشاهده شده تا یافتن اولین واحد معیوب. اگر X_1 تعداد واحدهای سالم بین (j-1) امین و زامین واحد معیوب باشد آنگاه $S_k = \sum_{j=1}^k X_j$ دارای توزیع دوجمله ای منفی تعمیم یافته است. بورک (1991) در حالتی که آزمایشها مستقل باشند، یعنی $d=1-a-b=0$ ، از S_7 به عنوان یک آماره کنترل استفاده کرده است.

فرض کنید یک دنباله آزمایش برنولی از یک زنجیر مارکوف با ماتریس احتمال تغییر وضعیت (1) 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:

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

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

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

و تابع مولد احتمال آن برابر است با:

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

میانگین و واریانس X عبارت اند از:

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

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

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 (1) باشد، تابع احتمال Z به صورت زیر است:

$$P(Z=n) = \begin{cases} (1-b)^2 & ; n=2 \\ rab(1-b)c^{n-2} + (n-2)a^2b^2c^{n-3} & ; n \geq 3 \end{cases} \quad (5)$$

تابع توزیع آن

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:

$$F(x;p) = \begin{cases} 0 & ; x < 1 \\ 1 - (1-p)^x & ; x \geq 1 \end{cases}$$

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 $S_T = X_1 + X_2$ قرار داده شده است. S_T دارای توزیع دوجمله ای منفی زیر است:

$$P(S_T = x) = \begin{cases} \binom{x+1}{x} p^2 q^x & ; x = 0, 1, 2, \dots \\ 0 & ; \text{سایر جاها} \end{cases}$$

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 $\alpha_U = \alpha_L = 0.025$ و $p = 0.001$ خواهیم داشت:

$$F_Z(z) = \quad (6)$$

$$\begin{cases} (1-b)^z & ; z=2 \\ 1 + [(z-1)b - 2]bc^{z-2} - (z-2)b^2c^{z-2} & ; z \geq 3 \end{cases}$$

که $c = 1 - \alpha$.

از آنجایی که X_1 و X_2 مستقل اند (حتی اگر دنباله برنولی مستقل نباشد) تابع مولد احتمال Z به صورت زیر است:

$$G_Z(s) = \frac{s^2(1-b-ds)^2}{(1-cs)^2} ; c = 1 - \alpha.$$

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 $S_k = \sum_{j=1}^k X_j$. به روش مشابه استدلال بالا تابع مولد احتمال S_k به صورت زیر است:

$$G_k(s) = \frac{s^k(1-b-ds)^k}{(1-cs)^k}$$

میانگین و واریانس S_k عبارت اند از:

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

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

$$\hat{a} = \frac{N_{01}}{N_{01} + N_{00}}, \quad \hat{b} = \frac{N_{10}}{N_{10} + N_{11}}$$

مقادیر a و b با استفاده از مقادیر p و d و روابط $p = \frac{a}{a+b}$ و $b = 1 - a - d$ به صورت $a = 0.00010001$ و $b = 0.99999999$ به دست آمده اند. با جایگذاری این مقادیر در معادله (۷) مقدار $LCL = 2425$ به دست می آید. همچنین با جایگذاری مقادیر a و b در معادله $P(S_r \geq LCL) = 0.025$ ، مقدار $UCL = 55710$ به دست می آید. برای سایر مقادیر d ، مقادیر LCL و UCL محاسبه شده و در جدول (۱) داده شده اند. لازم به ذکر است که d نمی تواند از $0.0001 -$ کمتر باشد، زیرا $b = (1-p)(1-d)$ و در این حالت برای $p = 0.0001$ مقدار b از ۱ بزرگتر می شود که غیر ممکن است.

می توان ایده بزرگ را به نمودارهای کنترل که بر اساس مجموع k طول گردش متوالی پایه گذاری شده اند، تعمیم داد. اگر S_k را به صورت $S_k = X_1 + X_2 + \dots + X_k$ تعریف کنیم، واضح است که اگر فرایند به طور متوالی مستقل باشد، S_k دارای توزیع دوجمله ای منفی است. این آماره ممکن است یک پایه برای ساختن یک نمودار میانگین متحرک از مرتبه k تشکیل دهد. با این وجود براون^۵ و ودریل^۶ (۱۹۹۱) نشان دادند که یافتن میانگین طول گردش (ARL) یک نمودار میانگین متحرک برای $k > 2$ مشکل است.

$$\alpha_L = P(S_r \leq LCL) = \sum_{x=0}^{LCL} \binom{x+1}{x} p^x q^x$$

$$= 1 - (LCL + 1)q^{LCL} + LCLq^{LCL+1}q = 1 - p$$

و در نتیجه $LCL = 2423$ و $UCL = 55715$.

اکنون فرض کنید فرایند در واقع بر اساس ماتریس (۱)، با همبستگی پایایی d همبسته باشد. با استفاده از رابطه (۶) برای مقادیر مختلف d و $\alpha = \alpha_L + \alpha_U = 0.05$ و $p = 0.0001$ و به کار بردن S_r به عنوان آماره آزمون، می توان حدود کنترل داده شده در جدول (۱) را به دست آورد.

جدول ۱

d	LCL	UCL
0.0001	2425	55710
0.000	2423	55715
0.005	2335	55894
0.010	2250	56075
0.050	1584	57572
0.100	825	59587
0.150	116	61872
0.155	45	62013

به عنوان مثال حدود کنترل بالا و پائین را برای $d = 0.0001$ به دست می آوریم. با استفاده از تابع توزیع (۶) داریم:

$$P(S_r < LCL) =$$

$$+ [(LCL - 1)b - 2]bc^{LCL-2} - (LCL - 2)b^2c^{LCL-2}$$

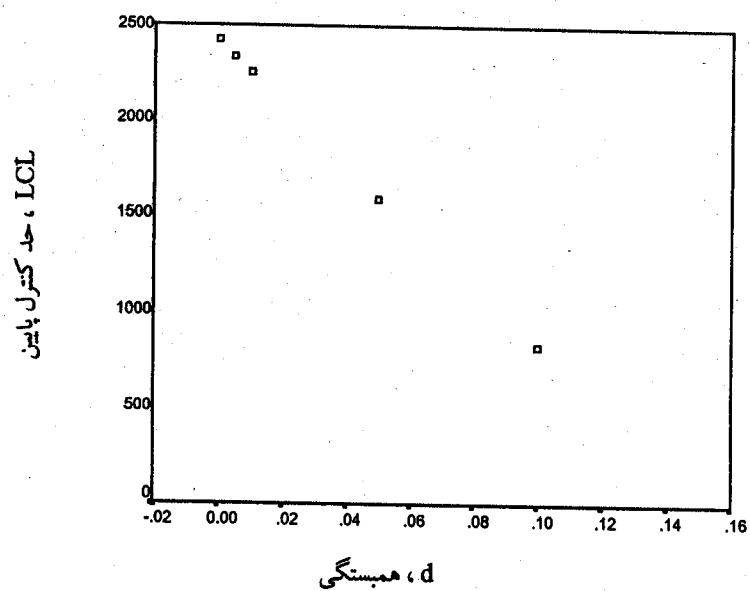
$$; LCL \geq 3$$

که در آن $C = 1 - a$ از طرفی $\alpha_L = \alpha_U = 0.025$. بنابراین

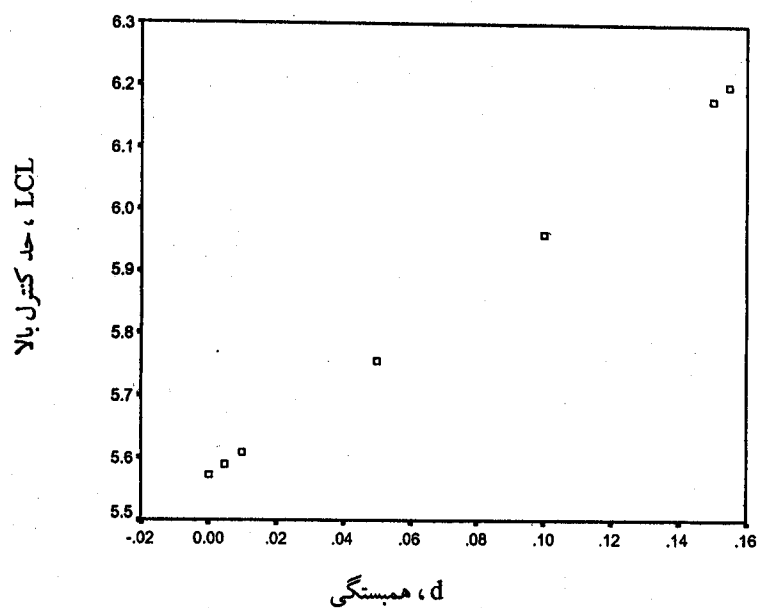
$$0.025 = P(S_r < LCL) \quad (7)$$

Brown^۵Wetherill^۶

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 Attribute 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.
- [5] خسروی، محسن، توزیع دوجمله ای منفی تعمیم یافته و کاربرد آن به کمک زنجیرهای مارکف دو حالتی مرتبه اول، رساله کارشناسی ارشد، گروه آمار دانشگاه شیراز ۱۳۷۹.
-