نشریه بین المللی مهندسی صنایع و مدیریت تولیـــد دانشگاه علم و صنعت ایران
جلد ۱۹، شماره ۴، زمستان ۱۳۸۷، صفحه ۶۱-۵۵

INTERVAL ESTIMATION PROCESS CAPABILITY INDEX

P. Bahmani
M.Sc.Degree Department of Mathematics and Computer Science Amirkabir University of Technology,
pbahmani210@yahoo.com

S.M.T. Fatemi Ghomi
Professor Department of Industrial Amirkabir University of Technology,Tehran,Iran
fatemi@aut.ac.ir

M. Nikoukar Zanjani
Assistant Professor Department of Mathematics and Computer Science Amirkabir University of Technology,Tehran,Iran

Abstract: Several sampling distribution properties of the estimator for Cpk are presented under the assumption that the data are normal, independent and identically distributed. In this paper using these assumptions, the expectation, variance and skewness are calculated by statistical methods, since the sampling distribution is weakly skewed, it is concluded that a symmetric interval estimator for Cpk might be reasonable. Also a symmetric interval estimator has been developed.

برآورد فاصله‌ای ضریب کارایی فرآیند

پریوش بهمنی، سید محمدتقی فاطمی قمی و مسعود نیکوکار زنجانی

چکیده: چند مورد از خواص توزیع نمونه‌ای برآوردگر C_{pk} عبارتند از: نرمال بودن، استقلال و همتوزیع بودن داده‌هاست. در این مقاله امید ریاضی و واریانس و چولگی C_{pk} به روش آماری محاسبه شده‌است. نظر به اینکه توزیع نمونه‌ای، چولگی ضعیفی دارد می‌توان نتیجه گرفت که برآوردگر فاصله‌ای متقارن C_{pk} منطقی است. نیز یک برآوردگر فاصله‌ای متقارن طراحی شده‌است.

واژه‌های کلیدی: حدود معنی‌داری، توزیع نرمال، توزیع نمونه‌ای، ضریب کارایی فرآیند، برآورد، حدود تلورانس، چولگی

۱. مقدمه

در اغلب مواقع ترجیح داده‌می‌شود که کارایی فرآیند بصورت یک کمیت ارائه گردد.

یکی از روشهایی که می‌توان برای بیان عملکرد یا کارایی فرآیند استفاده کرد نسبت کارایی فرآیند (ProcessCapabilityRatio) است.

این نسبت برای یک مشخصه کیفی که دارای حدود مشخصات فنی قابل قبول بالا (USL) و پایین (LSL) است، بصورت زیر محاسبه می‌شود:

$$C_p = PCR = \frac{USL - LSL}{6\sigma} \qquad (۱)$$

σ انحراف معیار فرآیند است. که در آن در عمل، انحراف معیار فرآیند نامعلوم است و از برآورد آن استفاده می‌شود. البته اگر در مطالعات کارایی فرآیند از نمودارهای کنترل متغیر استفاده شود، آنگاه می‌توان انحراف معیار فرآیند را توسط رابطه زیر برآورد نمود:

$$\hat{C}_p = P\hat{C}R = \frac{USL - LSL}{6\hat{\sigma}} \qquad (۲)$$

تاریخ وصول: ۸۳/۵/۲۰
تاریخ تصویب: ۸۴/۳/۲۹

پریوش بهمنی، دانشکده ریاضی و علوم کامپیوتر، دانشگاه صنعتی امیرکبیر، pbahmani210@yahoo.com
دکتر سید محمد تقی فاطمی قمی،دانشکده مهندسی صنایع، دانشگاه صنعتی امیرکبیر، fatemi@aut.ac.ir
دکتر مسعود نیکوکار زنجانی، دانشکده ریاضی و علوم کامپیوتر، دانشگاه صنعتی امیرکبیر، pbahmani210@yahoo.com

تعریف فوق توسط مونتگومری، کاتز و جانسون [۱و۲] ارائه شده‌است. در این منابع ضرایب دیگری نیز برای محاسبه کارایی فرآیند ارائه شده که در بین آنها ضریب C_{pk} بصورت زیر تعریف شده‌است[۳]:

$$C_{pk} = Min\left[\frac{USL-\mu}{3\sigma}, \frac{\mu-LSL}{3\sigma}\right] \quad (۳)$$

این ضریب کارایی کاربرد گسترده‌ای در صنعت دارد. این ضریب به حدود تلورانس طبیعی (3σ) و حدود معنی‌داری بستگی دارد. هدف ما در این مقاله توضیح تغییرپذیری برآوردگر $\hat{C}_{pk}$ است. توزیع آماری $\hat{C}_{pk}$ پیچیده بوده و استنباط آماری بر اساس آن مشکل است [۶ -۴]. بنابراین برآوردگر فاصله‌ای $\hat{C}_{pk}$را با استفاده از فرمول‌های امیدریاضی و واریانس و چولگی (نه بر اساس توزیع آن) به صورت $\hat{C}_{pk} \pm k\hat{\sigma}_{pk}$ محاسبه می کنیم و برای حالات $k=2$, $k=3$ محاسباتی را انجام می‌دهیم. در نهایت نتیجه می‌گیریم که تغییرات $\hat{C}_{pk}$ از نمونه‌ای به نمونه دیگر می‌تواند کاملاً بزرگ باشد حتی اگر فرآیند با تغییرات اساسی مواجه نگردد.

۲. برآوردگر C_{pk}

همان‌گونه که ملاحظه شد C_{pk} بصورت رابطه (۳) تعریف می‌شود. σ, μ میانگین و انحراف معیار فرآیند بوده که معمولاً نامعلوم می‌باشند و آنها را بر اساس یک نمونه nتایی $X_1, X_2, \dots, X_n$ با میانگین نمونه $\bar{X}$ و انحراف معیار نمونه S برآورد می‌نمائیم. آنگاه برآورد C_{pk} را بصورت زیر بدست می‌آوریم:

$$\hat{C}_{pk} = Min\left\{\frac{USL-\bar{X}}{3S}, \frac{\bar{X}-LSL}{3S}\right\}$$

$$; \ \hat{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

$$S = \sqrt{\frac{\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2}{n-1}}$$

می‌توان رابطه C_{pk} را بصورت رابطه زیر تعریف نمود:

$$C_{pk} = \frac{USL-LSL}{6\sigma} - \frac{\left|(USL-\mu)+(LSL-\mu)\right|}{6\sigma} \quad (۴)$$

$$\Rightarrow C_{pk} = \frac{USL-LSL}{6\sigma} - \frac{\left|USL+LSL-2\mu\right|}{6\sigma} \quad (۵)$$

$$\Rightarrow \hat{C}_{pk} = \frac{USL-LSL}{6S} - \frac{\left|USL+LSL-2\bar{X}\right|}{6S}$$

که وقتی n نامتناهی است، $\hat{C}_{pk}$ به C_{pk} نزدیک می شود. لذا $\hat{C}_{pk}$ یک برآوردگر سازگار برای C_{pk} است.

تحت این فرض که X_iها با میانگین μ و انحراف معیار σ مستقل و از جامعه نرمال هستند، $\hat{C}_{pk}$ در رابطه (۵) بصورت تفاضل معکوس یک متغیر با توزیع کای اسکور و مقدار قدرمطلق یک متغیر با توزیع tی غیرمرکزی با پارامتر غیرمرکزی: $\frac{\sqrt{n}\left(\mu-\left(\frac{USL-LSL}{2}\right)\right)}{\sigma}$ و درجه آزادی $(n-1)$ خواهدبود [۷-۸]. لذا محاسبه برآورد فاصله‌ای و آزمون فرض با کمک توزیع آماری $\hat{C}_{pk}$ مشکل است. از اینرو بهتر است که از عبارات دیگری مانند امید ریاضی و واریانس برای محاسبه برآورد فاصله‌ای استفاده نمائیم.

۳. امید ریاضی $\hat{C}_{pk}$

چو و اُون [۹] امیدریاضی و واریانس $\hat{C}_{pk}$ را به روشهای مختلف خود طراحی و محاسبه نمودند. مستقل از کار آنها در این مقاله امیدریاضی و واریانس $\hat{C}_{pk}$ را به روش آماری به صورت زیر محاسبه می‌نمائیم. در رابطه (۵) ملاحظه شد که:

$$\hat{C}_{pk} = \frac{USL-LSL}{6S} - \frac{\left|USL+LSL-2\bar{X}\right|}{6S}$$

با تغییر متغیرهای $a=USL-LSL$, $b=USL+LSL$ داریم:

$$\begin{aligned}\hat{C}_{pk} &= \frac{a}{6S} - \frac{\left|b-2\bar{X}\right|}{6S}\\ \Rightarrow E\left(\hat{C}_{pk}\right) &= E\left[\frac{a-\left|b-2\bar{X}\right|}{6S}\right]\\ &= \frac{1}{6}E\left(\frac{1}{S}\right)E\left(a-\left|2\bar{X}-b\right|\right)\\ &= \frac{1}{6}E\left(\frac{1}{S}\right)E\left(a-2\left|\bar{X}-\frac{b}{2}\right|\right)\end{aligned} \quad (۶)$$

از طرفی می‌دانیم:

$$\begin{aligned}X=\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)} \quad &\Rightarrow \frac{\sqrt{n-1}\,S}{\sigma} \sim \chi_{(n-1)}\\ \Rightarrow S \sim \frac{\sigma\chi_{(n-1)}}{\sqrt{n-1}} \quad &\Rightarrow \quad \frac{1}{S} \sim \frac{\sqrt{n-1}}{\sigma\chi_{(n-1)}}\end{aligned} \quad (۷)$$

نیز می‌دانیم که تابع توزیع کای‌اسکور با ν درجه آزادی عبارتست از:

$$f(x)=\frac{1}{2^{\nu/2}\,\Gamma(\nu/2)}\,x^{\nu-1}e^{-x^2/2}$$

با کمک این توزیع داریم:

$$E\left(\frac{1}{X}\right)=\int_0^\infty \frac{x^{\nu-1}\,e^{-x^2/2}}{2^{\nu/2}\,\Gamma\left(\nu/2\right)}\,\frac{1}{x}\,dx=\left(\frac{1}{2}\right)^{1/2}\frac{\Gamma\left(\frac{\nu-1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}$$

و با جایگذاری $\nu=n-1$ در رابطه امید ریاضی فوق خواهیم داشت:

$$E\left(\frac{1}{X}\right)=\left(\frac{1}{2}\right)^{1/2}\frac{\Gamma\left(\frac{n-2}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}$$

و از آن رابطه $E\left(\frac{1}{S}\right)$ حاصل می‌شود:

$$E\left(\frac{1}{S}\right)=\frac{\sqrt{n-1}}{\sigma}\cdot\left(\frac{1}{2}\right)^{1/2}\cdot\frac{\Gamma\left(\frac{n-2}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)} \qquad (۸)$$

حال برای محاسبه $E\left(\left|\bar{X}-\frac{b}{2}\right|\right)$ با استفاده از فرض نرمال بودن توزیع X_i ها داریم:

$$X_i\sim N(\mu,\sigma^2)\ \Rightarrow\ \bar{X}\sim N\left(\mu,\frac{\sigma^2}{n}\right)\ \Rightarrow\ \bar{X}-\frac{b}{2}\sim N\left(\mu-\frac{b}{2},\frac{\sigma^2}{n}\right) \qquad (۹)$$

$$E\left(\left|\bar{X}-\frac{b}{2}\right|\right)=$$
$$\sigma\cdot\sqrt{\frac{2}{n\pi}}.e^{-n\frac{\left(\mu-\frac{b}{2}\right)^2}{2\sigma^2}}+$$
$$\left(\mu-\frac{b}{2}\right)\left[1-2\Phi\left(\sqrt{n}\left(\frac{b/2-\mu}{\sigma}\right)\right)\right]$$

که در آن ϕ تابع توزیع تجمعی توزیع نرمال استاندارد است.
در رابطه (۹) متذکر می‌شویم که عبارت اول عمل جمع مربوط به حالت $\left(\bar{X}-\frac{b}{2}\right)\succ 0$ و عبارت دوم عمل جمع مربوط به حالت $\left(\bar{X}-\frac{b}{2}\right)\prec 0$ است.
حال از روابط (۶)، (۸) و (۹) داریم:

$$E\left(\hat{C}_{pk}\right)=\frac{1}{6}\left(\sqrt{\frac{n-1}{2}}\cdot\frac{\Gamma\left(\frac{n-2}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\right).$$
$$.\left[a-2\sigma\sqrt{\frac{2}{n\pi}}.e^{-n\frac{\left(\mu-b/2\right)^2}{2\sigma^2}}-2\left(\mu-b/2\right)\right]$$

$$E\left(\hat{C}_{pk}\right)=\frac{1}{6}\left(\sqrt{\frac{n-1}{2}}\cdot\frac{\Gamma\left(\frac{n-2}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\right).$$
$$.\left[\frac{a}{\sigma}-2\sqrt{\frac{2}{n\pi}}.e^{-n\frac{\left(\mu-b/2\right)^2}{2\sigma^2}}-2\frac{\left(\mu-b/2\right)}{\sigma}\left(1-2\Phi\left(\sqrt{n}\left(\frac{\frac{b}{2}-\mu}{\sigma}\right)\right)\right)\right] \qquad (۱۰)$$

بطوریکه وقتی $n\to\infty$ با کمک تقریب استرلینگ داریم:

$$\frac{\Gamma\left(\left(n-2\right)/2\right)}{\Gamma\left(\left(n-1\right)/2\right)}\sim\sqrt{2/(n-1)}$$

$$\frac{\left(\mu-b/2\right)}{\sigma}\left(1-2\phi\left(\sqrt{n}\,\frac{\left(b/2-\mu\right)}{\sigma}\right)\right)\sim\frac{\left|\mu-b/2\right|}{\sigma}$$

حال با استفاده از تغییر متغیرهای زیر بصورت:

$$c=\frac{\sqrt{n}(\mu-b/2)}{\sigma}=\frac{\sqrt{n}\left(\mu-\left(\frac{USL+LSL}{2}\right)\right)}{\sigma}\quad,\quad d=\frac{a}{\sigma}=\frac{USL-LSL}{\sigma}$$

و با جایگذاری آنها در رابطه امیدریاضی $\hat{C}_{pk}$ داریم:

$$E\left(\hat{C}_{pk}\right)=\sqrt{\frac{n-1}{2}}\,\frac{\Gamma\left((n-2)/2\right)}{\Gamma\left((n-1)/2\right)}\cdot\left(\frac{d}{6}-\sqrt{\frac{2}{\pi}}.\frac{e^{-c^2/2}}{3\sqrt{n}}-c\frac{\left(1-2\phi(-c)\right)}{3\sqrt{n}}\right) \qquad (۱۱)$$

لذا وقتی n نامتناهی است، $E\left(\hat{C}_{pk}\right)$ به C_{pk} نزدیک می شود. در نتیجه $\hat{C}_{pk}$ یک برآوردگر نا اریب C_{pk} خواهد بود.

۴. واریانس $\hat{C}_{pk}$

واریانس $\hat{C}_{pk}$ را با استفاده از فرمول کلی آماری بصورت زیر بدست می‌آوریم:

$$Var\left(\hat{C}_{pk}\right)=E\left(\hat{C}_{pk}^{\ 2}\right)-\left(E\left(\hat{C}_{pk}\right)\right)^2 \qquad (۱۲)$$

همچنین با استفاده از رابطه (۵) و با فرض اینکه $S,\bar{X}$ مستقل از یکدیگرند و با در نظر گرفتن تغییر متغیرهای $a=USL-LSL\ ,\ b=USL+LSL$، می‌توان $E\left(\hat{C}_{pk}^{\ 2}\right)$ را بصورت زیر محاسبه نمود:

$$\hat{C}_{pk}^{\,2}=\frac{\left(USL-LSL\right)^2}{36\,S^2}-$$
$$-2\left(\frac{USL-LSL}{6\,S}\right)\left(\frac{\left|USL+LSL-2\,\bar{X}\right|}{6\,S}\right)+$$
$$+\frac{\left|USL+LSL-2\,\bar{X}\right|^2}{36\,S^2}$$

$$\hat{C}_{pk}^{\,2} = \frac{a^2}{36S^2} - (a)\left(\frac{\left|b-2\bar{X}\right|}{18S^2}\right) + \frac{\left|b-2\bar{X}\right|^2}{36S^2}$$

$$\hat{C}_{pk}^{\,2} = \frac{a^2}{36S^2} - 2a\left|\left(b-2\bar{X}\right)\right|\left(\frac{1}{36S^2}\right) + \frac{\left(b-2\bar{X}\right)^2}{36S^2}$$

$$E\left(\hat{C}_{pk}^{\,2}\right) = E\left(\frac{a^2}{36S^2}\right) - 2a.E\left[\left|\left(b-2\bar{X}\right)\right|\left(\frac{1}{36S^2}\right)\right] + E\left(\frac{\left(b-2\bar{X}\right)^2}{36S^2}\right) \quad (۱۳)$$

حال با استفاده از تکنیکی که برای محاسبه امیدریاضی در قسمت قبل بکار بردیم خواهیم داشت:

$$E\left(\frac{a^2}{36S^2}\right) = \frac{a^2}{36\sigma^2}.\frac{n-1}{n-3} \quad (۱۴)$$

$$E\left(\frac{\left|b-2\bar{X}\right|}{36\,S^2}\right) =$$

$$= \frac{4a}{36\,\sigma^2}.\frac{n-1}{n-3}.\sqrt{\frac{2}{\pi}}.\frac{\sigma}{\sqrt{n}}e^{-n\left(\left(\mu-b/2\right)^2/2\sigma^2\right)}$$

$$+\frac{4a}{36\,\sigma^2}.\frac{n-1}{n-3}.\left(\mu-b/2\right)\left(1-2\Phi\left(\frac{\sqrt{n}\left(b/2-\mu\right)}{\sigma}\right)\right)$$

$$E\left(\frac{\left(b-2\bar{X}\right)^2}{36S^2}\right) = \frac{1}{36}.\frac{n-1}{n-3}\left[\frac{1}{n}+\frac{\left(\mu-b/2\right)^2}{\sigma^2}\right] \quad (۱۵)$$

حال با جایگذاری روابط (۱۱) ، (۱۳)، (۱۴) و (۱۵) در رابطه (۱۲) داریم:

$$Var\left(\hat{C}_{pk}\right) = \frac{a^2}{36\,\sigma^2}.\frac{n-1}{n-3} -$$

$$-\frac{a}{9\sigma^2}.\frac{n-1}{n-3}.\sigma\sqrt{\frac{2}{n\pi}}e^{-n\left(\left(\mu-b/2\right)^2/2\sigma^2\right)} -$$

$$-\frac{a}{9\sigma^2}.\frac{n-1}{n-3}.\left(\mu-b/2\right)\left(1-2\Phi\left(\frac{\sqrt{n}\left(b/2-\mu\right)}{\sigma}\right)\right)+$$

$$+\frac{1}{9}.\frac{n-1}{n-3}.\frac{\left(1/n\right)+\left(\mu-b/2\right)^2}{\sigma^2} -$$

$$-\frac{n-1}{72\sigma^2}\left(\frac{\Gamma\left(\left(n-2\right)/2\right)}{\Gamma\left(\left(n-1\right)/2\right)}\right)^2.$$

$$.\left\{a-2\sigma\sqrt{\frac{2}{n\pi}}e^{-\frac{n\left(\mu-b/2\right)^2}{2\sigma^2}}-2\left(\mu-\frac{b}{2}\right)\left[1-2\Phi\left(\frac{\sqrt{n}\left(\frac{b}{2}-\mu\right)}{\sigma}\right)\right]\right\}^2 \quad (۱۶)$$

از فرمولهای بدست آمده برای امید ریاضی و واریانس $\hat{C}_{pk}$ می‌توان نتیجه گرفت که در حالت کلی این عبارات تنها توابع صریحی از C_{pk} نیستند، بلکه توابعی از USL, LSL, σ, μ, n خواهند بود. لذا با تغییر متغیرهای زیر می‌توان آنها را تنها بصورت توابعی از n, c, d نوشت:

$$c=\frac{\sqrt{n}\left(\mu-b/2\right)}{\sigma}=\frac{\sqrt{n}\left(\mu-\left(USL+LSL\right)/2\right)}{\sigma} \quad , \quad d=\frac{a}{\sigma}=\frac{USL-LSL}{\sigma}$$

$$E\left(\hat{C}_{pk}\right)=\frac{1}{6}.\sqrt{\frac{n-1}{2n}}.\frac{\Gamma\left(\left(n-2\right)/2\right)}{\Gamma\left(\left(n-1\right)/2\right)}. \quad (۱۷)$$
$$\left(d\sqrt{n}-2\sqrt{\frac{2}{\pi}}.e^{-c^2/2}-2c\left(1-2\phi\left(-c\right)\right)\right)$$

$$\Rightarrow Var\left(\hat{C}_{pk}\right)=\frac{d^2}{36}.\frac{n-1}{n-3}-\frac{d}{9\sqrt{n}}.\frac{n-1}{n-3}\left[\sqrt{\frac{2}{\pi}}e^{-\left(c^2/2\right)}+c\left(1-2\phi\left(-c\right)\right)\right]+$$
$$+\frac{1}{9}.\frac{n-1}{n\left(n-3\right)}.\left(1+c^2\right)- \quad (۱۸)$$
$$-\frac{n-1}{72n}.\left(\frac{\Gamma\left(\left(n-2\right)/2\right)}{\Gamma\left(\left(n-1\right)/2\right)}\right)^2.$$
$$\left\{d\sqrt{n}-2\sqrt{\frac{2}{\pi}}e^{-c^2/2}-2c\left[1-2\phi\left(-c\right)\right]\right\}^2$$

که در آن میانگین و واریانس نمونه فرآیند که با σ, μ نشان می‌دهیم توسط $S, \bar{X}$ برآورد می‌شوند. انحراف معیار نمونه $\hat{C}_{pk}$ که با $\hat{\sigma}_{pk}$ نشان می‌دهیم جذر واریانس نمونه $\hat{C}_{pk}$ است.

۵. فرمول‌های تقریب امید ریاضی و واریانس $\hat{C}_{pk}$

فرمولهای محاسبه شده در قسمت‌های قبل برای امید ریاضی و واریانس $\hat{C}_{pk}$ طولانی بنظر می‌رسند با این وجود می‌توانیم آنها را توسط برنامه‌های کامپیوتری برنامه‌ریزی نمائیم. البته گاهی ممکن است علاقمند باشیم که از تقریب‌های این فرمولها استفاده نمائیم. با استفاده از تکنیک چان [۴] فرمولهای تقریبی برای محاسبه امید ریاضی و واریانس $\hat{C}_{pk}$ طبق شرایط خاص زیر مطرح می‌نمائیم:

$$n\geq 25 \;,\; 0.75\leq C_{pk}\leq 4 \;,\; \left|\frac{\sqrt{n}\left(\mu-b/2\right)}{\sigma}\right|\leq 100 \;,\; \frac{USL-LSL}{\sigma}\leq 24 \quad (۱۹)$$

$$E\left(\hat{C}_{pk}\right)\cong\sqrt{\frac{n-1}{2}}\;\frac{\Gamma\left(\left(n-2\right)/2\right)}{\Gamma\left(\left(n-1\right)/2\right)}.C_{pk} \quad (۲۰)$$

$$Var\left(\hat{C}_{pk}\right) \cong (n-1)\left\{\frac{1}{n-3}-\frac{1}{2}\left(\frac{\Gamma\left((n-2)/2\right)}{\Gamma\left((n-1)/2\right)}\right)^{2}\right\}C_{pk}^{2} \qquad (۲۱)$$

یادآوری می‌شود که شرایط موجود در (۱۹) برای فرمولهای تقریبی (۲۰) و (۲۱) ضروری هستند. بطوریکه تقریب‌های حاصل تنها توابعی از $\hat{C}_{pk}$ و n هستند. هارت[۱۰] نیز به روشی دیگر فرمولهای تقریبی برای امیدریاضی و واریانس $\hat{C}_{pk}$ برای مقادیر مثبت x ارائه نموده‌است.

۶. چولگی توزیع $\hat{C}_{pk}$

می‌دانیم که میزان عدم تقارن منحنی فراوانی را چولگی می‌نامند. فرض کنیم $\bar{X}$ ، m ، M ، S ، m_3 میانگین، میانه، نما، انحراف معیار و گشتاور مرکزی سوم باشند. هر یک از فرمولهای زیر را می‌توان به عنوان معیار چولگی بکار برد:

ضریب چولگی اول پیرسن:
$$b_1 = \frac{\bar{X} - M}{S}$$

ضریب چولگی دوم پیرسن:
$$b_2 = \frac{3\left(\bar{X}-m\right)}{S} \quad ; \quad m = E\left(X\right)$$

ضریب گشتاوری و چولگی:
$$g = \frac{m_3}{S^3} \quad ; \quad m_3 = E\left(X - E\left(X\right)\right)^3$$

در فرمولهای فوق از S استفاده شده است تا ضرایب به واحد اندازه‌گیری بستگی نداشته باشند. در صورتیکه داده‌ها نسبت به میانگین متقارن باشند ضرایب فوق برابر صفر هستند. ولی باید توجه نمود که عکس این موضوع کاملاً صحت ندارد. برحسب اینکه b_1 , b_2 , g مثبت یا منفی باشند منحنی فراوانی چوله به راست یا چوله به چپ خواهد بود. حال براحتی می‌توان ضریب چولگی زیر را برای $\hat{C}_{pk}$ تعریف نمود:

$$\beta = \frac{E\left[\hat{C}_{pk} - E\left(\hat{C}_{pk}\right)\right]^3}{\hat{\sigma}_{pk}^3} \qquad (۲۲)$$

که بعنوان معیار چولگی توزیع $\hat{C}_{pk}$ است.

از رابطه (۲۲) داریم:

$$\beta = \frac{E\left(\hat{C}_{pk}^3\right) - 3E\left(\hat{C}_{pk}^2\right)E\left(\hat{C}_{pk}\right) + 2E^3\left(\hat{C}_{pk}\right)}{\sigma_{pk}^3} \qquad (۲۳)$$

$$E\left(\hat{C}_{pk}^3\right) = \frac{1}{216}E\left(\frac{1}{S^3}\right) \qquad (۲۴)$$
$$\left\{a^2.E\left[a-3\left|2\bar{X}-b\right|\right]+3a.E\left[\left(\bar{X}-2b\right)^2\right]-E\left|b-2\bar{X}\right|^3\right\}$$

با استفاده از تکنیک‌های مشابه برای محاسبه میانگین و واریانس داریم:

$$E\left(\hat{C}_{pk}^3\right) = \frac{1}{216}\left(\sqrt{\frac{n-1}{2}}\right)^3 \cdot \frac{\Gamma\left((n-4)/2\right)}{\Gamma\left((n-1)/2\right)} \cdot$$
$$.\left(d^2\left[d-6\sqrt{\frac{2}{n\pi}}.e^{-c^2/2}-6\frac{c}{\sqrt{n}}\left(1-2\phi(-c)\right)\right]+ \qquad (۲۵)$$
$$+12\left(\frac{d}{n}\right)\left[1+c^2\right]-8\sqrt{\frac{2}{n\pi}}e^{-c^2/2}.\left(\frac{1}{n}\right)\left(c^2+2\right)$$
$$-8\left(\frac{c}{n\sqrt{n}}\right)\left(3+c^2\right)\left[1-2\phi(-c)\right]\Big)$$

با جایگذاری در رابطه (۲۲) می‌توان β را محاسبه نمود. برای اندازه‌های نمونه بزرگ $n \succ 100$ ملاحظه می‌شود که توزیع $\hat{C}_{pk}$ چولگی ضعیفی دارد، به گونه‌ای که طبق روابط بدست آمده برای واریانس $\hat{C}_{pk}$ در بخش‌های قبل، ملاحظه شد که واریانس $\hat{C}_{pk}$ (رابطه(۱۶)) تابعی از n است. لذا برای n کمتر از ۵۰ ضریب تغییرات بیشتری دارد ولی وقتی برای n مقداری بزرگتر منظور شود، تغییرات $\hat{C}_{pk}$ رفته رفته کاهش می‌یابد.

۷. برآوردگر فاصله‌ای $\hat{C}_{pk}$

بر اساس خواص $\hat{C}_{pk}$ از قبیل نا اریب بودن و چولگی ضعیف توزیع آن می‌توان یک برآوردگر فاصله‌ای برای $\hat{C}_{pk}$ بصورت زیر طراحی نمود:

$$\hat{C}_{pk} \pm k\hat{\sigma}_{pk} \qquad (۲۶)$$

که در آن k یک مقدار ثابت و $\hat{\sigma}_{pk}$ برآوردگر نمونه‌ای انحراف معیار $\hat{C}_{pk}$ است.

۸. تجزیه و تحلیل و نتیجه گیری

در این بخش ابتدا برای تشریح مطالبی که بخشهای پیشین ذکر شد، مثالی آورده می‌شود. سپس نتیجه گیریهای لازم بعمل خواهد آمد. در جدول (۸-۱) برای بعضی مقادیر n و $\hat{C}_{pk}$ به ازای $k=2$ و $k=3$ برآوردهای فاصله‌ای را محاسبه می‌نمائیم و برای مقادیر جدول، مقایسه‌هایی انجام می‌دهیم.

$$USL = 3 \quad , \quad LSL = -3$$
$$\mu = -4, -3, -2, -1, 0, 1, 2, 3, 4$$
$$\sigma = 0.25, 0.5, 1, 1.5, 2$$
$$n = 25, 50, 100, 150, 200$$

با کمک این مقادیر و رابطه‌های بدست آمده در بخش‌های قبل، $\hat{C}_{pk}$ را محاسبه نموده و بر اساس آن برآوردهای فاصله‌ای را بدست می‌آوریم. نتایج حاصله در جدول (۸-۱) ارائه شده است:

جدول شماره ۱. $K=2$

$\hat{C}_{pk}$	*n=25*	*n=50*	*n=100*	*n=150*	*n=200*
0/75	(0/52,0/98)	(0/59,0/91)	(0/64,0/86)	(0/66,0/84)	(0/67,0/83)
۱	(0/69,1/31)	(0/79,1/21)	(0/86,1/14)	(0/88,1/12)	(0/9,1/10)
1/33	(0/91,1/75)	(1/05,1/61)	(1/14,1/52)	(1/17,1/49)	(1/20,1/46)
۲	(1/37,2/63)	(1/58,2/42)	(1/71,2/29)	(1/77,2/33)	(1/8,2/2)

جدول شماره ۲. $K=3$

$\hat{C}_{pk}$	*n=25*	*n=50*	*n=100*	*n=150*	*n=200*
0/75	(0/40,1/10)	(0/51,0/99)	(0/59,0/91)	(0/62,0/88)	(0/64,0/86)
۱	(0/53,1/47)	(0/68,1/32)	(0/78,1/22)	(0/82,1/18)	(0/85,1/15)
1/33	(0/71,1/95)	(0/91,1/75)	(1/04,1/62)	(1/10,1/56)	(1/13,1/53)
۲	(1/06,2/94)	(1/37,2/63)	(1/57,2/43)	(1/65,2/35)	(1/70,2/30)

برای هر ترکیب از μ,σ,n نمونه‌های تصادفی از یک توزیع نرمال تولید شده‌اند. برای هر نمونه، میانگین، انحراف معیار و برآوردهای $\hat{C}_{pk}$ و $\hat{\sigma}_{pk}$ را محاسبه می‌نمائیم. سپس بر اساس این محاسبه برآورد فاصله‌ای $\hat{C}_{pk}$ را به ازای هر دو مقدار $k=2$ و $k=3$ بدست می‌آوریم.

جدول (۸-۱) اشاره به این موضوع دارد که در نمونه‌های کوچک، غیرمرکزی بودن برآورد $\hat{C}_{pk}$ نسبتاً زیاد است. بعنوان مثال وقتی $n=50 \;\; , \;\; \hat{C}_{pk}=1$ برآورد $\hat{C}_{pk}$ می‌تواند به اندازه ۲۰٪ تا ۳۰٪ دورتر از مرکز بوده و برای $n=25$ این برآورد به اندازه ۴۰٪ به بالا یا پایین نوسان دارد. فرض کنید که C_{pk} برای بعضی فرآیندها بطور متناوب برآورد شود. از یک برآورد تا برآورد بعدی، C_{pk} مقادیر مختلفی را برطبق توزیع نمونه‌ای خواهد گرفت حتی وقتی که در واقعیت، فرآیند تغییر زیادی نداشته باشد. بنابراین تغییرات مشاهده شده برآوردهای C_{pk} ممکن است آنقدر زیاد باشد که در فرآیند بطور واقعی نشان داده نشوند.

اگر افزایش C_{pk} به دلیل وسیع بودن توزیع نمونه‌ای نادیده گرفته شود، ممکن است مشخص کردن کارایی فرآیند مشکل باشد. بنابراین این نکته برای وقتی که ارزیابی فرآیند توسط $\hat{C}_{pk}$ انجام می‌شود، ارزنده است. از این لحاظ درمی‌یابیم که اریبی برآورد C_{pk} ناچیز است، هر چند که در اکثر موارد اریبی توسط انحراف معیار نسبتاً بزرگ نمونه، تحت‌الشعاع قرار می‌گیرد. محاسبات انجام شده دلالت بر این موضوع دارد که توزیع نمونه‌ای $\hat{C}_{pk}$ چولگی اندکی دارد.

این موضوع موجب می‌شود که برآوردگر فاصله‌ای متقارنی که در نظر گرفته شده است، منطقی باشد.

بنابراین C_{pk} چه در بالا یا پایین فاصله واقع شود، مشکلی ایجاد نمی‌کند. ضریب کارایی فرآیند C_{pk} کاربردهای بسیار وسیعی در صنعت دارد. ملاحظه شد که برآوردگر فاصله‌ای این ضریب برای اندازه‌های نمونه کوچک بسیار عریض است. برای وقتی که این ضریب بعنوان معیاری برای اندازه‌گیری میزان تغییرات فرآیند، استفاده می‌شود، برآورد فاصله‌ای اهمیت شایسته‌ای خواهدیافت.

مراجع

[1] Montgomery, D.C., Friedman, D.J., “*Statistical Process in a Control in a Computer Integrated Manufacturing Environment”*, in Statistical Process Control in Automated Manufacturing, Edited by J.B. Keats and N.F. Hubele, Marcel Dekker, Inc., New York, 1989.

[2] Kots, S., Johnson, N.L., *Process Capability Indices*, Chapman & Hall, London, 1993.

[3] Kane, V.E., *Process Capability Indices*, Journal of Quality Technhlogy 18, 1986, PP.41-52.

[4] Chan, L.K., Cheng, S.W., Spiring, F.A., “*A New Measure of Process Capability”*, Cpm, Journal of Quality Technology 20, 1988, PP.162-175.

[5] Spiring, F.A., ASQC Quality Congress Transactions-Toronto, 1989.

[6] Gunter, B., *The Use and Abuse of Cpk*, Quality Progress, Part 3 May, 1989, PP.79-80. .

[7] Pearn, W.L., Kots, S., Johnson, N.L., “*Distributional and Inferential Properties of Process Capability*

Indices", Journal of Quality Technhlogy, 24(4), 1992, PP. 216-233.

[8] Levinson, W.A., *Exact Confidence Limits for Process Capabilities*, Quality Engine., 9(3), 1997, PP. 511-518.

[9] Y.M., Chou, D.B., Owen, "*On the Distribution of the Estimated Process Capability Indices*". Communications in Statistics-Theory and Methods, 18(12), 1989, PP. 549-4560.

[10] Hart, R.G., "*A Close Approximation Related to the Error Function. Mathematics of Computation"*, 1966, PP. 600-602.

[11] Lancaster, H.O., Chi Distribution. Encyclopedia of Statistical Science,Vol.1.John Wiley & Sons, 1982.